198421

Assessing Technology, Methods, and Information for Committing and Combating Cyber Crime

A study sponsored by the National Institute of Justice under
grant number 2000-9614-NY-IJ

February 4, 2002

The Computer Forensics Research & Development Center (CFRDC)
At Utica College
&
WetStone Technologies, Inc.

Dr. Gary R. Gordon – Director, CFRDC at Utica College
Chet D. Hosmer – President, WetStone Technologies, Inc.
Christine Siedsma – Project Coordinator, CFRDC at Utica College
Dr. Don Rebovich – Associate Professor, Economic Crime Programs, Utica College

Supported under Award number 2000-LT-BX-K002 from the Office of Justice Programs, US Department of Justice. Points of view in this document are those of the authors and do not necessarily represent the official position of the U.S. Department of Justice

Acknowledgements

This report is the work of many individuals. The staff of the Computer Forensics Research & Development Center at Utica College, Christine Siedsma and Matt Ward, contributed many hours of research and analysis. Dr. Don Rebovich, Professor of Economic Crime Programs at Utica College, provided incisive advice on the research methodology and thoughtful editorial support. Thanks to the WetStone Technologies, Inc. staff who contributed to the effort: Chris Hyde and Todd Grant.

We wish to express our thanks to the Northeast Law Enforcement and Corrections Technology Center (NLECTC) staff: John Ritz, Fred Demma, and Jim Riccardi.

A special thanks goes to the Air Force Research Lab Information Directorate for funding several earlier efforts that provided the foundation for this work. We especially owe a great deal of gratitude to Joe Giordano, Technical Director.

We would like to thank NIJ for supporting this project, and specifically Trent DePersia and Amon Young for their direction and patience.

Most importantly, we owe a great deal to the numerous law enforcement personnel who provided comments, suggestions, and completed the survey. Thank you.

Gary R. Gordon
Chester D. Hosmer

Table of Contents

Introduction

Computer Crime: New Investigative Needs for an Emerging Crime Area

As we move forward into the 21st century, technological innovations have paved the way for us to experience new and wonderful conveniences in the how we are educated, the way we shop, how we are entertained and the manner in which we do business. Our day-to-day lives have been forever changed thanks to rapid advances made in the field of computer technology. These changes allow us to communicate over great distances in an instant and permit us, almost effortlessly, to gather and organize large amounts of information, tasks that could, otherwise, prove unwieldy and expensive. The technological treasures that have improved the quality of our lives, however, can reasonably be viewed as a doubled-edged sword. While computer technology has opened doors to enhanced conveniences for many, this same technology has also opened new doors for criminals.

Businesses that have grown to rely upon computerization to collect and assemble sensitive information on their critical resources now face the daunting, and costly, task of protecting this information from those who would seek illegal access to it. Criminals can now easily encrypt information representing evidence of their criminal acts, store the information and even transmit it with little fear of detection by law enforcement. Due to the extraordinary impact of the Internet, a computer crime scene can now span from the geographical point of the victimization (e.g., the victim's personal computer) to any other point on the planet, further complicating criminal investigative efforts. In effect, computer technology has dramatically altered the criminal justice terrain such that enterprising and opportunistic criminals have consciously turned to the computer to commit their illegal acts in situations in which the computer serves as the instrument of the crime, the means by which the crime is committed, as well as in cases in which the victim's computer, or computer system, is the target, or objective, of the act. And, as stated above, the presence of new computer technology aids cyber criminals in situations in which the computer's role is incidental to the crime; situations in which the computer is used to house and protect information that is evidence tying the offender to criminal acts. A commonality among these types of crimes is that the offender, to a great degree, depends upon the lack of technological skills of law enforcement to successfully commit the offenses and escape undetected. Based upon what empirical evidence has been available on self-assessed skills of investigators in this area, computer criminals would have good reason to feel some confidence in their chances to evade detection of their crimes. [1]

The goal of this report is to provide key insights to the law enforcement community on how to upgrade basic abilities to effectively investigate computer crimes. This report is

[1] Stambaugh, H., et. al, Electronic Crime Needs Assessment for State and Local Law Enforcement, National Institute of Justice Report, Washington, DC: U.S. Department of Justice, March 2001.

designed to reduce the "skill distance" between what computer criminals have learned to successfully commit their crimes and what law enforcers need to know to successfully bring these offenders to justice. By presenting this information in a clear, structured form, we believe great inroads can be made to gain a competitive edge over those who would misuse technology for criminal gain. The information contained in this report serves as a valuable guide to computer crime investigators. Properly implemented, the information should prove instrumental in controlling and preventing the highly damaging crimes committed against large portions of the general public and business community, crimes that, not long ago, would have been impossible to achieve with the ease with which it they can be achieved today.

Responding to a Growing Crime Problem For the 21st Century

Back in the 1960s, the term "computer" would bring to mind images of large, bulky mainframes, machines whose inner workings were, for many, cloaked in mystery. Only select parts of our population had direct access to computers, building the mystical aura surrounding computers, what they did and the type of knowledge needed to operate them. With IBM's introduction of its stand-alone "personal computer" in 1981, some of the layers of mystery about computers had been peeled away exposing many to the rewards of quick data access and manipulation that, up to that time, had been realized by few. Today it is estimated that 53.7 million households have personal computers, over 50% of the nation's households, and that the demographics of owners are finally beginning to reflect the overall demographics of the general population of the U.S. The lure of the Internet has enticed over 100 million in the U.S. to go online in year 2000 to join a world wide communications network that few envisioned when the Arpanet, the Internet's predecessor, was developed in the 1960s.[2] Likewise, few, at that time, could ever anticipate the opportunities computers, the Internet and its vast ocean of users would offer to technologically savvy criminals.

The process of criminalization of human behavior judged to be harmful to the public is typically one that builds slowly in common law jurisdictions. Momentum gained through problem identification and pressures exerted by special interest groups can easily span decades before undesirable actions are classified as "crime". In some instances, this process is accelerated through the occurrence of certain "catalyst events" that capture the attention of the public and the attention of lawmakers.

In the case of computer crime, legislators grew increasingly attentive in the 1980s as businesses became more dependent upon computerization and as catalyst event cases exposed significant vulnerabilities to computer crime violations. Cases like the Ian Murphy ("Captain Zap") invasion of White House switchboards to hack into classified military files underscored the seriousness of computer crimes and, thus, helped speed along the enactment of the Computer Fraud and Abuse Act of 1986 to replace laws that proved to be inadequate in addressing computer crime. In 1996, the Economic Espionage Act of 1996 was signed into law to, in large part, stunt the affect that the incredible

[2] U.S. Commerce Department, "Falling Through the Net: Toward Digital Inclusion," Washington, DC: U.S. Commerce Department, October 16, 2000.

growth of the Internet was having on the frequency of theft and destruction of trade secrets.[3]

Scope of the Problem

Recent statistics on the frequency of computer/Internet crimes point to the value of the enactment of computer crime-specific laws and their enforcement and demonstrate how computer crime has moved towards the front of crime concern priorities for the nation.

The Federal Trade Commission has reported that the number of consumer complaints, to the FTC, of Internet fraud and deception rose from less than 1,000 complaints in 1997 to over 25,000 complaints in 2000.[4]

The Internet Fraud Complaint Center announced in 2000 that the mean financial loss for Internet frauds reported to them was over $800, with victims tending to be clustered in the Northeast and West. Over 50% of the frauds were perpetrated through email.[5]

The Computer Security Institute/Federal Bureau of Investigation (CSI/FBI) 2000 Computer Crime and Security Survey of over 600 computer security practitioners in corporations and government agencies across the U.S. reported found that 70% experienced unauthorized use of computer systems, a 28% rise from 1996.

Nearly 75% of the businesses reported financial losses due to computer crime. Over $265 million was reported lost to computer crime victimization (the average annual total for the 3 prior years was just over $120 million). The most serious category of victimization was theft of proprietary information (over $66 million).[6]

The Computer Emergency Response Team (CERT) at Carnegie Mellon U., one of the most reputable sources of Internet security information, has revealed that the number of security attack incidents reported to them, nationwide, has more than doubled since 1998.[7]

These "hard" indicators of the frequency of crime commission and its associated damage highlight the growing threat of computer crime. Public surveys conducted by the Pew Research Institute have also illustrated how the issue of computer crime has crept into the public consciousness. According to the Pew Internet and American Life Project's most recent survey, 82% of the public are concerned that terrorists can commit their crimes via

---

[3] Jones Telecommunications & Multimedia Encyclopedia, Computer Fraud, (Available at www.digitalcentury.com/encyclo/update/comfraud.html).

[4] Stevenson, H. Testimony of the Federal Trade Commission before the Senate Finance Committee, April 5, 2001

[5] Internet Fraud Complaint Center, "Six Month Data Trends Report: May-November 2000." Fairmont, WV: National White Collar Crime Center/Federal Bureau of Investigation, February 2001.

[6] Power, R., *Tangled Web*, Indianapolis, IN: Que Corporation, 2000.

[7] CERT/CC Statistics 1988-2000

the Internet, 78% fear hackers getting access to government computer networks and 76% fear hackers obtaining access to business networks. Public perceptions of law enforcement, in this survey, proved to be quite supportive of law enforcement and the need to strengthen their abilities to enforce computer crime laws.[8]

Unfortunately, it has become apparent that the expertise required of law enforcers to competently battle the emerging menace of computer crime may not be matching the expectations of a public becoming increasingly aware of the gravity of the effects of computer crime. A recent National Institute of Justice survey of some of the most experienced law enforcement officials in computer crime representing over 100 law enforcement agencies at local and state government levels found that three quarters of the investigators believe "they do not possess the necessary equipment or tools to effectively detect and identify computer or electronic intrusion crimes."[9] Over 80% believed they required additional training on computer crime investigation to do there jobs properly and rated their abilities to deal with encrypted data as "low" or "doesn't exist".[10] It is not surprising that investigator participants in NIJ's study cited the availability and understanding of up-to-date forensic cyber tools as one of the most critical needs for computer crime investigators today.

The Approach of This Report: "Leveling" the Playing Field

Entrusted with the broad responsibilities of enforcing relatively new laws on computer crime is a growing army of investigators, like those surveyed in NIJ's computer crime needs assessment, specializing in computer crime investigation. Once found exclusively within the U.S. Department of Justice, computer crime investigators now populate many state attorney general offices as well as the offices of local district attorneys and police departments in urban and suburban areas throughout the U.S. Of course, simply having sufficient numbers of investigators dedicated to this crime area does not, in itself, guarantee effective enforcement of computer crime-related laws. The "new breed" of offender that takes advantage of the public's increasing use of computers requires a "new breed" of investigator, equipped with the requisite technological skills to level the new playing field of crime. The changing criminal environment demands a reassessment of what is needed to control "crime" as it is newly defined, or risk falling far behind methods employed by computer criminals.

The approach that the authors of this report take in addressing the needs of computer crime investigators, owes much to Cohen and Felson's routine activities theory[11] and Felson's 1998 work on the "chemistry" for crime commission.[12] Like routine activities theory, our approach is grounded in the understanding of situational activities that present

---

[8] Fox, S., and O. Lewis, "Fear of Online Crime: Americans Support FBI Interception of Criminal Suspects' Email and New Laws to Protect Online Privacy," Washington, DC: Pew Internet and American Life Project, April 2, 2001.

[9] Stambaugh et al, page 17.

[10] Stambaugh et al.

[11] Cohen, L.E., and M. Felson, "Social Change and Crime Rate Trends: A Routine Activity Approach." *American Sociological Review*, 44, 588-608, 1979.

[12] Felson, M., *Crime and Everyday Life*, Thousand Oaks, CA: Pine Forge Press, 1998.

special opportunities for the commission of crimes. Felson boils down predatory crime into three minimal elements – 1) a likely offender, 2) a suitable target, and 3) the absence of a capable guardian against the offense. The probability that someone will be an offender or target depends upon the "suitability" of the target from the offender's perspective. This suitability is typically measured considering the factors of – 1) the value of the target, 2) inertia of the target (e.g., rejection of theft of some items due to physical hurdles making theft impractical), 3) visibility of the target, and 4) access to offender with chance to exit easily.

Computer crime, in general, is a result of situations in which offenders capitalize on perceived opportunities to invade computer systems to achieve criminal ends or use computers as instruments of crime, betting that the "guardians" do not possess the means or knowledge to prevent or detect criminal acts. In many ways, these are old battles fought with new weapons accessing "unguarded" targets and permitting quick and unencumbered entry and exit. Cohen and Felson stress the importance of "target hardening" to counteract the criminal acts and help dissuade decisions leading to future criminal acts. Enhancing the abilities of the "guardians" is one of a number of ways to harden criminal targets. Viewing criminal investigators as the "guardians" against computer crimes and arming them with the best possible technological skills to close the gap between offender capabilities and those of law enforcement forms the core of this report.

For this report, the authors present the most up-to-date information on computer crime commission and investigation so the reader will understand, 1) how offenders use technology to commit their crimes (i.e., most popular and effective methods), 2) what enforcers must know to effectively detect/investigate these offenses and 3) in which areas offenders are still exceeding skills of law enforcement – areas where additional research and resources are needed for law enforcement to regain the competitive edge over the cyber criminal. To facilitate a better understanding of offender methods, investigative methods and the gaps between, the authors follow the lead provided by previously developed computer crime categorizations that consider computer crime from the perspective of the role the computer plays in the given crime – 1) the computer as target (e.g., intrusions, data theft, techno-vandalism, techno-trespass), 2) the computer as instrument (e.g., credit card fraud, securities fraud), and 3) the computer as being incidental to other crimes (e.g., data collection, protection and transmission for crimes such as drug trafficking, money laundering, child pornography).[13] This report takes these categorizations a step further and applies them to forensic tools used in computer crime cases.

The tools described as being used by offenders are logically grouped and categorized by function (e.g., Scanning Tools, Wardialing Programs, Password Crackers). The investigative tools presented address the investigative needs such as evidence source identification, evidence preservation, evidence extraction and evidence analysis. These

[13] Carter, D.L., and A.J. Katz, "Computer Crime: An Emerging Challenge for Law Enforcement," *FBI Law Enforcement Bulletin*, 1996 (Available at http://www.fbi.gov/leb/dec961.txt).

tools are grouped into the general categories of – 1) Evidence Collection and Preservation Tools, 2) Evidence Extraction Tools, 3) Evidence Examination Tools, 3) Evidence Organization Tools, 4) Network Forensic Tools, 5) Attack Analysis Tools, 6) Multi-Purpose Forensic Tools and Toolkits, 7) Honeypots, and 8) Trusted Time Stamping. The tools are separated further, by function, into subcategories (e.g., Intrusion Detection Tools, Trace Back Tools). The body of the report offers a general description of the investigative tools with directions on where more specific information on the tools can be found in the report's appendices.

The material contained in this report rests heavily on the technical expertise of the authors as well as previous research conducted by two of the authors (Gordon and Hosmer) for the Forensic Information Warfare Study (completed for the Air Force Research Laboratory in Rome, New York). To help ensure that the report is a "utility-based" research product, the authors drew upon information generated through the NIJ Law Enforcement Needs Assessment Study, mentioned above, and through the authors' own survey of law enforcement practitioners familiar with computer forensic tools. This survey was designed to determine what computer forensic tools law enforcement practitioners use most frequently, what are the perceived strengths of the tools and what are the perceived weaknesses. The authors have relied upon empirical data from these two studies for guidance in identifying those needs considered most critical for improving computer crime investigative skills and most essential for reclaiming the technological advantage over cyber criminals.

Task 1: Assessment of Tools Used in the Commission of Cyber Crimes

Objective

Task I provides a review of the role that computer technology currently plays in the commission of cyber crimes; the tools and techniques used by criminals in carrying out specific cyber crimes.

To achieve that objective, a description of the major categories of tools used by offenders follows. The purpose is to provide law enforcement practitioners with an accurate portrayal of those tools currently employed during the commission of cyber crimes.

The tools are described in the context of generic 'classifications' of tools. Additional information is provided that aids the practitioner in identifying and/or locating the "fruits of the crime" - the data that these tools have aided in gathering and/or producing during the commission of the crime. This data provides the necessary link between the perpetrator and the cyber crime under investigation.

Approach

The technology and software described here provide a broad cross-section of current and evolutionary technologies. While many of these tools have legitimate usages, for security testing and as diagnostic aids, the techniques used by commercially available 'penetration-testing' tools are the same as those tools used in the commission of cyber crimes.[14] Each of the technologies was assessed for its usefulness, and potential, for use as a tool in support of criminal objectives.

Individual versions of each type of tool are widely available, and easily obtained. They have been classified based on similar purpose and functionality. Many variations of the tools exist within each classification. The most common variations are those related to:

- Operating System Differences
- Command-Line vs. Graphical User Interface (GUI)

The effort will concentrate on evaluating tool classifications based on their similarities in operation and functionality, but noting advancements in the technologies that make, or promise to make, the tools a more formidable threat.

[14] Many commercial vendors of security software got their start by creating an early version of a penetration-testing tool for the underground community.

Structure of Task 1

For the purpose of this paper, the tools used to commit cyber crimes are grouped together into mutually exclusive categories. These categories delineate the different roles that a computer can play during the commission of a cyber crime. These roles are described in the following Introduction Section. Within each category, the different tools are discussed.

Description Section

This section provides the investigator with an explanation of the individual type of tool, its basic functionality, optional features employed by some versions of the tool, how it may best be used, and, where applicable, how it may be introduced into a target system (if the investigation involves a compromised system).

Evidentiary Value Section

This section is most relevant to the *forensic investigation* of a computer incident. It describes what additional evidence the investigator should look for once the presence of a particular type of tool has been identified on a system. This is the actual evidence that could provide the link between the perpetrator and the alleged cyber crime. The investigator must be aware that these systems can hold this additional evidence, and that it is up to him to collect all evidence that may be present on that system.

This paper does not attempt to describe *all* types of digital evidence that may be derived from the computers in their individual roles, but to educate the investigator on the types of cyber crime tools that may be present on the given system, dependent on the type of crime under investigation, and the collateral evidence associated with the presence of such tools.

Introduction

Carter and Katz proposed a set of categories and definitions in order to aid law enforcement in developing investigative strategies and procedures in the area of cyber crime. Their approach was to describe a computer associated with a cyber crime within the context of the role that the computer plays in the cyber criminal act. And, for each role of the computer, there are distinct sets of associated cyber crimes.

According to Carter and Katz, any computer encountered during the course of the investigation will fit into at least one[15] of the following categories:

- The computer is the *target* of the crime
- The computer is the *instrumentality* of the crime

[15] A computer involved in multiple cyber crimes may fit into more than one category

- The computer is *incidental* to the crime

The strategies and procedures for the investigative process differ, depending upon the role of the computer in a cyber crime, as does the evidence that can be collected. We have extended Carter and Katz's definitions to further provide a framework for the classification of those cyber crime tools ('cyber tools') that may be found within the particular 'crime scene' (the computer currently under analysis). By using the categories they have described, we have classified the tools accordingly.[16]

It must be kept in mind that, as with traditional crime, for every cyber crime there is a perpetrator (using the instrumentality of the crime) and victim[17] (the target of the crime). Typically, the perpetrator of the crime will use a particular tool for the job, but keep that tool in their physical possession, leaving behind only the indication that a tool was employed. This analogy is applicable to the use of cyber tools that will be present as the *instrumentality* of the cyber crime. These cyber tools may provide the investigator with the 'smoking gun' needed to connect the dots between perpetrator and victim.

But, unlike physical tools used to commit crimes, some cyber crime tools operate best when left behind by the perpetrator. These types of tools are analogous to a covert listening device, left behind after the initial compromise to assist the perpetrator in furthering their criminal ends. If located, the covert device may provide the investigator with information that could potentially be traced back to the intruder. In much the same way, cyber tools may be left behind on a computer that has been the *target* of the cyber crime.

While a particular tool may be thought of as most closely associated with the target of the cyber crime, and indeed is the tool used to commit that crime, the relevant evidence that will link the victim with the perpetrator is that evidence that may be present on the instrumental computer; the tool itself, or the output from that tool.

The Roles

In the context of cyber crime investigations, the perspective from which a computer will ultimately be analyzed is directly related to the role it has played in the cyber crime.

Target

When a computer or computer system is examined as the *target* of the crime, the investigation has determined that a *computer crime* has occurred. The computer system may have been accessed a) without proper authority or permission, b) legitimate access to

---

[16] It is at this point that we have departed somewhat from Carter and Katz's definitions. While his descriptions of the roles list the associated cyber crimes as an exclusionary set for each role, we take into consideration the fact that a computer may be an instrument, *as well as the target*, of those types of crimes that he exclusively associates the computer as the target of the crime. We have found that tools exist that are considered instruments in the commission of computer crimes, and should be classified accordingly.

[17] With the exception of victimless crimes; e.g. gambling.

the system may have been blocked or disabled, or, c) some type of malicious code has been introduced into the system. When the computer is the target, the investigator looks for evidence of the compromise or attack, and additional evidence that may assist in identifying the origin of this malicious activity.

The crimes under investigation are exclusively those crimes that are enabled through the proliferation of computers and networked systems. They are known as *non-traditional* crimes, the types of offenses that computer crime statutes were written to address.

Carter and Katz use examples of several types of cyber crimes to define the meaning of a target computer. Those crimes include, but are not limited to:

- Computer intrusion
- Data theft
- Computer vandalism
- Computer trespass

Where the computer is the target of one of these crimes, a distinct set of cyber tools has been identified that may be found on that system. These types of tools have been introduced into the system by the perpetrator, left behind in an attempt to collect, and subsequently provide, additional data to the perpetrator. This data could be:

- User account information (passwords)
- Administrative account information (passwords)
- Proprietary data
- Credit card numbers
- Personal information

Instrumentality

When a computer is examined as the *instrumentality* of the crime, the investigation has determined that there is sufficient reason to suspect that the computer was used as a tool to commit, or further advance, the crime under investigation.

- Computer applications were used to further advance a theft or a fraud.
- The computer was used to block or gain access to other computer systems, and to possibly manipulate these systems to produce a desired result.
- The computer may have been used to create malicious code (e.g. a virus), generate credit card numbers or bank checks that are used to facilitate a fraud, or commit an act of counterfeiting.

Computers that fit into the category of instrumentality may be used to commit traditional, as well as non-traditional, crimes. Carter and Katz used the following crimes as examples of those crimes within which the computer would serve as the instrumentality:

- Credit card fraud

- Telecommunications fraud
- Theft
- Fraud

While the processing power and automated applications available to cyber criminals could aid in the commitment of such traditional crimes, the computer may also be used as the instrumentality to commit non-traditional crimes, as well. These types of cyber crimes would include, but not be limited to,

- Unauthorized access to a computer (over a network)
- Denial of service
- Harassment
- Cyber stalking
- Create malicious code

The first cyber crime in the list, unauthorized access to a computer, was mentioned in the previous section as a category within *target.* But, when the computer is the *instrumentality used to commit* such a crime, the tools present on the system, and collateral evidence that the investigator would be seeking out, are totally different from those tools where the computer has been the target of this same cyber crime.

Incidental

When a computer is *incidental* to the crime, the investigation has determined that the computer will contain additional evidence that is relevant to the crime under investigation. In this situation, the computer itself is not an essential element for the crime to have occurred, but the technology that a computer provides has assisted in the commission of the crime.

A computer that plays the role of a system incidental to the cyber crime may contain evidence of traditional, as well as non-traditional, crimes. Carter and Katz mention the following types of crimes as descriptive of their definition:

- Copyright violations
- Software piracy
- Child Pornography

Documents, databases, records may be found on the system that are directly related to the commission of other traditional crimes, such as fraud (financial, credit card, etc.), the sale of illegal substances (drugs, foods), extortion, gambling, as well as identity theft.

Supporting evidence of a non-traditional cyber crime may be retained in a computer's logs, such as those found on e-mail servers or Internet Service Providers.

- An email server may contain copies of messages sent during the course of a cyber stalking.

- An Internet service provider may have records of users logged on during a given time frame.

Tools may be found on that computer that, while not directly associated to a specific type of crime, may be indicative of suspicious activity. These tools have been associated with those perpetrators that wish to hide illicit activity, and use the techniques provided by the tools to 'hide' or disguise the relevant evidence of their activity.

Introduction to the Tools

Specific categories of tools can be associated with each role the computer has played in a crime. The investigator may expect to find any or all of the cyber tools associated with this predetermined role.

These tools are classified using the previously outlined framework. A thorough discussion of each individual tool is beyond the scope of this paper, as there are too many tools within each category. The purpose of this section is to familiarize the investigator with the types of tools and techniques used to compromise protected computers and networks, and/or commit associated cyber crimes. These are the tools that will typically be encountered during the examination of a computer involved in a cyber crime. The tools are categorized for each separate role the suspect computer plays in the commission of a cyber crime.

Computer as the Instrumentality of Cyber Crime

Gaining Unauthorized Access

Within this section is a discussion of cyber weapons. These tools allow an individual to automate techniques used to commit a cyber crime, these techniques that would otherwise be labor-intensive and time-consuming.

These tasks could be performed manually, but would involve many steps and a great deal of time in order to achieve the desired goal. Alternatively, cyber weapons allow an individual to complete these and other tasks in an automated fashion, taking a fraction of the time that the manual methods would take. And, because of the ease with which the tools operate, the bar is lowered on the level of knowledge that the individual needs in order to perform these tasks.

Scanning Tools

Description

Probably the most useful tools that an attacker can have in his arsenal are network-scanning programs. A scanner is a program that can identify active[18] networked

[18] Systems that are currently receiving and sending computer network communication.

computers, and gain valuable reconnaissance data about the type of operating system that computer is running (as well as the version), open system services (email, FTP, HTML sever, etc.) and a host of other data, depending on the capabilities of the scanning program.

Some scanners are designed to scan only a single networked computer, gaining as much reconnaissance data about that system as possible. Others can scan any entire range of network addresses, seeking out those that appear to have a particular operating system or service running that may be vulnerable to an attack. Still others are designed to scan and map out entire Local Area Networks (LANs), identifying each host that resides on that network.

Once mapped, it is simple to single out those systems that may have security weaknesses. It is now possible for an attacker to determine which other tools and scripts[19] from their arsenal they can now deploy against a selected target.

Without scanners, and the information they provide, an attacker could spend an enormous amount of time blindly throwing every possible exploit script at the target, not knowing which operating system or version was being used, what service packs or patches[20] had been applied, what services were running on which port[21], or if a proxy[22] or firewall[23] is in place that could defeat many attacks.

In short, scanning tools may be able to do any or all of the following:

- Find a 'live' target network or system by pinging[24] a range of Internet Protocol (IP) addresses[25], and recording those that respond;
- Identify and list all active services running on the target server, by creating a list of open ports;
- Identify the operating system of a particular server, and possibly indicate which service pack has been applied;
- Scan a target server, seeking out a specific service (e.g. file transfer protocol (ftp)), and attempt to exploit it for any number of known vulnerabilities;
- Seek out trojan[26] servers that are installed and running on a remote machine;
- Probe firewalls for configuration errors;

---

[19] Scripts are small applications written to exploit a vulnerability related to a ***specific*** application, operating system, or networked process/service (such as an email program).

[20] Patches are actual codes that fix a known bug or vulnerability in software. Services packs are updates to software programs that improve or enhance the product.

[21] Ports are communication gateways into a computer system.

[22] Proxies are tools used to filter network communication, and improve the performance of groups of users.

[23] A firewall is a system designed to prevent unauthorized access to a network.

[24] Pinging is a means of communication between computers. One computer will send a packet of information to another computer, and wait for a reply. If a reply is received, the computers are properly connected.

[25] IP addresses are unique identifiers of a networked system, and these addresses can be matched to provide a tentative link between the suspect and the victim.

[26] Refer to the *Target* section in this task for a further description of Trojans.

- Probe Windows hosts, looking for open shared resources.

The first step for planning an attack is the reconnaissance, or information-gathering stage. From a network perspective, this means using one or more scanning tools. An experienced attacker will have in his arsenal a variety of tools that will scan a target computer, range of IP addresses, or Local Area Networks (LANs), looking for ways into systems.

In order to map the target network, in preparation for an attack, scans are conducted against systems to see which hosts are up and running. For this, they use a ping sweep.[27] Once the hosts are found, further scans may be carried out against them. Using any number of protocols,[28] the next step is to look for open ports on the target system[29]. These scans generate a list of services on systems that have responded to network pings. It is then a matter of examining this list, and choosing a target based on the information derived from the scan.[30] After this, it is up to the attacker to use the appropriate exploit script[31].

These tools are popular because they are widely available, they are free, they can legitimately be used as security products, they are legal, they are available for every operating system, and they provide anonymity for the user.

Evidentiary Value

The presence of a scanning application does not itself indicate intended malicious activity. Since the advent of scanning tools, security and vulnerability scanners have found legitimate use as a way for system administrators to analyze the status of security on their networks. And, because of their non-invasive nature, there are no existing laws that might serve to deter their illegitimate use. "The courts have described the use of a scanner as virtual "doorknob rattling."[32,33] While victims seek redress on the issue of minimized bandwidth[34] capabilities, the courts have found that the amount of bandwidth used does not reach the threshold of depriving a target of a significant resource. Until an overt act is committed against these targets, no crime exists.

While the presence of the tools themselves proves nothing, it is the output from these tools that provide the incriminating evidence of the user's illicit activity. These scanning tools generate lists. These lists contain, among other things, IP addresses of potential

---

[27] A ping sweep is where the offender pings a range of IP addresses recording those that respond.
[28] The most common protocols are for scanning activity are TCP and UDP.
[29] This means looking for communication gateways that might available for use.
[30] As an example, using a vulnerability scan would produce a list of potential targets that appear to be running a flawed or unpatched version of an application or a service.
[31] An exploit script is a set of commands that attempt to break into a computer system.
[32] Testing a computer to see if it has vulnerabilities that can be exploited.
[33] Moulton v. VC3, N.D. Ga., Civil Action File No. 1:00-CV-434-TWT.
[34] Bandwidth is the amount of data that can be received in a certain amount of time.

targets, as well as port numbers and susceptible services[35] that may be running on the target machine.

Using these lists, the investigator can potentially tie the system that was used as the instrument of the crime to the system that was the target or victim of the attack, by using the IP addresses. The evidence can further be strengthened if it can be shown from this list that the victim was compromised by one of the vulnerabilities that the scanning utility had identified.

Password Crackers

Description

Passwords are everywhere. Banks, credit card companies and telephone companies, as well as many others, incorporate the use of passwords, or Personal Identification Numbers (PINs), to authenticate the user's of their services.

When the term *password* is applied to computers, it is referring to the measures that are in place that authenticate the user to that system, or, referring to protections placed by the individual users to prevent unwanted access to their personal information, files and applications. Password cracking applications are computer programs that attempt to circumvent these protections.

All computers store passwords within the system, in order to authenticate that the users are who they claim to be. Early versions of operating systems stored their password files in plain text. All an individual needed to do was to find a way into the system, and steal this file. Today, in order to protect password files from this type of activity, they are stored as encrypted[36] files. So, even if access is gained and this file is captured, it is useless in this encrypted state.

A password-cracking program does not actually "decrypt" the passwords. The CPU time it would take to decrypt even one password would make this approach unfeasible. What a typical cracking utility will do is accept individual words from a "dictionary" (a list of words that could be used as passwords). The program then encrypts the individual words, and the encrypted value is compared to the captured password file. Because many users are known to choose weak passwords, it is not long before an attacker has a list of passwords that can be used to enter the target system.

One drawback to using the password-cracking utility is that it takes a very long time to run.[37] Every word from the "dictionary" must be encrypted, and compared to every entry in the stolen password file.

---

[35] These are programs on the computer that could be vulnerable for an attack.
[36] Encryption is the transformation of data from plain code, to a secret code.
[37] LockDown, The Home Computer Security Centre, (Available at http://www.lockdown.co.uk/security/combi.php).

There are many different cracking programs available, but they typically run through a series of stages:[38]

1. Try common passwords, such as "password" or the name on the account in question.
2. Run through all the words in the dictionary and lists of common passwords.
3. Add numbers to the end or the beginning of these dictionary words.
4. Run through all the words in foreign dictionaries and special "crack" dictionaries.
5. Try all combinations of letters out to a certain size, such as 5 letters (brute force method).
6. Try all combinations of letters, upper/lower case, numbers, and punctuation out to a certain size, such as 3 characters.

When a password has been compromised, the attacker has full access to the user's account and associated permissions on the system. He can use this account as a platform for an attack, thereby disguising his true identity, and leaving the legitimate owner of the account unaware that his account has been used in such a way. It can also be used as a way to springboard to other systems. A sophisticated attacker will chain together several compromised accounts, and effectively hide his actual location. The more accounts an attacker can compromise and use for this purpose, the less likely it is that a successful trace can be made.

Many common applications allow the user to apply password protection to select files. Word processing documents, spreadsheets, databases, etc. may be 'locked' in such a way that the creator of the file can allow or deny access to them on a selective basis.

Password cracking programs are available that may allow the user to circumvent these protections. Typically, these programs are specific to the file that it will be used against (i.e. a 'Zip cracker' will only work against password-protected zip files). Other than this difference, these programs operate in much the same way as the aforementioned crackers, employing brute force techniques in an attempt to guess the password.

Evidentiary Value

What an attacker will do after the break-in is download the captured password-protected file(s), and run the cracking program on their own system, because the cracking programs are CPU intensive[39], and a spike in CPU activity on a compromised machine will be easily spotted.

The captured password file(s), as well as a list of 'cracked' passwords and their associated usernames, would provide definitive evidence of a computer compromise.

---

[38] Network Ice, Password Cracking, (Available at http://www.networkice.com/Advice/Underground/Hacking/Methods/Technical/crack/default.htm).

[39] This means that these programs use a lot of the computer's memory, therefore slowing the computer's performance.

Wardialing Programs

Description

Wardialing involves using the computer's modem to call a range of telephone numbers, seeking out and saving the numbers that answer with the telltale 'handshake tones'[40] used by computer modems or fax machines. Wardialing programs use the computer to automate the process. The program will accept, as parameters, the first and last numbers for a range of telephone numbers, dial all numbers within that range, and record those that answer in a database or log file.

Those numbers that are logged indicate potential entry points to computer or telecommunications systems. Some of these programs can distinguish between modem, fax, or Private Branch Exchange (PBX)[41] tones, and log each one accordingly. If a modem is detected, they can capture certain details of the system to which that modem is attached. Some wardialers can then further assess the security of the system by attempting an array of login attempts. Those systems determined as vulnerable in this manner can then be prioritized as viable targets.

Using these tools, an attacker can scan an entire business exchange in several hours,[42] identifying all hosts with modems or other networked devices in that range. It is generally easy to determine a range of phone numbers to dial by finding the target's main telephone number or fax number. This is often publicly available information.

The task of locating targets is now automated. In this way, an attacker may find any unregistered or unsecured dial-in modems that may be installed within that telephone exchange.[43] Securing the network perimeter will not prevent the use of an unauthorized modem. A modem is a means of bypassing the perimeter defenses that protect the network from intrusions. By using a wardialer to distinguish the modem telephone number, and a password cracker to break a weak password, access can be had to the system. Once a connection is made, a connection to any other locally networked computer can be made.[44]

Not only is this tool useful for attacking computers, it is also one of the most important tools in the phreaker's[45] tool kit. The wardialer is to them what the port scanner is to a computer attacker. It gives them a list of potential targets for their illicit activities.

---

[40] The signature tones transmitted over communication lines that enable one computer to recognize and initiate contact with another.

[41] A PBX is a private telephone network used by companies.

[42] Most organizations have a block of sequential phone numbers.

[43] While most business security policies do not allow these types of dial-ins, it is not unusual for users to install their own modems for remote access during non-business hours.

[44] U.S. Army Space & Missile Defense Command. (Available at http://www.smdc.army.mil/SecurityGuide/v1comput/Modems.htm).

[45] The term associated with the individuals that specifically target the telephone system. Rather than looking for access points into computer systems, phreakers attempt to locate entry points into telecommunications systems.

Changes in phone networks have made this activity much less appealing, but with the introduction of voicemail systems that allow connectivity to an IP network, this may change. The newer phone switches are now Transmission Control Protocol (TCP)/Internet Protocol (IP)[46] based, which may appeal to an attacker, always on the lookout for a new avenue to exploit.

Increasingly more people are using mobile phones to pick up email, access the net etc. Many users would not traditionally think of these as modems either. By default, when a computer is hooked to a mobile phone, it automatically answers incoming data calls.[47] Unsolicited "data" calls to a mobile phone, with caller ID withheld, could possibly point to the use of a war dialer.

Many telephone companies have equipment to detect wardialing, and can block an attack once this activity has been identified. However, this equipment only detects sequentially dialed number attacks. To prevent this means of detection, many wardialer programs allow the randomization of the order in which they dial telephone numbers.[48]

Evidentiary Value

The logs or databases of targeting information that the tools generate provide evidence that can link the owner of the computer to a particular system attack. Not only do these logs provide telephone numbers that belong to the target of an attack, but also the more sophisticated wardialing programs provide additional data about weaknesses in a target system, these weaknesses being those that may be exploited by the attacker.

Denial of Service

Description

Denial of Service (DoS), in its simplest terms, means rendering a network service (e.g. email or HTTP) unavailable to others. Generating and sending so much traffic to a target network, that all bandwidth is consumed, and no legitimate traffic can pass, can accomplish this. Other DoS attacks direct exorbitant amounts of messages to a target server, thus filling up all available space within which the service runs (i.e. mail server queues). Or, by exploiting a flaw within a network service, they cause the target machine to crash.

Reasons that an attacker would want to use a DoS attack that crashes a target computer might include the following:

---

[46] A suite of communication protocols used to connect systems on the Internet.
[47] NFR Security. (Available at http://www.nfr.com/pipermail/firewall-wizards/1999-December/007449.html).
[48] Network Ice Wardialers, (Available at http://www.networkice.com/Advice/Countermeasures/Scanners/War_Dialers/default.htm).

1. A Trojan has been installed, but the system must be rebooted in order to install it.
2. The attacker wishes to cover their tracks, or excessive CPU activity, with a system crash.

Many simple flooding/nuking programs exist that will send the traffic in a variety of ways.[49] Each program has a unique approach[50] to their creation of this illicit communication. But, unless an attacker is using a spoofing[51] technique, DoS attacks are relatively simple to trace back to their source.

Email Flooding Programs

Email flooding programs, a subset of DoS tools, are designed to attack and render useless email services. The tools generate many messages in a short period of time, and transmit these messages to the targeted user or email server, this provided by the attacker. The receiver's mailbox is quickly filled to overflowing with the massive amounts of email. Email services for the specific user, or to an entire organization, may be blocked or brought to a halt by the influx of messages, these messages containing random 'garbage' as their content. This could be devastating to individuals or businesses that are dependent on email for purposes of communication.

Distributed Denial of Service (DDoS)

A DDoS is a special kind of Denial of Service attack. While the 'distributed' concept may suggest more than one participant, these attacks typically originate from a single attacker. The attacker begins by compromising many networked computers, and obtaining administrative or root privileges on all of them.[52] He then installs specially designed DoS 'agent'[53] software on them. This software will allow the computers to be controlled in a coordinated manner when the attacker decides to launch attacks on the target systems. These compromised computers (also known as 'zombies') are unwitting participants in the attack.

These agents will await commands from a central handler[54], the portion of the program that sits on the attacker's computer. The handler will then contact all the agents, and instruct them to send as much traffic as they can to one target. The tool coordinates the timing of the flooding of a target system, and directs the activities of all available DoS agents, thus the distributed concept. These attacks will typically exhaust bandwidth, router processing capacity, or other network resources, blocking network connectivity to the victims.[55]

---

[49] Using different network protocols.
[50] There are many variables that can be manipulated in the header of an IP packet, each having a different effect on the target system.
[51] See section on 'IP Spoofing'.
[52] To gain access, scanning tools are used to probe for systems with specific vulnerabilities. These vulnerabilities are then exploited using freely available scripts.
[53] Also may be referred to as the 'server'.
[54] Also may be referred to as the 'client'.
[55] Amis, R., Recommended Daily Requirement, G21 Magazine, February 17, 2000, (Available at http://www.g21.net/daily0217.htm).

Once the DDoS attack has been launched, it is very difficult to stop. It is possible to block packets at the victim's firewall, stopping the flood from directly attacking the victim's internal systems. But, the flood will continue to overwhelm the Internet connection, making the target unreachable by legitimate network requests. If the source can be identified, it may be possible to contact the administrators of the 'zombies', inform them of their role in the attack, and ask them to stop the traffic. If the source IP addresses of the packets have been 'spoofed' (faked), there is no way of quickly determining the source of the attack until the traffic has been traced, and the owners contacted.[56]

As devastating as these attacks were, the tools used were considered to be first generation. A paper entitled "TFN3" outlines future evolutional possibilities for such tools.[57]

Evidentiary Value

DoS, email flooding, and DDoS tools are readily available, and any Internet host is a potential target, either as a zombie or as the focus of the attack. Distributed attacks are the most difficult type of denial of service attacks to deal with, because they are very hard to block and shut down, especially when the traffic is found to have originated in countries that don't have the a legal infrastructure in place to deal with this type of crime. The traffic is arriving not from one source, but many. It takes time to identify these sources, and block the traffic.[58]

Potentially useful evidence may be obtainable from the DDoS client portion of the tool, as it requires a list of server agents. Finding a system with a list of agents makes the task of uncovering other agents much simpler. Additionally, some of the agents themselves may include an encrypted list of master clients, but breaking the encryption may prove to be very difficult and time-consuming.

Anonymous Email

Description

Anonymous email, also known as email spoofing, is the deliberate misconfiguration of source or return email information, such as the username or originating domain, within any email. In other words, a user receives email that appears to have originated from one source, when it actually was sent from another source.[59]

---

[56] Farrow, R., Distributed Denial of Service Attacks, Network Magazine, March 1, 2000, (Available at http://www.networkmagazine.com/article/NMG20000512S0041/2).

[57] TFN3k is a paper about the future of DDoS tools, how they can be used, and the dangerous features that can and probably will be implemented in the future. Tribe Flood Network 3000, (Available at http://packetstorm.widexs.nl/distributed/tfn3k.txt).

[58] Bell, M., Undernet IRC Network Under Siege, Monitor Magazine, (Available at http://www.monitor.ca/monitor/issues/vol8iss7/online.html).

[59] CacheNet, (Available at http://www.cache.net/acceptable_use.shtml).

Anonymous email is actually a combination of software and service. The application provides a seamless interface to the anonymous email service. These services are available on the Internet for the exclusive purpose of hiding the origination of email communication, and promote their service as a way of protecting the privacy rights of the user.

The user goes to the anonymous server site and downloads the appropriate software. After registration, the user is set up with one or more electronic pseudonyms[60]. The anonymous server accepts messages sent by the user, and replaces the actual return address of the message with the return address of the user's pseudonym. The message is then encrypted and submitted for delivery via the anonymous server, hiding the message's point of origin. If the recipient responds, the anonymous service's server will take the message, encrypt it, and deliver it back to the user's e-mail address.[61]

Anonymous email can be used in an attempt to fool a victim into making an unguarded statement, or releasing security information (such as passwords). By impersonating a trusted contact, the actual sender will use this deception to gain the target's faith in the return address of the email, and unwittingly give out sensitive information.[62]

Evidentiary Value

If it is important to first verify that a suspect has been using such a service, a thorough search must be conducted of bookmarks, temporary Internet files, and the cache to extract the addresses of these services (this would assume that the investigator has a current listing of all available anonymous email services with which to compare the output to). If the suspect's machine is within a LAN, any intermediary hosts (firewall, proxy server, etc.) that do logging may also reveal the use of such a service.

The examination of illicitly sent email on a suspect's computer has long been a valuable source of evidence for the criminal investigator. The use of these programs and services alters that information in an email header that would provide the most sought after clues. It randomizes the return address, or uses a fictitious return email addresses, thus making it impossible to determine the originator of the message.[63]

Online anonymity makes it more difficult for a law enforcement officer to successfully catch and prosecute Internet-based criminals. There are many computer crimes (e.g. child pornography) that may be committed online; this anonymity can significantly complicate an investigation.[64]

---

[60] Username, or identity.

[61] Newton, M., Hide Your E-Mail Tracks With New Privacy Tool, PC World Magazine, March 2000, (Available at http://www.pcworld.com/news/article/0,aid,14930,00.asp).

[62] CERT Coordination Center, Spoofed/Forged Email, April 26, 1999, (Available at http://www.cert.org/tech_tips/email_spoofing.html).

[63] Esper Systems, (Available at http://www.esper.com/aup.html).

[64] MacMillan, R., Attorney General Complains About Net Anonymity, May 23, 2001, (Available at http://www.infowar.com/law/01/law_052301b_j.shtml).

IP Spoofing Tools

Description

IP spoofing[65] involves the creation of network traffic that appears to have originated on one machine, but is actually from another. This is accomplished by changing the source information (IP address) contained in the header of a network packet to an address other than that of the originating machine.

Routers only use the destination IP address to forward TCP/IP packets; they do not verify the source IP address. The only time the source address is needed is when the destination machine uses this source address to respond back. Forging the source IP address causes all responses to this communication to be directed to a machine other than the origin, thus effectively disguising the source of an attack that implements this technique.

Illegitimate traffic may be allowed onto a local network that would not normally be allowed. A LAN that blocks traffic based solely on source addressing would allow in this type of traffic. In this way, an attacker can insert any type of traffic into the LAN (including DoS), because the source information makes it appear to have originated from a trusted associate.

Evidentiary Value

By using IP spoofing tools and techniques, an attacker can achieve virtual anonymity. By changing or obscuring the originating address of illicit Internet traffic, there is no effective way to traceback this traffic to the perpetrator. Traceback capabilities are becoming more widely implemented in security product suites, but will produce misleading information to the investigator that trusts the results on their face value. A sophisticated attacker will use spoofing techniques to cover his tracks, and protect his identity.

Advancement of a Crime

Credit Card Number Generators

Description

These programs are based on the algorithmic formulas that the major credit card companies use to generate their credit card numbers. Every company has its own approach to generating these numbers. Therefore, while some of these programs are designed to generate numbers that fit just one company's formula (such as MasterCard numbers), others will give the user the option of creating other types of numbers (Visa, American Express, etc.).

---

[65] Daemon9, IP Spoofing Demystified - Trust-Relationship Exploitation, Phrack Magazine, June 1996, (Available at http://www.fc.net/phrack/files/p48/p48-14.html).

Included also within this category of tools are those that are designed to generate telephone calling card numbers. Again, they use the same formulas that the phone service providers use to generate numbers.

This type of application can produce as many numbers as the user requests. While these may be *valid* numbers, there is no guarantee that they belong to an active account.

There is the possibility that the generated numbers may indeed belong to an active account, which can be tested in several ways. There are sites on the Internet that offer such services. The user need only provide the service with the credit card number. As long as the account remains active, the fraud will continue.

Evidentiary Value

The list of the credit card and calling card numbers that these programs generate would be the most definitive proof of the use of such programs. A search of the suspect computer should be conducted for number sets that match the patterns of numbers of common credit cards.

Additionally, users of such programs may have in their possession the equipment to produce physical copies of credit cards. This equipment would be used to facilitate the credit card fraud.

Virus Generators

Description

Virus[66] generating programs[67] give the user the ability to create custom virus code. They allow the user to select and customize the characteristics of the virus they are designing. Users can usually specify the following characteristics:

- Virus name
- Author name
- Whether to implement encryption or not
- Whether to implement anti-debugging techniques or not
- Minimum and maximum file size of the host file
- Maximum number of infections
- Whether it is a COM or EXE infector
- Whether it infects COMMAND.COM
- The trigger date for payload

[66] A virus is a piece of code that runs on a computer, and has the capability of causing damage to a system.
[67] Also referred to as constructors, creators, and factories.

The most effective viruses are written in assembly language. These programs simplify the process of writing a virus by providing a high level interface to the underlying assembly code (ASM).

While many of these new viruses will not get past modern protection software, there is always the exception to the rule. Using these types of programs, individuals with little or no knowledge of how to program a virus can produce potentially malicious code.

Evidentiary Value

The use of these types of programs have been linked to several well-publicized viruses released in the recent past.[68] While the program may not be directly linked to a particular virus, earlier versions of the code may remain on the suspect machine.

Computer as the Target of Cyber Crime

Within each category of the tools to follow, there are a widening variety of tools available for multiple operating systems, and capability levels of the user. In the right (or wrong) hands, these tools provide the user with a powerful set of weapons. The tools are widely available and accessible through many Internet 'security' sites (the term security is used loosely). Many come with a user-friendly graphical interface, providing relative ease of use of these tools. These weapons enable many individuals that do not have the requisite technical expertise to launch attacks.

Packet Sniffers/Analyzers

Description

A packet sniffer (or just sniffer) is a simple program that passively listens to network traffic, recording all of the traffic, or selected portions of it. The sniffer then produces analysis based on the recorded traffic, and provides the analysis in a readable report.

A sniffer puts the Network Interface Card[69] (NIC) of the target computer into a mode known as promiscuous mode. To explain, each computer on a network will normally receive all traffic passing along that network, but will ignore the traffic that is not destined for that computer. However, a NIC set to promiscuous mode accepts, records, and examines any and all packets it receives, monitoring all traffic being transmitted over the network.

Most sniffers only monitor one connection at a time. The reason for this is to make the sniffer harder to detect, due to smaller logs and less use of CPU power. A small number of sniffers monitor all connections.

---

[68] The VBS worm generator version 2 was used to create the Anna Kournikova virus.

[69] Also known as an Ethernet card, one of the necessary pieces of hardware to physically connect computers together.

Sniffers have different methods of logging. Some sniffers will only record the first x (x being a certain number) number of bytes of a packet, in order to capture a user's login/password combinations. Another method will capture the entire session. Some of the more versatile sniffers will support both methods. The specific type used will vary depending on the intruder and the desired end result.

One method that has been used to break into secure machines has been to break into another machine, either directly or in-directly, that the target machine trusts.[70] Therefore, if the attacker can access a trusted computer, he can abuse that trust, and use it as a compromise into the rest of the network. By monitoring the traffic on a trusted system, an attacker is likely to gain important intelligence from the information transmitted between the two systems.

Evidentiary Value

Often times, looking at the CPU usage and file system are the only ways to detect such sniffers. If CPU usage is higher than normal, or there is consistently unexplained loss of disk space, it may point to the presence of a sniffer.

Investigators must be mindful that the presence of a sniffer within a network indicates a serious security problem, as a network card operating in promiscuous-mode requires root privileges on a majority of UNIX and Linux operating systems. But, the sniffer is only an indicator of an incident, and does not itself provide any evidentiary value in identifying the source or the perpetrator of the crime.

Keylogger Programs

Description

A key logger is a small application (usually only a few Kbs in size), installed directly on a user's machine, and used to record the user's every keystroke, saving these to a file (log file).

The standard features of a key logger include:

- Record all keystrokes, including numbers and special characters. Key combinations are also recorded (e.g. ctrl +alt +delete)
- Log startup and shut down time
- Run automatically at startup, invisibly
- Log file encryption
- Password-protected controller
- Specify characters to be logged
- Specify the logger path and log file location

[70] Trust, within the scope of a network environment, means that some machines are configured to 'trust' other computers to share resources. Security between trusted computers is minimal, if it exists at all.

- Options to automatically clear the log file
- Run in system tray, so the menu can be accessed easily

Some also have special features, such as:

- Automatically send the log file
- Remote commands
- System information

There are legitimate uses for such programs; they create work-in-progress backups that can be useful in the event of power failure or accidental deletions. They can be used to keep track of chat room conversations. They also provide the absent computer owner a level of security, allowing the owner to see if others are using the computer without their knowledge.[71]

Some programs are able to record both online and offline actions. In online recording mode, they detect that the victim is online, record every keystroke, and email the log to the attacker at regular intervals. When in offline recording, everything typed after Windows starts up is recorded and saved on the victim's disk, to be later collected by the attacker.

Because key loggers use very little disk space, they are difficult to find. They can masquerade as important system utilities, making them difficult to identify. Some key loggers also highlight passwords found in text boxes with titles such as "enter password" or just the word "password" somewhere within the title text.[72]

Evidentiary Value

Because several key loggers use email as a way of sending the collected logs back to the individual who planted the program, it may be possible to extract email destination information from the key logger program. If it is password protected, this task may be difficult and even if the traceback is successful, the address will most likely be an anonymous email service.

Rootkits

Description

A powerful mechanism used to hide activity on compromised systems is known as a "rootkit." A rootkit is typically a suite of programs that are used by a cyber criminal to cover up any evidence of an intrusion, by replacing system commands that would

---

[71] Apel, W., Protect Your Computer From Unauthorized Access, PC World Magazine, May 2000, (Available at http://www.pcworld.com.eg/protect_may2000.htm).

[72] Maniac, and Raven, Computer Trojan Horses, Black Sun Research Facility, March 11, 1999, (Available at http://packetstorm.decepticons.org/papers/virus/trojans.txt).

normally be used to reveal the intrusion. They are also used to hide trojans, and other applications and data (such as DDoS tools).[73]

A rootkit gets its name not because it is composed of tools to obtain root, but because it contains tools to maintain the attacker's hold on root. The intruder achieves invisibility by relying on an administrator to trust the output of various system programs.[74] As an example, on a Unix system, the administrator will trust the *ps* command to display all running system processes, and *ls* to list all files on the system.[75] As an example, *ls* (a listing command) is altered in such a way that it will not display the files added by the attacker. The *ps* (a running process listing) is modified not to display the processes that may be running attack commands. By replacing these system utilities with the revised versions found in the rootkit, these commands will not provide the system administrator with an accurate picture of the system, because it will not display the activities or added files of the attacker.

To replace these programs, the attacker must already have root access. In order to get to that point, they have found a vulnerability (possibly through a port/vulnerability scan), and launched a successful attack against the system (exploit script). This attack has given them root, administrator, or 'super-user' access. But, once this level of access has been achieved, they will want to ensure their ability to return. He leaves a backdoor in order to avoid the necessity of using the same exploit again, which may be patched the next time he returns to the system.

Various versions of rootkits are available at many hacker sites. The most accessible versions are for open-source operating systems such as Linux and FreeBSD. Also commonly reported, are versions for Irix, SunOS, and Solaris.

Evidentiary Value

Much like the packet sniffer programs, the detection of the rootkit only provides supporting evidence of a system compromise, but the trojanned applications and utilities provide no evidence as to the identity or source of the attack.

The act of trojanning these system utilities will effectively destroy any evidence of the intruder's actions on the system. This can prevent a thorough investigation of the incident, and make it impossible to collect usable evidence.

---

[73] Pedestal Software, Intact Integrity Protection Driver, (Available at http://pedestalsoftware.com/intact/iipdriver.htm).

[74] The italicized items in subsequent section refer to common Unix commands and locations of system logs.

[75] Brumley, D., Rootkits - How Intruders Hide, Theory Group, (Available at http://www.theorygroup.com/Theory/rootkits.html).

Trojan Horse Programs

Description

Today, there are more than 600 known trojans on the net, with the possibility of many, many more. Well-known hacking groups regularly release new versions of their own signature trojans, and commercial software sites continue to release new products that are marketed as 'Remote Administrative Tools' that have the same basic functionality as many trojan programs.

A trojan horse program typically falls into one of the following categories:[76]

1. Legitimate application designers will often insert unauthorized instructions within their products, as either a backdoor mechanism[77], or as a way of collecting personal information about the users of their product. These instructions perform these operations without the knowledge or permission of the user.
2. A legitimate-appearing program that has been obtained from a questionable source is altered by the placement of unauthorized instructions within it. These instructions perform secondary functions unknown to the user.
3. Any other program that appears to perform one operation or function but that, because of the unknown instructions within it (by design), performs functions unknown to the user.

The typical mode of insertion of the trojan involves an attacker sending the victim a file that, when run, fools the target into believing it is something that it is not. When the victim runs the executable, the trojan installs an additional component on the target, a component that the victim will have no idea exists.

If it is the server portion of a Remote Administrative Trojan (RAT), it installs itself, opens a specific port, and listens for attempts to connect on that port. RATs communicate like any client and server. The victim runs the server, the attacker sends commands to the infected server with his client, and the server follows whatever directions the client gives it. The attacker has all of the same rights and privileges as the victim on that system. The attacker can relay proprietary information out of the system via e-mail or file transfer, or take full control over the system, leaving the legitimate user powerless.

Most victims assume that, if after running an executable the computer is still working with all data still available, no damage has been done. If it had been a virus, their data would be corrupted, their computer would have stopped working, or there would be some other indication of a virus infestation. The victim is aware that there has been an attack, and can begin repairs. On the other hand, the trojan is a tool with a long useful life, because it will run in the background, and perform its functions without giving telltale indications to the user.

---

[76] Maniac et al.

[77] For maintenance purposes.

Evidentiary Value

Note: most of this analysis must be conducted prior to the system being taken offline.

Once the presence of a trojan on a system has been established, the first step is to type netstat -n. This may provide the investigator with the IP address of the attacker, making it possible to trace them back to where they have come from.

To trace an infiltrator back to his/her source, there are online resources that may aid in identifying the source of the connection. The site http://www.samspade.org may reveal information about the intruder, including the administrator of their ISP, just by entering the IP address gleaned through the use of the netstat command.

If a trojan horse program is found on a computer, it is important to determine how it was placed there. If it was sent via email, examining the header information of the email message may provide clues as to who sent it, assuming it was not sent anonymously. The investigator can find out where the information was being sent each time the user goes online by checking ports that are open on the computer, and what IP addresses they are connected with.

Computer Incidental to Cyber Crime

Steganography Tools

Description

Steganography[78] is the science of hiding the existence of a message. It is typically used to describe the hiding of information within other information. This is not to be confused with cryptography, which is generally concerned with protecting the secrecy of the content of a message. While hidden or stegoed images do not need to be encrypted, using encryption adds an extra layer of security if the message is discovered.[79,80]

Modern steganography takes advantage of the fact that most computer files contain unused or insignificant areas of data and uses these spaces to hide information. Once a message is hidden within an innocent looking file, a picture, for example, the file can be sent. The covered message will now appear to the casual observer as an innocent exchange. Only the sender and receiver know that a secret message has been communicated, even if a third party intercepts the message. For example, an image of a family portrait could conceal a private letter to a conspirator or a digital audio clip of a song might contain a company's plans for a hostile takeover.[81]

---

[78] The word steganography is of Greek origin, and literally means "covered writing."
[79] Kahn, The History of Steganography, Page 1, 1996.
[80] Petitcolas, The Information Hiding Home Page.
[81] Milbrandt, What is Steganography?

"Several commercial and freeware programs offer steganography, either by themselves or as part of a complete communications security package."[82] The technology does have legitimate uses. Proprietary graphics, images, sound, and video files or documents can receive a digital watermark to establish ownership and to deter "image piracy" on the Internet.[83]

Aside from the ability to hide covert information within other file formats, steganographic applications are available that purport to encrypt complete partitions steganographically under Linux. This means that the data cannot be recovered without the correct pass phrase, and that no one can prove that any data exists on a Steganographic File System (SFS)[84] encrypted partition. The steganographic file system accomplishes this by creating random information on the device, and then hiding the actual information inside this information.

Steganography does not just scramble information like cryptography does. When cryptography is used, there remains evidence that a file exists, though the contents of that file may be illegible. With steganography, the information is hidden inside of another file. Potential evidence remains virtually unobtainable. Without the correct program to unhide the information, or having the original cover image that was used before embedding (for comparison purposes), there is no indication that the file is anything other than it appears. Steganographic images have a great capacity in which to hide contraband images or illicit data.

To date, there have been in excess of 100 such tools identified for use as a means of hiding information within various types of files. Any of these tools can be used in the commission of a variety of crimes, such as information warfare, industrial espionage, and the exchange of child pornography.

Steganography is becoming increasingly important as governments seek to limit the use of cryptography. In certain countries, the use or possession of encrypted files is against the law. Where this is the case, steganography can be used to replace or conceal the use of cryptography.

Evidentiary Value

To date, there has been little or no way available to law enforcement to identify steganographic carrier files, much less separate an embedded file from the carrier.

---

[82] Schneier, B., Crypto-Gram Newsletter, Counterpane Internet Security, October 15, 1998, (Available at http://www.counterpane.com/crypto-gram-9810.html).

[83] Mendell, R., Steganography - Electronic Spycraft, September 20, 2000, (Available at http://www.earthweb.com/article/0,,10456_624101,00.html).

[84] StegFS - A Steganographic File System for Linux. (Available at http://www.mcdonald.org.uk/StegFS/).

Encryption

Description

Encryption is any procedure used in cryptography to convert plaintext into ciphertext. This procedure is done in order to prevent anyone except the intended recipient from reading that particular data. There are many types of data encryption, and they are the basis of most network security procedures. Two of the most common types include Data Encryption Standard and Public-Key Encryption.[85]

1. Data Encryption Standard - A product cipher that operates on 64-bit blocks of data, using a 56-bit key.
2. Public-Key Encryption - A type of encryption where each person gets a pair of keys, called the public key and the private key. Each person's public key is published while the private key is kept secret. Messages are encrypted using the intended recipient's public key and can only be decrypted using his private key. This is often used in conjunction with a digital signature. Diffie and Hellman introduced Public-Key Encryption in 1976.

Evidentiary Value

Digital evidence is easily modified. Criminals routinely hide evidence from storage media using encryption or freeware/commercial utility programs.[86]

The creation and eventual widespread use of encryption applications poses challenges to law enforcement. Criminals are using encryption more and more to hide their activities. While investigators have a variety of tools used to collect electronic evidence of illegal activity, these tools will be virtually useless when encryption is used to scramble the evidence. Therefore, law enforcement cannot decipher it in a timely fashion, if at all.[87]

Secure File Deletion Programs

Description

With normal file deletion, the first letter of the filename is changed, and reference to the file is removed from the File Allocation Table. This allows the disk space to be reused when new files need to be saved. But, all of the information contained in that file is still present on the storage media after deleting it. However, the data is in unallocated space, and is not readily accessible. It will remain until the disk space is reallocated, and written over by a new file.

---

[85] Noesis, Introduction to Encryption, (Available at http://www.digitalnoesis.com/resources/encryption/cryptointro.shtml).

[86] Champlin, L., E-Commerce Legal Issues Can Ensnare Unwary Merchants, The Business Journal, March 24, 2000, (Available at http://kansascity.bcentral.com/kansascity/stories/2000/03/27/focus2.html).

[87] Cyberspace Electronic Security Act Fact Sheet, Center for Democracy and Technology, September 16, 1999, (Available at http://www.cdt.org/crypto/CESA/CESArevfactsheet2.shtml).

"Information in files that are deleted, but not "secure deleted" can often be recovered and/or viewed using simple tools like the "undelete" command in DOS, or common disk utility programs."[88]

Secure file deletion programs delete files in a totally different manner. Instead of mere deletion, these programs wipe disk storage beyond recovery by scrambling the file name and associated dates, effectively removing it from the disk. These programs can completely destroy any data from previously deleted files that might still be accessible on the disk. To accomplish this task, these utilities will destroy the files from the Recycle Bin, will wipe the free disk space (this space usually contains data from previously deleted files), and will also wipe the slack portion of existing files.[89]

Some can clear the contents of folders that usually contain cookies and other important data, such as the web browser cache, swap files, locked files, temporary Internet files, the recent document list, etc. Others erase entire folder structures, even entire drives. The features of these tools may include the ability to automatically perform erase operations from batch files or scheduling software, password protection, optional confirmation, logging support, etc.

Evidentiary Value

These programs are purported to defeat all types of forensic software. The only way that data may be recoverable after the use of one of these programs is with the assistance of an electron microscope.

While the discovery of secure deletion software on a suspect machine may provide a basis for suspicion, it has no output (associated files or logs) that may be recovered and used as evidence. The tool has one purpose, to wipe all of the free space on magnetic storage media.

Concluding Remarks

Cyber weapons have, and will continue to pose a significant threat to the Internet, and all users of networked computers. Based on our assessment of past and present cyber weapon performance, these weapons appear to pose as significant a threat today as their earlier counterparts. The tools exploit inherent design weaknesses in network and computer procedures and protocols. It is impractical to eliminate these weaknesses, as it would require a major reimplementation of the basic infrastructure upon which the technologies have been built.

The tool authors continue to release updated versions of their tools through any number of relatively easy to locate Internet sites. These authors get a significant amount of

---

[88] Nuker Details, Genio USA, (Available at http://www.geniousa.com/nuker_product_details.htm).
[89] Cyberscrub Overview, Secure File Deletion/Internet Privacy Utility, 2000, (Available at http://165.121.190.90/page2.html).

development assistance from their 'peers', as many cyber weapons are developed in open source project environments. Peers offer suggestions as to current design improvement, and additional features, that would make the weapons easier to use and more powerful; they may even port the applications to other operating systems, making them that much more widely available.

Appendix A of this paper provides the investigator with a sampling of versions of current tool categories. The list for each may be short, but this is not to suggest that there are not many more such tools available. For many, there may be literally dozens more such tools. The tools listed in the appendix are representative of the tools and the capabilities of tools within each classification.

To address the investigative challenge that the use of these weapons pose, tools are available that aid investigators in their analytical tasks. The following section describes tools used in the cyber forensic process, tools used to analyze suspect systems for telltale signs of cyber weapon usage and criminal abuse.

Task 2: Assessment of Cyber Crime Technologies Available to Law Enforcement

The following section is a discussion of the types of tools that are available to law enforcement for the investigation of cyber crimes.

Introduction

The purpose of this task is to identify technology-related tools, methods, and information that are presently being used by, or may otherwise be available for use, by law enforcement agencies in investigating cyber crimes.

The use of a computer to create and store information leaves behind "electronic footprints" that can actually make or break a criminal case. Sensitive data such as e-mail, documents, temporary files, passwords, time and date stamps, and other potentially valuable information are written to remote locations on computer hard disk drives and floppy disks as part of the normal operating process. Most perpetrators are unaware that such information exists, and, therefore, are extremely careless in covering their tracks.

The tools, technology, and software that are currently available for use in uncovering "electronic footprints" are described here. The techniques employed and, where, appropriate, the limitations of these investigative aids are identified. It is important for investigators to know how these tools work and what they can and cannot expect them to do.

Computer forensic specialists provide the legal profession with services that allow them to use the seized computer or computer data in court or in their discovery process. These specialists follow strict guidelines in order to provide acceptable data to the court. Descriptions of those guidelines are outside the scope of this paper. However, all reputable computer forensic investigators follow some basic practices for the preservation of crime scene evidence.

Tool Selection and Assessment Criteria

The types of tools that are discussed were chosen based on our research of currently available technologies that purport to have features and capabilities that may be beneficial to cybercrime investigators. To complete our research, we have interviewed law enforcement personnel, monitored newsgroups and list servers, reviewed developer and vendor information, searched the Internet, and utilized selected tools. While there are many tools currently available, the majority of tools are not widely used by law enforcement.

We have attempted to identify the majority of forensic tools that are currently available and used by cyber crime investigators. But, the field is evolving, changing, and advancing at a brisk pace, and we anticipate that the number of tools that are available at the time of this writing will not reflect the tools that will become available in the near future.

Cyber Forensic Investigation Methodology

Goals of an Investigation

The current methodology in the investigation of a computer suspected to have been involved in a crime includes:

- Identify sources of evidence;
- Preserve evidence;
- Extract evidence;
- Examine/analyze evidence;
- Organize/report results.

This applies to the investigation of traditional, as well as non-traditional crimes. In the event of a system attack or compromise (a non-traditional crime), it is desirable to gather additional evidence that will allow the investigator:[90]

- To understand how the intruder is entering the system, and possibly where the attack may have originated;
- To gather as much evidence of the intrusion as possible;
- To ensure that all applicable logs and evidence are preserved;
- To discover why the intruder chose the computer;
- To obtain information that may narrow the list of suspects;
- To obtain the information to justify a trap and trace of the phone line the intruder is using or to procure a subpoena to obtain information from an ISP.

To this end, law enforcement currently uses tools that fall into the following categories.

- Computer Forensic Tools & Techniques:
  - Evidence Collection & Preservation Tools
  - Evidence Extraction Tools
  - Evidence Examination Tools
  - Evidence Organization Tools

- Incident Forensic Tools & Techniques:
  - Statically Linked Binaries

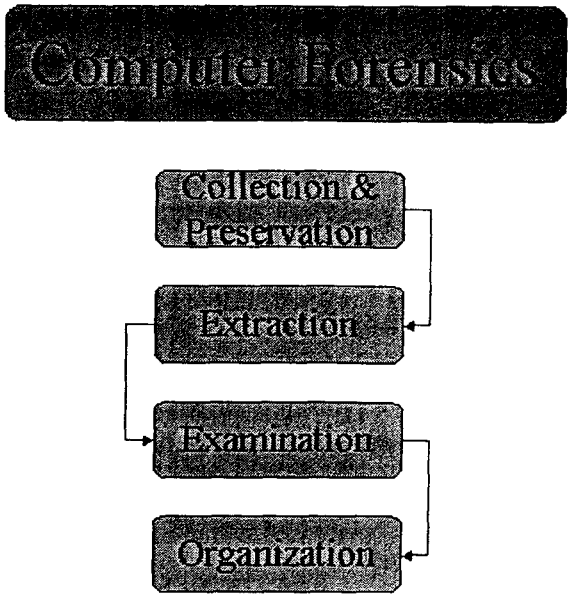
---

[90] Stephenson, P. "Investigating Computer Security Incidents." (Available at http://www.cerias.purdue.edu/secsem/presentations/11-8-2000.pdf).

- Network Forensic Tools & Techniques:
  - System and Firewall Log Analysis
  - Intrusion Detection Tool Analysis
  - Trace Back Tools

- Honeypots
- Trusted Time Stamping

Computer Forensics



There are four basic steps that are taken in a computer forensic investigation. These steps include collection and preservation, extraction, examination, and organization. These steps are used in the investigation of traditional crimes, as well as post mortem investigations of systems that have been compromised.

During the collection and preservation step, the investigator is concerned with the retrieval of electronic evidence, and maintaining the integrity of the evidence they have retrieved. This can be done by using imaging tools to take a bit stream image of the suspect drive, properly documenting the chain of custody, using write-block technologies to assure that the information on the drive is not altered, and creating and maintaining a hash library of the evidence files.

After the collection and preservation of the evidence is complete, it is time to extract the information that is needed for the investigation. Data is extracted from remote areas of the storage media, such as swap space, slack space, and unallocated memory to retrieve swap files, deleted files, hidden files, and temporary cache files. Known file filtering techniques are employed to eliminate common files, and reduce the amount of data to be examined.

Once the data has been extracted, the investigator must now examine all of the filtered data. File listing utilities, file-type identification, keyword searches, and file image/viewers are used to sift through the vast amounts of data. This is the step in which information with evidentiary value is identified.

The evidence now requires logical organization, and must be provided in a format suitable for use in a court of law. Utilities are available that assist the investigator in the organization of their data in easily to read reports, charts and graphs.

These four steps in the computer forensic process are described in great detail throughout the paper.

Evidence Collection and Preservation

The first step in the cyber forensic process is the collection and preservation of electronic evidence. The investigator, before determining the appropriate tool to use for this phase of the investigation, must determine the following:

- What type of computer is it?
- What is its primary purpose (workstation, web server, network server, ISP server, etc.)?
- What operating system is running on the computer?
- Is the computer password-protected?
- What is the size of the computer's memory?
- Is the computer networked or stand-alone?
- What peripherals are attached?
- What, if any, external storage media may contain additional information (floppies, CDs, tapes, etc.)?

By answering these basic questions, the investigator can systematically begin to eliminate from consideration those forensic tools that do not apply to the target evidentiary media.

Investigative Considerations

The investigator must take care to maintain the integrity of the evidence. Electronic evidence is very fragile.

> "Evidence is usually in the form of data fragments and it can be easily overwritten by something as simple as the booting of the computer and/or the running Microsoft Windows. When Windows starts, it potentially creates new files and opens existing ones as a normal process. This situation can cause erased files to be overwritten and data previously stored in the Windows swap file can be altered or destroyed. Furthermore, Windows has a habit of updating directory entries for files as a normal operating process. As you can imagine, file dates are very important from an evidence standpoint." [91]

Another concern for the investigator is the choice of what media will be used in the duplication process as the storage media for the mirror image of the evidence.[92] It is this mirror image on which the investigators will perform their forensic examination. This

---

[91] Anderson, M.R. "Computer Evidence Processing The Third Step - Preserve the Electronic Crime Scene." (Available at http://www.secure-data.com/art7.html).

[92] See Appendix F for a detailed discussion of imaging media considerations.

leaves the evidence in its original state, free from questions of alteration during the investigative process.

These choices include hard drives, tape drives, optical drives, CDs, and DVDs. All types of media choices have their advantages and disadvantages. A careful examination must be made to see which one will provide the most effective medium for storing evidence in both short and long-term situations.

Disk Imaging Considerations

The most immediate objective of the collection and preservation process is the imaging of the evidentiary device. In order to produce a 'forensically sound'[93] image of the computer media, the tool must be able to create a true bit stream copy of every sector of the electronic storage media, without regard to the content (essentially, it must duplicate from the first sector to the end of the physical device). Also, the tool must not make any changes to the original media during the imaging process, and the image copy must contain identical content to the original. "Any additional data in the copy occurring as a result of differing drive geometries must be exclusively HEX 00."[94]

> "Ideally, at least two copies of the data should be taken off computer (or another device that has been identified as a possible source of evidence). One of these is sealed in the presence of the computer owner, and then placed in secure storage. This is the *master copy*, and it will only be opened for examination under instruction from the court. This may happen in the event of a challenge to the evidence presented after forensic analysis on the second copy. If, however, the computer itself is seized and held in secure storage by the police, this will constitute "best evidence." Otherwise the *master copy* will become best evidence.[95]
>
> It would seem on the surface that this approach would be sound and complete. However, the key point is it is not the backup media itself that needs protecting, but rather the integrity of the content that needs ensuring and protecting.
>
> In this way, even if the original source of the evidence is lost, stolen, becomes defective, or is destroyed, the integrity of the backup or the copy being used can be proven. Furthermore, it should be possible to prove the integrity of only a portion of the copy, if it is just this subset that contains the relevant evidence. If only a small part of the computer evidence needs to be accessed in court, a high degree of confidence will be afforded if the

---

[93] For our purposes, 'forensically sound' means every bit or byte in the data area on the original evidence media is accurately reproduced, not just 'most' of the bits and bytes.

[94] Holley, James. "Computer Forensics." (Available at http://www.scmagazine.com/scmagazine/2000_09/survey/survey.html).

[95] Bates, Jim. "Data Integrity Verification and Authentication (DIVA)." (Available at http://www.foresic-computing.com/archives/diva.html).

integrity of a specific file or directory can be proved as well. In order to accomplish this, a secure method of verifying the integrity of the copy is required, even in the absence of the original."[96]

Once a tool or series of tools is procured that can accomplish the required functions, the rest is up to the knowledge, skills and abilities of the investigator.

Evidence Collection and Preservation Tools

An investigator at a crime scene must be able to identify computer systems or media possibly containing digital evidence relevant to the case. Devices that may contain evidence include, but are not limited to:

- Workstation Computers
- Off-Site Computers (Laptops, Notebooks, Home Computers, Senders and Recipients of E-mail, PDAs, etc.)
- Removable Storage Devices (Zips, Jaz, Orb, Floppy Diskettes, CDs, Sony Memory Sticks, Smart Media, Compact Flash, LS - 120, Optical Disk, SyQuest, Bernouli, Microdrives, Pocketdrives, USB Disk, Firewire Disk, PCMICA.)
- Network Storage Devices (RAIDs, Servers, Sans, Nas, Spanned, Remote Network Hard Drives, Back-up Tapes, Etc.)

Upon entering a 'crime scene,' the investigator may be faced with any or all of these devices or media, including the possibility of multiples of each. The investigator would ideally preview the media or systems, determine which ones may have relevant evidence, and preserve evidentiary images only of those systems deemed to contain relevant evidence.

Previewing Electronic Evidence

Upon entering the electronic crime scene, the investigator may be faced with instances when the seizure, or the collection, of the entire contents of all electronic devices present is not possible. Several networked systems may be involved, making seizure impractical. In such a case, conducting an initial 'preview' of the contents of the media aids in the identification of those systems that would most likely contain the sought after information.

"It may not be simple to determine which systems contain information of evidentiary value. In cases where an office has multiple computers, and only a few of those computers contain relevant evidence, then seizing or

[96] Hosmer, C. "Using SmartCards and Digital Signatures to Preserve Electronic Evidence." (Available at http://www.wetstonetech.com/digital.htm).

imaging all the computers at that site could be a tremendous waste of time and resources."[97]

There are good and bad aspects associated with previewing, and the potential copying, of electronic evidence. However, wherever possible, it should never be used in place of imaging. The preview option is used only to get quick results that could potentially lead to new evidence.

A preview of the contents of a suspect's computer may allow the investigator to glean additional evidence from the crime scene. That is, it may aid in the identification and seizure of additional 'paper-based' material from the immediate area of the computer. This new information can only be used if a link can be established that the newly found information is related to the case.

One disadvantage to the preview method is that after a preview of a drive has been made, and an investigator does not immediately identify any relevant documents on that drive, an image of the target drive may be disallowed, due to that fact that the drive does not contain any relevant case material.[98]

Once the preview process has determined that different computer systems or media contain information relevant to the ongoing investigation, an investigator must have the tools capable of performing a forensically sound image of those systems or media, in a way that does not alter the original evidence in any way. This step is essential, especially if the evidence is to be used in subsequent legal proceedings. The investigator must be able to prove that all evidence retrieved is an exact, complete, unaltered duplicate of the original.[99]

Refer to Appendix D for further discussion of evidence-previewing tools. This discussion introduces the reader to remote previewing features of EnCase and alerts the reader to possible ramifications of the use of these features.

Disk Imaging Tools

Software Imaging Tools

For the collection phase, bit stream disk imaging software is used. These applications are used to create 'true mirror-image'[100] backups of individual partitions on hard disks drives, and/or to make a true mirror-image copy of an entire hard disk, which may contain multiple partitions and/or operating systems. Imaging tools are also used to create exact duplicates of portable disks (floppy disks, zip disks, etc.) and other media. The

---

[97] Holley, J. "Meeting Computer Forensic Analysis Requirements." (Available at http://www.scmagazine.com/scmagazine/sc-online/2001/article/016/article.html).
[98] Earnshaw, C. "Re: Preview in Encase (or other package) rather than image." (Available at http://lists.jammed.com/forensics/2001/07/0008.html).
[99] Holley, J. et. al.
[100] A true mirror image is a sector-by-sector copy of every sector of the original media.

tools must possess the capability of collecting a multitude of evidence types from a variety of sources.

Most imaging applications are operating system independent, as they run under MS-DOS, and access the computer media at a purely physical level.

Several software tools on the market advertise they make true bit stream images[101] of hard drives. The selection of tools available to the investigator to perform this function is varied, all having certain advantages and disadvantages. Some of the tools are well suited for certain circumstances, but are ill suited for forensic purposes. They are not designed from a forensic standpoint, leaving the validity of copies in question. Many tools have the ability to operate only on limited media, or on a narrowly defined environment.

The current capabilities that can be found among the choices of true imaging utilities are:[102]

- All areas of the hard disk drive can be copied.
- Duplicate copies of hard disk drives can be made from hard disk to hard disk in direct mode.
- Image files can be stored as one large file or separate files of fixed sizes; this is helpful in making copies for archive on CD's.
- Copies can be made in either physical or logical mode at the option of the user.
- Multiple partitions containing one or more operating systems can be copied and restored.
- Date and time stamped audit trails are maintained by the application, keeping a record of operations during an imaging session.
- Drives can be accurately copied and restored.
- Bit image copy can backup and restore (see Image Restoration) to:
    - Another internal or external hard disk
    - Hard drive to tape
    - Tape to tape
    - SCSI to SCSI
    - IDE to IDE
    - SCSI to IDE
    - IDE to SCSI

Limitations:

- *Speed* - the speed of the imaging process can vary greatly based on a number of factors: physical state of the media, the processor, interfaces.
- *Size Limitations* - the target media for the image needs to be larger than the evidentiary media that is being copied.[103]

---

[101] Other terms, such as cloning or mirroring are often used in association with hard drive duplication. While these tools are useful for backup of data, they often do not collect the entire contents of the drive, i.e. the unused space between active files, where additional evidence may be hidden.

[102] SafeBack Mirror Image Backup Software. (Available at http://www.forensics-intl.com/safeback.html).

- *Platform Requirements* - most of the tools will work only on certain Windows and/or *nix distributions.
- *Cost* - the price of these tools may be prohibitive for most law enforcement cyber crime units.
- *Hardware Differences* - see Appendix F for a discussion of media problems that may affect the imaging process.

Refer to Appendix D for further discussion of software imaging tools. The software imaging section of Appendix D identifies and lists five specific software imaging tools that clone or image files from hard drives and also collect ambient data. The software imaging tool descriptions provide information on tool capabilities, how they function, and the manner in which they ensure data integrity. This section of Appendix D also describes some limitations of the respective tools that investigators should be aware of.

Hardware Imaging Devices

Hardware imaging devices are much faster than their software counterparts. They enable the copying of a suspect hard disk onto some other storage media, typically another clean hard disk. Vendors claim copying speed that can exceed 1 GB per minute, depending on certain characteristics of the suspect drive, such as age or condition. The hardware devices range from the handheld units, to complete desktop PCs designed specifically for use as a forensic workstation.

Handheld Forensic Devices

Handheld devices are small, lightweight, durable units, used as portable forensic imaging devices. Aside from their imaging capabilities, they provide additional functions, including:

- Wiping of the target storage drive, if it has not been previously wiped;
- Preventing unwanted writes to the evidentiary drive;
- Scans the destination drive to ensure that there are no bad sectors;
- Provides printed reports, when attached to a portable printer;
- Audit trail reporting.

Some known limitations of the handheld devices are:

- They may have limited data authentication or validation;
- They normally make only one copy at a time;
- Copies are made onto hard disks, which are easily damaged;

---

[103] One notable exception is Byte Back: if the target media is smaller than the image file, Byte Back will prompt the user for additional media, automatically breaking the image down to individual volumes. It also allows the user to pre-set the volume sizes, i.e. if they wish to use CD-R to store an image, a 640 MB volume size can be selected. This allows the use of CD-R software to burn the individual volumes created by Byte Back.

- The devices require additional adapters for imaging SCSI drives;
- They cannot obtain images through the parallel port.

Forensic Workstations

Several vendors of forensic software now provide their own versions of forensic workstations as complete solutions to the investigator's evidence processing requirements. These can be either laptops or PCs that come installed with the vendor's proprietary software, and are configured specifically for a complete forensic analysis.

These workstations provide the speed and the processing power to produce an exact copy of the original suspect machine from the bit-image copy. This is then analyzed using the proprietary forensic software that is installed on the dedicated analysis drive. Also, the reconstructed drive can be booted on to the workstation to replicate the performance of the original machine.[104]

Each workstation comes with a variety of optional features tailored to the needs of forensic investigator, which may include any or all of the following:

- ZIP or JAZ drives
- A tape drive
- A variety of preinstalled operating systems
- A high-speed SCSI card and cable
- A CD ROM reader/writer
- DVD reader/writer
- Removable hard drive racks

Image Restoration Tools

Applications that support image restoration are capable of restoring the captured image of the evidentiary hard drive to a hard drive that is *identical* in physical geometry to the original. The restored image is digitally identical to the original. If identical media is not available, then any additional area in the restored image occurring as a result of different drive geometries is zeroed. The applications may restore stored images from tape or disks.[105]

> "Where the functionality of an application is in question, just analyzing the files comprising the application at a physical or logical level is not sufficient. There may be the need to run applications (executables) that have been preserved as evidence, generally for the purpose of examining files that have been created with this particular application. This cannot

[104] The DIBS® Forensic Workstation. (Available at http://www.computer-forensics.com/products/welcome.html?workstation.html).
[105] Holley, J. "Computer Forensics." (Available at http://www.scmagazine.com/scmagazine/2000_09/survey/survey.html).

currently be done from within the image files, so the image must be restored."[106]

Refer to Appendix D for a further discussion of image restoration tools. This section identifies several of these types of forensic utilities.

Imaging Validation Tools

The purpose of these utilities is to reassure the analyst that the bit-stream image from the evidentiary computer is the same as the original drive. This is determined by using mathematical formulas called 'hash functions', which are used to calculate hash values of both the original media, and the imaged copy. A valid copy has been produced when the two calculated hash values are identical.

Forensic tools either incorporate validation techniques as an additional function of the toolkit, or provide it as a standalone utility as part of a suite of tools.

The validation of the evidentiary copy may be done in one of two ways. A message digest (or signature) may be created of the original drive, and then compared to one created on the image. The utilities are used to mathematically create a unique signature for the entire contents of a computer hard disk drive or other media. The signatures can then be used to validate the integrity of forensic bit stream images made during the evidence collection process.

Or, an integrity check may be run in parallel to the collection process. For example, a Cyclical Redundancy Check (CRC) is run every time a specified number of sectors are reached during the collection. This CRC value of the image copy is then compared to a CRC value of the original media. If the values match, the process continues. If not, an error has occurred during the processing, and the imaging utility must be restarted. Using this type of error checking ensures that the image that is created is an exact duplicate of the original media.

MD5 hashes, Cyclical Redundancy Checks (CRCs) and digital signatures are the most common method used today to verify the authenticity of evidence.[107] Digital signatures are by far the preferred method, because they provide a much higher degree of cryptological certainty.

Refer to Appendix D for further discussion of imaging validation, describing the imaging validation capabilities of three forensic tools.

---

[106] Holley, J. "Meeting Computer Forensic Analysis Requirements." (Available at http://www.scmagazine.com/scmagazine/sc-online/2001/article/016/article.html).

[107] See Appendix G for a detailed discussion of these procedures.

Write Protection/Write Blocking Tools

"File viewing, and most file operations, will attempt to alter the last access date of a file." [108] And once this date has been altered, any evidentiary value that the file may have contained is now rendered useless. A defense attorney can now argue, convincingly, that the contents of that file may have been changed during the course of the investigation. For this reason, great care must be taken to prevent this from happening.

Write blocking utilities are used to prevent inadvertent changes to file attributes on an evidentiary system. When a write blocker is used, it must not allow the last access date to change, as maintaining file dates and times, and thus the integrity of the evidence, is critical to any cyber crime investigation.

Basic requirements for write blockers are:

- They must always prevent writing to the device they purport to protect.
- They must not alter the content of the device they purport to protect.
- They must permit unaltered examination of the content of the device they purport to protect.

Certain problems exist with non-Windows operating systems such as NT, UNIX, Novell, etc., due to the facts that write blocking programs do not work with these systems.[109]

Refer to Appendix D for further discussion of hardware write protection/write blocking tools. This section of Appendix D is divided into write protection/write blocking software and hardware tools. The software subsection offers information on the functions and capabilities of five specific write protection/write-blocking tools along with tips on how they are operated. Similar material is provided on three hardware devices.

Hardware Write Blockers

These devices are hardware-based adapters, and thus work at the physical level to block writes to IDE drives. These adapters simply accept all 'write' commands, but fail to act upon them.

Refer to Appendix D for write protection/write blocking hardware devices such as Acard, Daten-Airbag and FastBloc.

System Time Recognition

The reliability of the file dates and times depend on the accuracy of the system settings for date and time on the subject computer. Therefore, it is important to immediately

---

[108] Mares, D. "What Time Is It?" (Available at http://www.dmares.com/maresware/time.htm).
[109] Flax, Jeff. "Understanding the Computer and How Child Pornography Cases are Made." (Available at http://www.dcfpd.org/2000seminar/flax.pdf).

document the accuracy of the system clock as soon as the investigation begins. The correctness, or incorrectness, of the current system time could be a valuable piece of information when attempting to establish a timeline of events based on file times.

A suspect may have purposely reset the system time in an attempt to throw the investigator off of the track. Time zone differences or day light savings time changes could also be sources of system clock inaccuracies. When reviewing the dates and times that files were created, modified or last accessed, the current system time information is vital.

Refer to Appendix D for further discussion of a system time recognition tool.

Evidence Collection and Preservation Assessment

Although there are many cyber forensics tools available for collecting and preserving evidence from an evidentiary computer, there are limitations to their effectiveness.

- Today's off-the-shelf software backs up evidence slowly, and is prone to errors that often require that the duplication process be repeated.
- Significant technical savvy is required, and in some cases, expensive specialized training by the software vendor is necessitated.
- The backup is not always reliable, and often requires the investigator to use separate software programs to ensure authenticity with cryptographic certainty.
- Backing up large hard drives or multiple computers may be impossible or infeasible, as the data cannot be contained on a single backup device.

Evidence Extraction

Evidence extraction represents one of the most mature technological areas within the forensics arena, though the current tools and methodologies for the extraction process are not necessarily systematic, and approaches to the investigative process are greatly dependent upon the case under investigation.

Evidence Extraction Tools

Hidden Data Recovery Tools

The second step in the cyber forensics process is evidence extraction. The initial task conducted within the area of extraction involves the location and retrieval of certain types of 'hidden'[110] data. These types of utilities selectively seek out and identify areas within the storage media containing one or more of the types of data from the following list:

[110] Hidden data is found in areas of the electronic storage media that are not normally accessible by the operating system; the operating system does not retain an active reference to this data.

- Deleted files
- Slack space
- Unallocated memory
- Swap space
- Temporary files
- Temporary Internet cache files
- Hidden files

Deleted Files

When a user deletes a file, two things happen; the reference to that file is eliminated (erased) from the file allocation table; the first letter of the file name contained in the directory listing is changed to a 'special character' (usually E5 hex). However, the data in that specific storage area remains unchanged until it is overwritten with new data.

Slack Space

Disk storage space is divided into 'sectors'; these sectors are usually 512 bytes in size. When a file is stored, it may not take up the entire 512 bytes of that particular sector. That remaining area sits unused by that file. This unused area is referred to as slack space. Slack resides between the end of the file stored in that sector, and the end of the physical sector.

Unallocated Memory

Allocated memory contains data that is currently 'in use', having a corresponding file entry in the file allocation table. Conversely, unallocated clusters may contain data, but this data is not stored in disk space that the system's file allocation table currently recognizes as being in use. Therefore, although unallocated memory frequently contains residual data, until it is eventually overwritten to store new data.

Swap Files

Windows-based systems utilize a swap file, temporarily allocated space on the hard drive, which is written to when active memory[111] resources are low, to extend the capacity of RAM. This is a file that can hold complete documents, emails or other data that can be of significant interest in an investigation.

Temporary Internet Cache Files

Web browser applications (i.e. Internet Explorer) retain various temporary files, Internet cache, favorites (bookmarks) and history files. These files provide a 'record' of visits to Internet sites. They also keep copies of other files that were viewed on that site, including all graphic files from that site.

---

[111] Random Access Memory (RAM)

Hidden Files

By changing the attributes of a file, it is possible to hide files that common DOS or WINDOWS commands (such as *dir* - directory listing) will not reveal to the user. The hidden file will not be displayed, and the casual user may not find its contents.

Hidden files are those that are traditionally sought during the initial stage of a cyber crime investigation, and the capability to locate and extract this type of information is a core requirement for any forensic tool.

Other Extraction Tools

What follows are those additional capabilities that further the processing of the collected data. Historically, the following tasks were done manually, and were very labor intensive. But, as the capabilities of forensic tools have evolved, many tools now automate one or more of the following functions.

File Identification and Processing

A common technique for hiding a file is to 'change' the file from one type to another by renaming the file, and appending a different file extension. Forensic practitioners need the ability to identify these changes to the file's extension.

The utility will compare the file's current extension (e.g. *.exe*) with the file's actual 'signature' in the file header[112] to determine if an attempt has been made to 'hide' the file. For example, if a file was created as an Adobe Acrobat (*.pdf*) document, and the extension was later changed to *.jpg*, the utility will identify that file as being suspicious.[113] Several forensic tool suites will examine and compare file signatures with file extensions.

Refer to Appendix D for further discussion of file identification and processing tools. The file identification and processing section of Appendix D explains how five specific forensic tools verify file signatures and how the respective tools report mismatches.

Known File Filtering

Examining computer files is one of the most time-consuming and labor-intensive activities performed during a forensic analysis. The challenge is to sift through and eliminate the extraneous data as quickly as possible, leaving only that data that bears

---

[112] Most graphic and text files contain a few bytes at the beginning of the sector that constitute a unique 'signature' of the file. For example, the first 6 bytes at the beginning of a .gif file are either GIF89A or GIF87A.

[113] Computer Forensics International. "The Basics of Digital Evidence Discovery." (Available at http://www.cf-intl.com/evidence_recovery_basics.htm)

further investigation. This should be completed through the use of a technique referred to as known file filtering.[114]

Known file filtering eliminates from consideration those files that are commonly found on systems, including commonly used utilities, tools and applications (as an example, the Microsoft Office Suite). The filtering tool creates a hash value for every file encountered on the evidentiary media, and compares that value previously created hashes of common files known not to have any evidentiary value.

File filtering tools, using a Reference Data Set (RDS) of known file profiles and signatures, can eliminate a high percentage of files from criminal investigations, allowing investigators to concentrate on those files that are not eliminated through the file filtering process.

The National Software Reference Library (NSRL)[115] project has created a database of known file profiles and signatures that can be used as a reference data set in legal proceedings concerning criminal evidence investigation, software piracy, copyright infringement, child pornography, etc. The library provides four different hash values for each application in its library. The algorithms used to produce these hash values are CRC32, MD4, MD5 and SHA-1. A list of products in their RDS is available for download.[116]

Another function closely related to known file filtering involves the comparison of the newly created file hash values to a database of hash values for files that have been predetermined to be illegal (such as child pornography images).

PERKEO[117]

> "The German Federal Criminal Police (Bundeskriminalamt - BKA) have supported the creation and maintenance of German software called PERKEO. PERKEO was developed in order to reduce the time-consuming work of analyzing computers that are suspected to contain child pornography. It produces checksums (comparable to electronic fingerprints) of files that were classified definite child pornography according to German law. The checksums are integrated into a regularly updated database, and used as a basis of comparison when analyzing media seized during an investigation of child pornography. At present, this database comprises about 14,000 checksums of child pornography, and about 4,000 check sums of bestiality files, as the distribution of bestiality is punishable according to German law."[118]

---

[114] Fisher, Gary. "National Software Reference Library (NSRL)." (Available at http://www.itl.nist.gov/div897/docs/nsrl.html)

[115] http://www.nsrl.nist.gov/index.html

[116] http://www.nsrl.nist.gov/inventory.txt

[117] http://www.perkeo.net

[118] Kind, H. "Combating Child Pornography on the Internet." (Available at http://www.asem.org/Documents/99ConfVienna/pa_kind.html).

File filtering programs can also be used to identify the use of specified applications. A search for the presence of these programs could lead to evidence of criminal activity. As an example, the use of a sophisticated graphics program like Photoshop or Illustrator could be an indication that forgery or counterfeiting is taking place.[119]

Refer to Appendix D for further discussion of known file filtering tools. Five specific known file filtering tools are identified and described in this section of Appendix D. Descriptions of the tools include explanations of how the tools extract benign files, how files are flagged for the investigator, and integrity assurance capabilities of the tools.

Special File Formats

When analyzing the data, investigators may encounter files and images that they cannot view. During the initial forensic search, the investigator may not recognize encrypted data, various compressed data formats, password-protected files and steganographic files. The data requires additional processing, as these types of files are not written to the disk in plain text, and search utilities cannot identify text data stored in these file formats. Additionally, various other formats[120] require special translators or viewers in order to be examined. Manual evaluation of these files is required, and in the case of encrypted files and steganographic carriers, much work may be involved. Investigators need to be technically prepared to deal with evidence found in these conditions.

Encryption Identification Tools

Recently, encryption technology has developed rapidly and has become very popular. Encryption plays a role in protecting confidential or personal information. However, criminals may also use encryption to protect their computer records and e-mail communications, which make an investigators job even harder. Before any attempt can be made to open an encrypted file, it must first be identified within the storage media. Without the proper tool to locate an encrypted file, it may appear as random, meaningless characters.

Refer to Appendix D for further description of capabilities of a specific encryption identification tool.

Decryption Tools

There are many commercially available decryption programs that purport to break a variety of encryption schemes. The ability to break a particular scheme is directly related to the length of the encryption key; the longer the key, the more difficult it will be to break.

---

[119] Holland, G. "P C P.D." (Available at http://www.usc.edu/isd/publications/networker/98-99/v9n4-Mar_Apr_99/sidebar-pc_pd.html).

[120] Steganographic carrier files; various graphic, video and audio files; .PDF format files; executable files or binary data files; files housing email archives and/or active email content; swap files or virtual memory files, and other such file formats that obscure their plain text content.

A discussion of decryption technology is beyond the scope of this paper. But, it should be noted that the Digital Millennium Copyright Act may hinder the use of existing tools, as well as further research into developing better tools for law enforcement purposes.

From a legal standpoint, there is serious concern related to the production of admissible evidence from encrypted data. It must be prove beyond reasonable doubt that the decryption method or technique used was the right one, and that it has produced the correct information from the encrypted data.[121] At this time, there are no decryption programs that can accomplish this with any legal certainty.

Given the large number of encryption programs, and the even greater number of encryption possibilities, it will continue to become increasingly more difficult to break encrypted code, and extract admissible evidence for prosecution.

Compression/Decompression Utilities

There are many different applications that will of compress files. Data compression is used quite frequently in backup utilities, spreadsheet applications, and database management systems, to name a few. But, data compression can also be used to hide and/or disguise 'sensitive' data. An investigator must be aware that data may be compressed using any one of the many formats, be able to recognize it, and be equipped with the proper tools in order to decompress and obtain the information contained therein. Assuming the file extension has not been changed, the investigator should be able to identify compressed files based upon the extension of that file.[122]

Refer to Appendix D for further discussion of compression/decompression utilities of two specific forensic tools.

Password Recovery Utilities

Some password-protected files can be manipulated in a manner to remove or expose the password. There are many products and services on the market that claim to 'crack' the passwords for a variety of applications. Vendors of multiple file-type password applications include Access Data,[123] Lost Password,[124] Office Recovery,[125] and Elcomsoft.[126] Their products claim to recovery the passwords for a large number of applications (see websites for listing of all applications).

---

[121] Hong Kong Internet Service Provider's Association. "Inter-departmental Working Group on Computer Related Crime." (Available at http://www.hkispa.org.hk/pdf/ComputerRelatedCrime.pdf).
[122] Guide to Data Compression File Formats. (Available at http://www.diffuse.org/zip.html).
[123] http://www.accessdata.com/
[124] http://www.lostpassword.com/
[125] http://www.officerecovery.com/
[126] http://www.elcomsoft.com/prs.html

No computer forensic products available today currently incorporate password recovery applications within their toolkits. However, this is not to say that the functionality is not available through the vendor as a plug-in to their product.[127]

These utilities vary in price, with the high-end recovery kits approaching one thousand dollars, so they may be prohibitively expensive from a law enforcement perspective.

Refer to Appendix D for further discussion of password recovery utilities.

Steganography[128] Detection Tools

When examining a computer seized as evidence, a law enforcement investigation could be seriously hindered by the possible use of steganography. A suspect, using steganography, could embed evidence in innocuous files, thus avoiding detection.[129] This could a particularly difficult problem when investigating a child pornography suspect, as the only tangible evidence that may be used against him is the possession of the actual images. If these images are 'hidden' within other files, no observable evidence may be located.

In order to prevent this threat, investigators must have access to steganography detection and extraction tools. At this time, very few applications exist in the area of steganalysis. In order to attempt to defeat steganography, investigators must have access to these utilities.

Refer to Appendix D for identification and description of one steganography detection tool that is currently available and two steganography detection tools that are, at the time of the writing of this report, in developmental stages.

Virus Detection Capabilities

Seized computer hard disk drives and floppy disks should be scanned for the presence of malicious code, such as viruses or worms, that could potentially contaminate both the evidentiary media and the analyst's work station. Any viruses found should be documented (the name of the virus, and it's location on the media), and then removed to avoid future threat of contamination.

Current forensic tool suites do not include virus-detecting capabilities. Instead, they require the use of third party virus scanning software.

Refer to Appendix D for further discussion of virus detection utilities.

---

[127] Access Data offers a password recovery product that works in conjunction with their computer forensic product.

[128] A method where a message (either text or image) can be hidden within other files containing text, images, or even sound, without a perceptible change in the original file's quality.

[129] McCullagh, D. "Secret Messages Come in .Wavs." (Available at http://www.wired.com/news/print/0,1294,41861,00.html).

Evidence Extraction Assessment

While this may be the most developed area for forensic processing capabilities, it is evident that there are many areas in need of work within evidence extraction.

The initial extraction capabilities of the forensic tools are quite adequate. That is, locating, identifying and collecting the contents from those portions of storage media that are not accessible by common computing techniques (deleted files, slack and unallocated space). When identified, these areas can be searched for keywords that are relevant to the case at hand.

Once this task has been accomplished, it is at this point where the tools show their weaknesses. Most forensic tools are incapable of identifying those special file formats that may contain additional information related to the cyber crime.

The reason forensic tools cannot yet find many types of data is that they do not have the ability to identify or open files in their logical format in order to view the contents of the file, as with compressed files, or those in the Adobe Acrobat format.

Even those tools that can identify special formats are limited in the types and number that they recognize. Many times, it is up to the investigator to supply additional file signatures to the utility in order to enable the product to seek out these formats.

Once identified, the investigator is then again limited in the viewing capabilities of the individual products. Most will only view a handful of file formats. This leaves the investigator the added burden of seeking out and obtaining file viewers that support the given format. There is an urgent need for expanded universal file type identification.

A broad array of technologies exists in the extraction area, and significant investments, from both public and private sources, have been made in creating these technologies. But, there is limited multi-platform support (i.e. Windows, Solaris, Linux, mobile, and network extraction) within a common class of tool.

Furthermore, there are a limited number of qualified personnel who are skilled in using these tools. Clearly, more tools need to be developed to meet these requirements, and more individuals need to be trained in using both the tools that currently exist and those that are being or will be developed.

Investigators need be aware that encrypted data and various compressed data formats will not allow searches until the data is uncompressed or decrypted. There is a need for forensic tools that can assist in decrypting data, breaking passwords, or accessing protected information contained in electronic organizers, which are becoming more popular. If the data owner refuses to turn over the encryption keys, the investigator is forced to try and "break" the encryption. This type of brute force attack is time consuming, costly, and often doesn't work. Successful brute force attacks depend on the strength of the encryption algorithm and the strength of the password. There are very

limited rudimentary tools available to aid in breaking passwords and encryption algorithm.

If there is any back door access to these devices, the investigators must procure it from the manufacturers and software vendors.

Evidence Examination

Once the digital evidence has been imaged and filtered during the extraction phase, the cyber forensic investigator must refine and further examine what has been collected. During the examination phase, the investigator uses the available tools to target specific digital evidence. A number of forensic tools exist that allow investigators to further control their search for evidence in the storage media.

The examination of a computer should be a methodical process. By doing some initial groundwork, an investigator can save time and make the examination more successful. "Knowing what information to search for in a forensics examination involves a mixture of background investigation, deductive reasoning, and common sense."[130]

> "Depending on the particular crime being investigated, and its relationship with various computer applications, there can be a number of specific files types to look for. As an example, when investigating child pornography, a search should be conducted not only for the graphic images, but any associated communication and transfer programs that might have been used to capture, download, modify, view and produce the image. Programs and files such as e-mail attachments, original compressed files, news & file retrieval agents, browser programs, dial-up information, file captures, session logs and many others can have a wealth of valuable information. Associated computer evidence found in various computer files can and often will reveal the time, date, manner, location, email address, history logs, web site, file transfer location, IP Internet address and other useful information."[131]

Due to the shear size of modern hard-disk drives, it is all but impossible for a computer investigator to manually view and evaluate every file on a computer hard drive. Therefore, investigators need to use specialized forensic tools to locate relevant evidence, and to shorten the time it takes to complete the investigation.[132]

---

[130] Betts, Bill. *Information Security*. March 2000. p. 32. (Available at www.nlectc.org/inthenews/crimeseen.html)

[131] McLean, J. "Basic Considerations in Investigating Computer Crime, Executing Computer Search Warrants and Seizing High Technology Equipment." (Available at http://www.bileta.ac.uk/99papers/maclean.html)

[132] Betts, B. et. al.

Evidence Examination Tools

Evidence examination tools aid the investigator in sifting through all of the collected evidence in order to locate those files that have relevance to the case being investigated. Many of the same search type tools that are used during the extraction phase are used during this phase. The difference is that the investigator has now refined his search, so that it focuses on specific pieces of relevant data or information.

The investigators use their experience and training to search the computer for documents, deleted files, images, e-mail, slack space and unallocated disk space that will provide them with evidence. Additional 'metadata'[133] related to these files must also be collected, such as creation, access, modification and deletion dates, to aid in the creation of a timeline of activity. There are very few tools specifically designed to aid in this process.

File Listing Utilities

A simple listing of all active and deleted files stored on the suspect computer can be of great use to the investigator, as the mere name given to a file can help direct the investigator to an area worth further examination.

In general, these programs can be used to determine the makeup of a computer hard disk drive. The applications typically create a file list upon the completion of the scanning function. All files and directories are included (as well as deleted ones). They can capture drive, path, file names, times, and dates. The investigator can then browse or search for specific file names, extensions, or creation dates.

Certain utilities have options for viewing directory folders and files, including a Windows Explorer-like tree structure view, and a table containing every file in a given case. The information for each file, such as name, attributes, type, size and creation date is displayed.

Refer to Appendix D for further discussion of file listing utilities provided by three separate sources.

Keyword Search

> "When analyzing retrieved information, computer forensic specialists look for keywords and phrases within the stream of data obtained during a search. They are trying to determine if the computer was being used to store important information such as dates, phone numbers, names of contacts, etc., in order to piece together materials and provide evidence to support strategies. The keywords, in many cases, are the words of the street, for example, drug "street talk," arsonist vocabulary, child pornography descriptors, slang phrases, or other criminal language.

[133] Data about data

> In addition to using keywords to find evidence, investigators must also search for words that would affect their ability to examine a document or file, because that file contains information that may be privileged. For example, in the search of a person's house for paper documents that may be incriminating, detectives use care to ensure they do not examine documents that are communications between the suspect and his spouse, attorney or priest. The same care must be taken when examining electronic documents."[134]

Certain keyword search programs provide the ability to search 'in context', that is, the ability to see terms identified by the search, and several words on either side of the term. In this way, the investigator is able to discern the context within which a word or phrase is used, to determine if the usage of the term is relevant to the investigation.

Refer to Appendix D for further discussion of keyword search tools. The discussion contains information on six types of keyword search tools available and the extent of their search capabilities. The description of the keyword search tools also explains the level of versatility of the respective tools (e.g., options to search slack space, abilities to make the searches case sensitive, etc.).

Dictionary/KeyWord List

While all of the previously mentioned tools allow the user to create or import their own custom wordlists for the application, very few come with pre-compiled lists of words and phrases that would be associated with a specific crime. The investigator must take the time to generate the set of search terms. This step can be very time consuming. The investigator's time would better be spent on other tasks.

Few forensic tools ship with dictionaries of words and phrases categorized by predefined cyber crime types[135].

Refer to Appendix D for further discussion of tools used to develop dictionary/keyword lists.

File Extension Searches

A file extension is the string of characters that follows the 'dot' in a file name. These are used to identify the file format to the computer. The computer recognizes the extension, and uses the associated application to open that file. The ability to sort data by specific file formats (identified by the extension) is a timesaving feature, and allows the investigator to quickly locate relevant information.

---

[134] Hosmer, C., J. Feldman and J. Giordano. "Advancing Crime Scene Computer Forensic Techniques." (Available at http://www.wetstonetech.com/crime.htm).

[135] These crime types may include categories such as "controlled substance & drug paraphernalia", "burglary & robbery", "sex crimes", "terrorism" and "gambling".

Often, the analysis focuses on looking for specific types of files. As an example, if an investigator were looking for a letter that the accused criminal may have written, he would do a search on file extensions such as .txt for a plain text document, .doc for a Microsoft Word document, or .wpd for a Microsoft WordPerfect document.

Many tools that do keyword searches also allow the investigator to do searches based on file extensions.

Other Searches

The standard *Find* command permits a search for occurrences of a specific word, such as 'child'. But sometimes, instead of searching for a specific word or phrase, it is more beneficial to search for a certain 'pattern' of characters used in a word or phrase. This pattern could be any word that starts with 'c' and has just four letters; or, any word ending with 'd' and having five letters, and the second and third letters are within the a certain range of letters (such as the first half of the alphabet).

In like manner, the investigator can search for patterns or 'general formats' within files, such as telephone numbers, IP addresses, or credit card numbers. Additional delimiters, such as a country, or a word that starts with a specific letter, can narrow the search. If a pattern can be conceived and described, it can be searched.

These types of searches are known as wildcard searches, regular expression searches or grep-type searches.

Refer to Appendix D for further discussion of tools or utilities that perform other searches.

File/Image Identification and Viewing Utilities

Once suspect files have been extracted, they are now in need of further examination. File viewers that will recognize and allow the viewing of many different file formats are used.

Viewing tool options may include file type selection, file header examination, and file size range. The ability to identify image files quickly is crucial, as images are a very important piece of evidence in child pornography cases, as well as many other types of cyber crimes.

In general, tools in this category have the following capabilities:

- Powerful file-viewing capabilities, supporting over 250 different file formats.
- Displays images of *.bmp*, *.jpg*, *.gif*, and *.tiff* formats.
- Allows a rapid review of graphic files by the user; the user need only set a minimum file size as criteria of the search.

Most computer forensic tools or suites do not include an accompanying or integrated image viewer, and those that do are limited in the number of supported images. Most investigators must resort to third party image viewers during the course of their investigation. Several good third-party viewers are on the market, including QuickView Plus, ThumbsPlus, ACDsee, and IFRANview.[136]

In some cases, a recovered file may have a valid header, but corrupted or incomplete data (a 'partial image'; as an example, 10Kb of an original 80Kb image[137]) can confuse the viewing program. Many viewers will not display partial files. But, for *.gif* and *.jpg*, ThumbsPlus, IRFANview, or QuickView Plus will usually displays partial images.

Refer to Appendix D for further discussion of file/image viewing utilities, as well as other viewing utilities.

Evidence Examination Tools Assessment

Evidence viewing requires the investigator to have numerous software programs to view a large variety of file types in their natural format. There are no tools currently available which act as a universal file format viewer. This makes it difficult for the investigator, especially if the suspect has been using obscure or outdated software to create files and documents. A universal information viewer is needed, as well as utilities that will associate uncommon file extensions with the specific program used to create that file type. This feature would be useful, as it is not possible for an investigator to have every native application used to create every file type.

At this time, there are no standard taxonomies of words, phrases, data formats, or data organization that can be applied to specific crimes under investigation and used to search the data. Without them, data searching is rudimentary and very time consuming.

There is also no tool available that allows investigators to identify possible authors based upon their known writings, i.e. vocabulary, grammar, or style.

While investigators are under an obligation not to examine privileged information, there is currently no way for them to know that what they are about to read is privileged, without actually reading it.

As the use of digital evidence becomes more prevalent in court cases, it will be increasingly necessary to develop cyber forensics tools to meet these needs.

---

[136] See Appendix H for additional information about utilities.

[137] The original image size was 80Kb before it was deleted. If the system used 4Kb clusters, then the image was assigned 20 clusters. Then, one or more of those clusters that were used by the original image is reused. So, the first 10Kb of the original image is good, but the remainder of the file is not, because after it was deleted, it was over-written.

Evidence Organization

The organization of digital evidence is critical to any investigation. An investigator must be able to take a piece of evidence and determine how it fits in the larger framework of the case. This is true whether the evidence is digital or non-digital. The correlation of digital and non-digital evidence is critical to many cases. It is rare that one single source will provide enough evidence to solve a case.

In most cases involving digital evidence, it is easy for the investigator to become swamped by so much data that it is hard to decipher the key pieces of information. It becomes the proverbial 'search for the needle in a haystack'. With proper case management and information chaining, the investigator can narrow his search to the most likely sources of key evidence. Unfortunately, today, because of the lack of integrated tools, and limited availability investigative software in general, the investigator must spend a disproportionate amount of time performing manual case management and reporting tasks.

Evidence Organization Tools

Evidence organization tools allow investigators to correlate evidence in several different ways:

- Among separate investigators/investigations
- Among digital and non-digital evidence
- Among separate locations
- Among separate incidents and or suspects

One of the main reasons that evidence organization is a crucial step in the computer forensic process is the need for case management. Case management involves the day-to-day organization of digital and non-digital evidence. When a new piece of evidence is located there are many tasks that need to be completed.

- The evidence must be checked for accuracy.
- The evidence must be compared to other evidence.
- The evidence must be properly documented and stored.
- The evidence must be analyzed to determine if it provides leads to other pieces of potential evidence.
- The admissibility of evidence must be examined.
- A strict chain of custody must be established.
- A way of proving the evidence's authenticity must be utilized, i.e. digital signatures.

Many of these tasks must currently be performed manually and are time consuming. By using forensic evidence organization tools, an investigator can practice better case management. These tools can assist an investigator with organizing the forensic evidence

that they find, and help them follow the necessary steps that need to be taken to get the evidence ready for trial.

Link Analysis Tool

The key operation in an investigation is not just the collection of evidence, but understanding how each piece of evidence relates to another. Linking analysis is a very important tool in the investigative process.

The ability to create linked charts provides a powerful case visualization tool. These computer-based link charts enable an investigator to create sub-links, allowing the expansion or collapse of the main chart. Multiple views permit analysts to alternately focus on the main issues, or to examine the background information in detail.

Some of the capabilities of these tools include link analysis or association charts, commodity flow charts, activity charts, network or high volume link analysis charts, timeline/sequence of events charts, case flow/transaction charts, and combined charts showing events and flows. The charts that are produced make it easier for the investigator to document the evidence as they progress with the investigation. This allows them to get a better understanding of the case, and could result in the investigator being able to solve the case much faster. By using this software, an investigator can draw associations between seemingly disparate pieces of evidence, and make it presentable in a court of law.

Refer to Appendix D for further discussion of link analysis tools. A more detailed description of what options some specific link analysis tools can offer (e.g., commodity flow charts, timeline/sequence of events charts, etc.) is provided in the link analysis section of Appendix D.

Time Lining

> "In the investigation of a criminal case involving a computer or computers, the time-line of "computer events" may provide critical information relating to the prosecution of involved persons. Timelines of computer usage can provide valuable information about the computer user and the sequence of events tied to the computer or multiple computers. This information can help to pinpoint the location of certain individuals, can assist with the determination of alibis, can undercover conversations and correspondences, and ultimately may help to determine the guilt or innocence of those facing criminal charges. The following computer events or evidence may provide direct clues to not only the means, but also the motive, of a criminal act.
>
> - Content or update time of electronic documents & files
> - Time and content of e-mail communications and messages
> - Information about system logon and logoff events

- Indication of access to specific Internet documents or sites
- Content of communication with known individuals in chat rooms or through other collaborative means
- Evidence of document destruction or hiding
- Knowledge of the forwarding of messages to external devices such as pagers, voice mail accounts or fax machines"[138]

Time Lining Utilities

While many of the utilities that provide file listing and searching provide the ability to *sort* files based on specific time-related criteria (date created, modified, last accessed), they do not provide additional functionality for further analysis of this list.

These utilities view and sort a file list. They provide a timeline analysis of file dates and times regarding files from one or multiple computer hard disk drives and floppy disks. With it, the investigator can:[139]

- Create a timeline of activity based on file access dates;
- Create a timeline of activity based on file creation dates;
- Create a timeline of activity based on file modification dates;
- Create a timeline of activity associated with deleted files.

There are very few tools in this specific area of cyber forensics. Encase has some features, including a time lining feature released in version 3.0, which help the investigator organize evidence, but most stand alone tools do not offer any help in this area.

Evidence Organization Tools Assessment

While there are several methodologies currently being used to organize and manage digital evidence, the available tools have limitations that can curtail the effectiveness and efficiency of an investigation.

As of yet, there are no tools available that automatically correlate non-digital evidence with digital evidence (including phone records, credit card receipts, eye witness testimony, Internet Service Provider (ISP) records, or other forensic evidence).

There is no tool that can effectively correlate computer information from the same computer or case, or make associations among cases or evidence files.

Tools need to be developed to help deal with the sheer volume of digital evidence created by even small networks of computers.

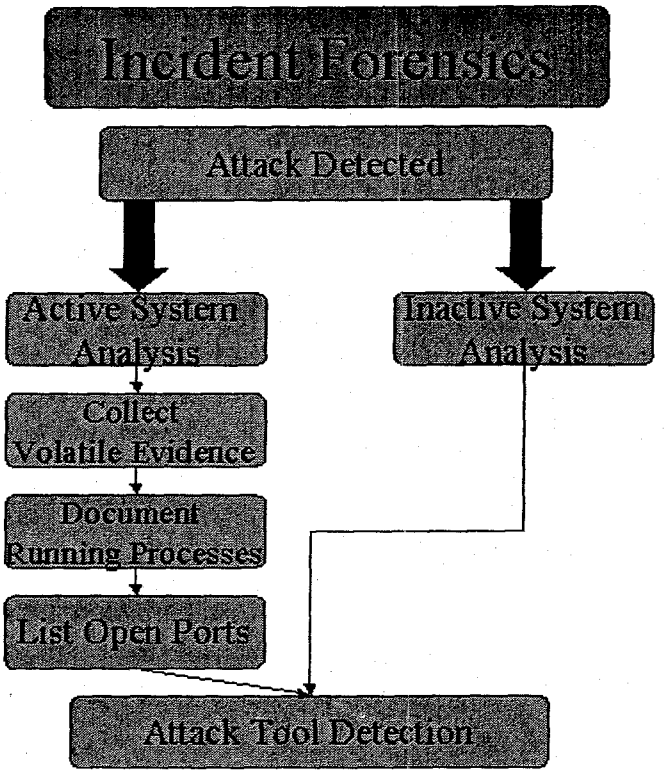
---

[138] Hosmer, C. "Time-Lining Computer Evidence." (Available at http://www.wetstonetech.com/timpaper.htm).
[139] FileList – Time line analysis tool. (Available at http://www.forensics-intl.com/filelist.html).

The programs that are being used have limited capabilities to correlate the evidence from computer break-ins.

In order to have the most effective and efficient investigations possible, new tools must be developed to meet these needs.

Incident Forensics



Incident forensics involves the investigation of a compromise or attack that has occurred on a system. After an attack has occurred, there are two approaches to the analysis of that system: active[140] system analysis, or inactive[141] system analysis.

The analysis of an active system requires additional processing steps before the standard computer forensic methodology is applied. The analysis of an inactive system uses the same basic methodology associated with the computer forensic process, but with a different set of objectives, these being the identification and recovery of modified system files and processes.

First, an investigator attempts to recover 'volatile' evidence. Volatile evidence is data that is available only within active areas of system memory. Once the system is taken off line, or powered down, this information is lost. This data consists of RAM, active processes, active network connections, and the computer screen.[142]

While there are a few utilities available that may successfully retrieve this information, forensics tools do not currently implement this type of functionality.

---

[140] An active system is performed *before* the system is shut down.
[141] Inactive analysis is the traditional computer forensic approach.
[142] Winterton, E. Incident Response Fundamentals Class. Arca Systems, Inc.

After volatile evidence is collected, investigators can use system commands to document running processes, and open ports. Again, even though these tasks can be accomplished manually, forensic tools with these capabilities have not been developed.

Incident Analysis Tools

An attacked system is a crime scene. During the course of the recovery process, the system administrator should gather evidence needed to prosecute intruders if they are found. For instance, if a hacker installs malicious programs on a system, and these same or related files are later found on his or her computer, this digital fingerprint can be an important piece of evidence.

The approach to examining a compromised machine breaks the long-standing rule that discourages operations performed on the suspect system. There are many good reasons to break this rule, including:

- The computer must be operated to determine if a crime has been committed;
- Data must be captured or it will be lost when the system is powered down;
- The system cannot be powered down.

Suites of these incident response tools exist that are intended for post attack analysis of Unix or Windows based machines (NT/2000). Most of the utilities are geared toward data collection rather than analysis. The simplicity of these utilities allows anyone to operate the tool, and then send the output to a skilled forensic investigator for further analysis.

Refer to Appendix D for further discussion of attack analysis tools. This discussion includes a description of four specific attack analysis tools pointing out the "user friendly" nature of some.

Statically Linked Binaries

Often the attacker will replace system binaries with a 'rootkit', so any application that makes use of native system commands cannot be trusted. Common programs such as ps, ls, ifconfig, lsof, etc. may be trojanned or backdoored. When a compromised system is found, it is desirable to have access to 'clean' system binaries.[143]

For the forensic analysis of a potentially compromised system, the investigator may need a set of statically linked executables that can be carried onto subverted systems (where the shared libraries cannot be trusted).

---

[143] Dittrich, D. "Anti Rootkit Tools." (Available at http://staff.washington.edu/dittrich/talks/blackhat/blackhat/antirootkit.html).

The incident-response.org website provides set of these binaries for doing analysis of compromised systems.[144] Currently, sets can be found for the following operating systems:

- Linux 2.2 Sparc
- Linux 2.2 x86
- Solaris 7
- Windows 95/98/ME
- Windows NT/2000

Incident Response Tools

After an incident has occurred, difficult questions arise, such as:

- Which systems are affected?
- Which files are altered?

The tools described in this section may be installed within the environment in which a system attack or compromise occurs. The tools may be in the form of a separate, standalone utility, or as part of an integrated package of intrusion detection and response tools. They all provide a different source of security checking, and address a particular aspect of a potential compromise technique.

But, while the tools may exist within the environment, the degree to which they may assist in or provide relevant information to an investigation may be limited. While they all detect and react to different triggers, the ability to capture relevant data about the incident may be limited, or completely absent, depending on the type of software used, the way it is configured, and/or the limited storage capacity of the user.

Port Scan Detection

While operating systems do not come packaged with the ability to monitor incoming Internet scans, most systems connected to the Internet implement some form of security. Ideally, the first line of defense should be a well-implemented firewall, followed by packet filtering on all Internet-accessible machines, use of TCP-Wrappers, and logging (access, system, event).

More importantly, automated software to examine the logs is often used to ease the task of log analysis (it is not feasible for an administrator to manually read log files because of the sheer volume of transactions). All of these measures have the potential for providing the investigator with information that could point to the source of scanning activity.

While a discussion of every detection tool is beyond the scope of this paper, there are a few noteworthy tools that are freely available that many system administrators have

[144] http://www.incident-response.org/irtoolkits.htm

chosen to employ as a means of detecting potential threats, thus providing them with the means to be more proactive in their implementation of security measures.

- ipfwadm/ipchains - Linux firewall that supports logging of data; using automated filters, it is possible to detect trends, hostile access attempts, etc. Because most firewalls log via the syslog facility, firewall packet logging can easily centralize on a single, dedicated host (requires a large amount of hard drive space).
- PortSentry[145] - a program that detects and logs (and optionally blocks) port scans, including stealthy scans. It will detect scans for exploitable services (old versions of imap, ftp) and scans for Trojan horses (Back Orifice, Netbus etc.).
- scanlogd[146] - a TCP port scan detection tool that monitors network packets and, if a threshold is exceeded, logs the packets.
- TCPWrappers - serves as an additional Unix authentication/logging mechanism to control access to a system.

War Dialing Detection

Many telephone companies have equipment to detect wardialing, and can lock out the attacker once this activity has been indicated. However, this equipment only detects sequentially dialed number attacks. To prevent this means of detection, many wardialer programs allow the randomization of the order in which they dial telephone numbers.[147]

Telephone call accounting systems found on most PBX systems capture, record, analyze and organize information about telephone calls coming into or originating from an entity. Details such as originating call number, number called, time of call, and duration of call are collected. This information can be used to detect incoming wardialing activity.[148]

Packet Sniffer Detection[149]

There is no reliable way to detect a packet sniffer, due to the fact that it operates in a passive state. On a Unix system, one approach is run: ifconfig -a. This will list the available network interfaces of that machine, and show all the information about them. The word PROMISC means that the interface is in promiscuous mode.

There is also a tool called AntiSniff,[150,151] which probes network devices and sees if their response indicates a network card in promiscuous mode, as opposed to normal modes of operation.

---

[145] www.psionic.com
[146] http://www.openwall.com/scanlogd/
[147] http://www.networkice.com/Advice/Countermeasures/Scanners/War_Dialers/default.htm
[148] http://www.att.com/isc/docs/war_dial_detection.pdf
[149] See http://www.robertgraham.com/pubs/sniffing-faq.html#detect for Robert Graham's discussion several techniques that may be used to detect NIC cards in promiscuous mode.
[150] http://www.securitysoftwaretech.com/antisniff/
[151] A product review can be found at http://www.nmrc.org/lab/antisniff.html.

Unfortunately, merely detecting the presence of a packet sniffer only supports the fact that a compromise has occurred. Further analysis by the detection program will not provide any further information as to the source or the identity of the perpetrator.

There is one point that may bear further investigation. The logs that the sniffer programs produce, again, while not yielding any identifiable information related to the perpetrator, will be retrieved at some later date. It may be possible to implement a trap-and-trace on the sniffer log, and trace the attacker to his location.

Password Cracking Tool Detection

It is unlikely that the password cracking tools will be found on a compromised machine. What an attacker will do after the break-in is download the captured password file, and run the cracking program on their own system, because the cracking programs are CPU intensive, and a spike in CPU activity on a compromised machine will be easily spotted.

File Integrity Checkers

File Integrity Checkers create a message digest or checksum for each selected file or directory at the time of installation, creating a 'baseline' database. Thereafter, at predetermined intervals, the program re-calculates the checksum values, and compares the baseline to the newly created value. It then will flag any differences or changes to that file, this indicated by values that do not match.

The use of checksums is important, as attackers often alter system files as a result of a compromise.

- They may alter or replace system binaries (e.g. Trojan Horse).
- They leave backdoors in the system, allowing reentry into the system at a later time (e.g. rootkit).
- They cover their tracks, so that system administrators will not be aware of the attack (e.g. log cleaners).

Although the primary use of file integrity checkers is to determine whether attackers have altered protected files or programs, they may be very valuable to those conducting a forensic examination of systems that have been attacked. A report produced by these applications can be used to aid in the construction of a timeline of the events. The application may also be used for documenting the evidence of an intrusion. By comparing the original checksums with the checksums of the altered files, it can be proved in a court of law that an incident has occurred, and that these checksums provide the proof of that incident. [152] The tools can also be used to verify the integrity of the supporting evidence. The checksums created by the tools are compared at to new values taken after the

[152] "Tripwire: Beyond Intrusion Detection." (Available at http://www.peapod.co.uk/downloads/tripwireuses.pdf).

evidence has been analyzed to provide the requisite proof that the evidence has not been altered.

The freeware product, Tripwire (www.tripwiresecurity.com) is perhaps the best-known example of a File Integrity Checker.

DDoS Detection

It may be possible to discover server software before the DDoS attack, either by using intrusion detection tools to detect commands between the client and the server, or by administering tools such as Remote Intrusion Detection (RID), which can send commands to servers, and detect them when they respond.[153]

Often, when the owner of a system discovers that his system has been compromised, he will reformat the hard drive and reinstall the operating system, thus destroying any traces that the attacker has left behind. These traces can lead to other compromised systems. It is possible to find a list of compromised hosts (usually because an intruder is using rcp[154] over a rootkit and the rcps are logged in SYSLOG) on another site. This list is invaluable to the investigator, but can be overlooked or destroyed by the uninitiated.

Key Logger Detection

Because key loggers are very small programs, and often run in DOS mode, the user would probably never notice them running.

One method to determine if a key logger is running on the computer is to bring up the Task Manger. This will display a list of all the programs currently running. Programs that are not recognized, especially those named "keylog.exe", are suspect.

Some can be renamed, making them harder to spot. The attacker may have given the program a discrete name such as "explore.exe", or "winprog.exe".

One indicator of the presence of a key logger is a growing log file. To check for log files, all programs need to be shut down. A text editor (such as Notepad) is then launched. After waiting a few minutes, a few lines of text are typed. A search for the most recently modified file can then be performed. The file may be a key log. Another approach is to type an unusual phrase and do an advanced search for that phrase. As long as the key log has not been encrypted. The file will be easily located.

Rootkit Detection

Identifying the system processes that have been replaced and/or modified can be very difficult. If the system administrator has generated checksums for all system utilities

---

[153] Farrow, R. "DDoS is Neither Dead nor Forgotten." (Available at http://www.networkmagazine.com/article/NMG20010125S0003/2).

[154] The remote copy command (rcp) is often used during the installation of DDoS software.

installed on the system, at the time of installation, the process of identifying a rootkit will be much easier. Programs like Tripwire are used for such purposes. It is also wise to check the file(s) size and timestamp of those utilities that a rootkit would normally alter.

If the system is Red Hat Linux, the administrator may have used the package distribution mechanism (Red Hat Package Manager, or RPM for short), which has a checksumming capability. The command 'rpm -V -a' will verify all the packages against the RPM database on the local hard drive. It is best to verify the files on the system by comparing the packages with the originals on a CD-ROM or a Red Hat distribution site.[155]

A tool that will try to find rootkits on Unix and Linux systems, chkrootkit, is available for free download.[156] It looks for known "signatures" of trojaned system binaries, by comparing system binaries to the binary signatures of several known rootkits.

Trojan Detection

Once a Trojan takes hold of the system, it can be very hard to detect, and even harder to completely purge from the system. Many trojans have the ability to hide in the system, and will restart every time the operating system is loaded. If the Trojan restarts every time Windows is loaded, it has placed something in the registry, in win.ini, or in another system file, allowing the Trojan to restart. Additionally, the Trojan creates a file in the WINDOWS\SYSTEM directory. The file is always trying to look like something that the victim computer will assume is a normal WINDOWS executable. Many trojans are also designed to hide from the Ctrl+Alt+Del menu, which gives a list of currently running processes on the system.

One way to find out if a Trojan is present on the system is to try the 'netstat' command. By going to the DOS prompt when the system is offline and type in netstat –a, it is possible to see the list of current connections and open ports on the computer. Compare the open ports with a Trojan port list to see if a Trojan port is open. Any unusual response from netstat, especially those indicating high UDP or TCP ports, must be considered suspect. This is because Back Orifice (and many other trojans) do not always use the default port, and can be modified to use any port on the system.[157]

Some good Trojan detection programs are Jammer[158] and NoBackDoors[159]. Jammer attempts to block the introduction of a remote Trojan. It monitors Internet traffic on the computer or network, and can tell if attempts are made to analyze or exploit system weaknesses. NoBackDoors operates in a different way. It lets the user know when it finds a Trojan present on the system. It then allows it to be moved, for further analysis, or kills

---

[155] http://www.dsinet.org/textfiles/faqs/Rootkits_FAQ.htm
[156] http://www.chkrootkit.org/
[157] "Preventing, Detecting and Removing Virus and Trojans." (Available at http://www.safenetworks.com/dica02US.html).
[158] http://www.agnitum.com/products/jammer/
[159] http://home.swipnet.se/technotel/

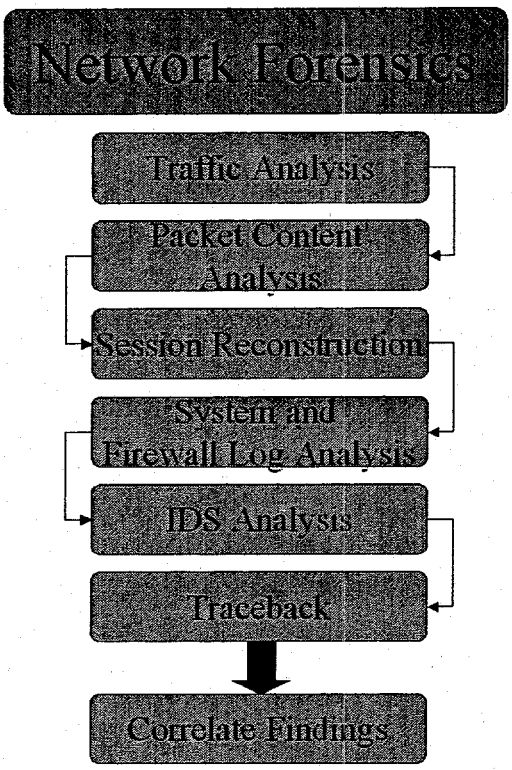
the running Trojan. It will also indicate which program originally contained or started it.[160]

Incident Response Tools Assessment

One problem that investigators have been struggling with is the collection of "volatile" data, this data lying within those areas of the computer that, when they lose power, lose all data stored within that area. This problem is compounded when faced with the investigation of a system compromise. The investigator is not looking for the traditional paper trail of evidence, but specifically for that data that would inevitably be lost if the system was powered down, and returned to a lab for further investigation. And, these target systems are those that may be running valuable network services, and would be cost-prohibitive for the target of the attack to take these systems offline for analysis.

[160] Apel, W. "Protect Your Computer From Unauthorized Access." (Available at http://www.pcworld.com.eg/protect_may2000.htm).

Network Forensics



When conducting a network forensic process, there are steps that should be taken in an attempt to identify the source of the attack or compromise. These steps include; traffic analysis, packet content analysis, session reconstruction, system and firewall log analysis, IDS analysis, traceback, and a correlation of all findings.

Analyzing network traffic is a step in the network forensic process that involves recording network traffic at a high level, and then examining it to determine what type of information is being exchanged, and who is sending and receiving the data. It involves the use of packet sniffers technologies to collect the network traffic. Subsequent analysis may then be performed on the data in a variety of ways.

The investigator may analyze the contents of the collected packets. Special software needs to have been implemented on the compromised system for this level of analysis to take place. It requires a significant investment in system resources, as collecting the contents of the packets requires a large amount of memory. The analysis of the collected data can prove to be very time consuming.

Session reconstruction involves parsing log files and other stored information about network activity, and then correlating the individual events together to reconstruct the original communication session. This allows the investigator to get a picture of what actions took place on the system during the actual compromise.

System and firewall logs usually contain too much information to be properly analyzed, and must be reduced before proper analysis can take place. An investigator could find information such as source host, destination host, time of day, duration of connection, protocol utilized, and number of bytes transferred.

Intrusion detection systems could provide an investigator with valuable information, if forensic capabilities are incorporated into the application. These systems inspect all inbound and outbound network activity, and identify suspicious patterns of activity that may indicate a network or system attack. They record the event, notify the appropriate security administrators of the suspicious event, and take the appropriate action to block further activity from the source of the unusual activity. All alert logs are analyzed to try and ascertain what went on during the event.

Traceback programs attempt to identify the location of the suspicious activity in real time. Using available information, these programs graphically trace the traffic back through each Internet node the traffic has used passed through in an attempt to identify the originating IP address.

Finally, the investigator correlates all of these findings and tries to determine the source.

Introduction

When networks are involved, the usual problems encountered with digital evidence collection are exponentially compounded. The collection of digital evidence in a network environment is termed Network Forensics. There are several additional complications to a cyber investigation when a network is involved. These complications can hamper the most skilled investigator, and investigations often require special equipment and training in order to be completed correctly.

Network forensics involves, "The reconstruction of events on a client network deduced from the clues at hand."[161] In order for an investigator to figure out what went on within the network, they have to piece together the evidence that they have collected with forensic tools.

Network-based digital evidence may reside on a variety of network devices, and requires the analysis of network hardware for relevant evidence, such as hubs, servers and routers.

---

[161] "Network Forensics." CTX Corporation. (Available at http://www.ctx.com/ctx_htmlcode/ctx_page.cfm?SectionID=Forensics).

Traffic Analysis

Traffic analysis is the process of recording network traffic at a high level and then examining it to determine what type of information is being exchanged and who is sending and receiving the data. For example an investigator may use a sniffer or network analyzer to determine if two individuals are communicating and what type of traffic they are generating i.e. encrypted email, FTP, IRC, etc. This information is useful to determine relationships between people and entities. The investigator can then focus his investigative efforts where they will be most effective. This type of analysis can also provide links to new suspects or sources of digital evidence.

Packet Content Analysis

Packet content analysis involves the inspection and extraction of evidence from individual network packets. To facilitate data transfer, networks break up the information sent over them into uniform sized parcels. These packets contain a wealth of information for a trained investigator. Not only do they contain information such as emails and files, but also routing information such as sender IP address, receiver IP address, etc. These pieces of information can be used to locate and identify suspects, victims, and further sources of digital evidence. With the tools available today this is still a fairly labor intensive task. There are now tools available that allow keyword searches on the data or that will extract all IP addresses, but the investigator must know what he is searching for and must be technically sophisticated enough to understand networks and how data is transferred over them.

Session Reconstruction

Session reconstruction allows the investigator to 'see' what the suspect saw. This means that, instead of looking at ASCII text, HTML code, or log files, the program will reconstruct the actual HTML document or email as close to the original form as possible. It also allows the investigator to follow what the person did on the computer network, step by step. This is done by parsing log files and other stored information about the network session and then bringing the individual pieces back together to reconstruct the original session. This makes the investigator's job far easier and allows him to view evidence in an easier to understand format. Unfortunately, this technology is still new and does not work for many applications.

Network Forensics also involves analysis of security applications that may have been running on a target system, such as activity and system logging, firewalls, IDS, tracebacks, honeypots, etc.

Network Forensics Tools

Network forensics is still a very young field, and there are a limited number of dedicated tools available to aid in the recovery of digital evidence in a networked environment.

Many network security tools are claiming to employ forensic capabilities, but these capabilities have many shortcomings, and lack standardization. The tools address security in different ways, collecting disparate information, and storing that data in proprietary formats.

System and Firewall Log Analysis

System logs (access, activity, firewall logs, etc.) are too large to be efficiently analyzed. The data in them must be reduced before they can be properly analyzed. Additionally, many types of computers (e.g., small, mobile, or embedded systems) do not themselves have the resources for maintaining audit-logging capabilities.[162]

The majority of information recorded in a firewall log is statistical connection detail. At least one such entry is written to the log for each connection that is made through the firewall. These entries contain details such as source host, destination host, time of day, duration of connection, protocol utilized, and number of bytes transferred. The logs created by firewalls can easily become large, making it difficult to determine the activity of server.

While log file analysis tools do exist, and are sometimes implemented on a system, many firewall analysis tools only allow administrators to generate reports showing information related to various protocol usages and generate traffic summaries based on Internet usage and network load summaries (such as the busiest time of day). From the forensic investigator's point of view, this information is useless. The data that would be most useful, such as source and destination IP addresses of packets that violate the firewall rule set, is typically discarded.

Intrusion Detection System Analysis

An intrusion detection system (IDS) is designed to inspect all inbound network activity, identify suspicious patterns of activity that may indicate a network or system attack, record the event, notify the appropriate security administrators of the suspicious event, and possibly take some action to block additional activity from the source.

Most organizations employ a variety if IDS to protect the variety of systems running within their network, these systems based on any number of platforms and operating systems. Different types of IDS are used to detect different forms of attacks. Some IDS are network sniffers, which look for suspicious packets. Some identify suspicious activity on individual hosts. Others observe interactions at the application level. Some are better at detecting certain kinds of attacks than others, while others detect the same kinds of attacks using different mechanisms.[163]

---

[162] Wee, C. "Audit logs: to keep or not to keep?" (Available at http://www.raid-symposium.org/raid99/PAPERS/Wee.pdf).

[163] Loshin, P. "Eliminating IDS Babble." (Available at http://www.infosecuritymag.com/articles/june01/columns_standards_watch.shtml).

The one thing that they have in common is their lack of commonality, this meaning lack of standardization, even among the tools that operate in similar fashion. Most individual vendors use proprietary formats for IDS information, presenting collected information in formats that may be difficult for anything but an experienced systems administrator to decipher.

There are several ways to categorize IDS:

- Misuse Detection vs. Anomaly Detection
- Network-Based vs. Host-Based Systems
- Passive System vs. Reactive System

Misuse Detection vs. Anomaly Detection

There are two basic methods used by intrusion detection systems on the market. These systems consist of anomaly detection, and misuse detection.

> "Misuse detection systems resemble a virus scanner attached to a network. These systems are usually programmed with signature sets representing the types of connections and traffic that indicate a specific attack. Other forms of these systems rely on host platform information such as C2 audit logs (which record information such as file accesses), to detect patterns of suspicious activity."[164]

Anomaly detection systems learn what constitutes normal network traffic, and develop models of this traffic, which are updated continually as the conditions of the network change. These models are compared to new traffic that comes in, and if the new traffic does not match the normal model, it is flagged as suspicious.

Additionally, many modern systems use a combination of both misuse and anomalous detection engines, creating yet another format for output.

Network-Based vs. Host-Based Systems

Network-based systems examine the individual packets flowing through a network. Unlike firewalls, which typically only look at IP addresses, ports and ICMP types, network based intrusion detection systems ("NIDS") are able to understand all the different flags and options that can exist within a network packet.

While network-based IDS look at all the traffic flowing by on the network, host-based intrusion detection systems are concerned with what is happening on each individual computer or "host." They are able to detect such things as repeated failed access attempts, or changes to critical system files.

---

[164] Ranum, M. "Is Network Intrusion Detection Software Being Used Correctly?" (Available at http://www.securitymanagement.com/library/000556.html).

Passive System vs. Reactive System

In a passive system, the IDS detects a potential security breach, logs the event, and signals an alert. In a reactive system, the IDS is designed to respond to the suspicious activity by logging off a user, or by reprogramming the firewall to block network traffic from the suspected malicious source.

Intrusion Detection Weaknesses

Intrusion detection systems have many inherent problems that have been exploited by users and professional in the field of network security. Some of the most common problems, and those that a network forensic investigator must be made aware, are:[165]

- Data overload - how much to record, how long to keep it, and how to present it to the end user
- False alarms - common and hard to distinguish from actual intrusions
- Not knowing what to do with alerts once they are received

Agencies that are concerned with and researching steganography will want to record all of the image files that pass through their intrusion detection system. Image files can range in size from a few kilobytes to several megabytes in size and a typical website may contain dozens of graphics. Storage of these files, along with the associated audit trails for each image can quickly overwhelm a system.

Intrusion detection tools have so much data to deal with that it is hard for them to detect attacks accurately. Most of the data that is analyzed by intrusion detection systems is not harmful. For example, the investigator must be aware of the possibility that the IDS may be set to a highly sensitive state, and that much of the data collected will be due to innocuous scanning activity.

In order to reach their full potential as forensic tools, the intrusion detection systems' role must evolve to include better logging and the implementation of forensic capabilities to better use the information as evidence. Many of the newer IDS products now claim to have better facilities for collecting forensic evidence, which could be used to aid in the identification of intruders, but there are no established guidelines to determine what information should be saved, no standard format for the preservation of such data, and little information available as to proper storage of such data to maintain its evidentiary value.

Many logs are easily modified or deleted. There is also the problem of the perpetrator assuming the user ID, and therefore the identity, of an innocent party. Trust in these

---

[165] Power, R. "CSI Roundtable: Experts discuss present and future intrusion detection systems." (Available at http://www.gocsi.com/roundtable.htm).

forms of digital evidence can be increased if logs are encrypted and/or digitally signed, and if there are strict and secure access control procedures.
The second area where trust is a problem is in the internal clocks and time stamping of logs. Changing the element of time on a computer or network is a simple task for an experienced user. This can create all kinds of problems for evidence correlation and time lining of evidence. In order for digital and network forensics to take the next step forward, a protocol must be implemented within networked environments that employ trusted time stamping.[166]

Traceback

Anyone who is connected to the Internet via DSL, cable modem, or 56k modem is susceptible to a hacking attack. After an attack occurs, there are tools available that attempt to traceback the route that the attack came from. The traceback utilities are commonly used with network utilities and intrusion detection tools to combat the attacks.

After the attack has taken place, the traceback utility collects valuable information about the intruder, such as the intruder's IP address, computer name (NetBIOS name), and hardware address (MAC address).[167] These tools can keep logs of everything that a hacker sends to the target computer. Law enforcement, ISP's, and network professionals investigate attacks that occur by using these tools.

Not only can these utilities traceback information about the attacker, they can also trace the route that the attack took to get from the attacker's computer to the victim's computer. This trace routing involves identifying the route between the two computers, which includes all intermediate nodes, and their registrant information.[168] This option will inform the user of router information, which includes the cities of their origin. These tools work by sending small ping packets from the users PC to each hop on the way to a specified destination.

Whois is a tool included with some traceback utilities that can find information about a computer located in any part of the world. This tool will also deliver all of the related records within a few seconds.[169] The user types in the IP address of the computer, and this tool will give all of the desired information to them.

Most traceback utilities are included in a suite with other network utilities. All of these network utilities work together to gather valuable information. The following is a list of some of the more popular network utilities that traceback tools work with:

- Ping - short for *Packet Internet Groper,* a utility to determine whether a specific IP address is accessible

---

[166] See Appendix E for detailed discussion of trusted time.
[167] Network Ice Corporation. http://www.networkice.com/products/blackice_defender.html
[168] Neoworx. http://www.neoworx.com/products/neotrace/default.asp
[169] TamoSoft, Inc. http://www.tamos.com/products/smartwhois/

- Trace Route - a utility that traces a packet from a computer to an Internet host, showing how many hops the packet requires to reach the host and how long each hop takes
- Forward and Reverse Lookups - translating IP addresses to the domain name, and vice versa
- Network Scanning- scanning a network to recover packets, and to search for intruders that are trying to access the system
- Whois - an Internet utility that returns information about a domain name or IP address
- Finger - a UNIX program that takes an e-mail address as input and returns information about the user who owns that e-mail address
- WhoAmI - the whoami command displays your login name. Unlike using the command who and specifying am i, the whoami command also works when you have root authority since it does not examine the /etc/utmp file
- ISP Lookup - utility that allows user to find out the name of the Internet service provider being used by a particular computer
- Domain Finder - this utility is capable of finding domains that are being used on a network.
- MTU Tester - utility that tests an MTU; MTU is short for *Maximum Transmission Unit,* the largest physical packet size, measured in bytes that a network can transmit. Any messages larger than the MTU are divided into smaller packets before being sent
- Subnet Calculator - calculates the subnet portion of a network. A subnet is a portion of a network that shares a common address component. On TCP/IP networks, subnets are defined as all devices whose IP addresses have the same prefix
- Telnet - a terminal emulation program for TCP/IP networks such as the Internet. The Telnet program runs on your computer and connects your PC to a server on the network
- NSLookUp - used to find the physical IP address or the host name of a machine
- Time Synchronizer- synchronizes the computer clock time with Internet time
- DBScanner - is responsible for monitoring the EDBServers and distributing the OA/LAL section file databases

Traceback utilities are equipped with many capabilities that make them even more powerful. The most popular capabilities that appear on almost every tool are IP address and host name lookup, location, host name, computer name, and computer address.

There also are some capabilities that only exist on certain traceback tools. A few of the unique capabilities that are available on some of the utilities include:[170]

- Detailed printable traceroute reports
- Provide network names and contact information
- Mail server tracing

[170] See Appendix B for additional information about traceback tools.

- Hostname resolution and DNS caching
- Can load and process lists of IP addresses
- Ability to view the traceroute visually on a map of the world
- Perform round-trip and remote trace routes
- Track down spammers and contact network administrators
- Fight SPAM e-mail by using contact information provided by popup network and domain WHOIS information
- Detect routing loops

Traceback utilities are very helpful to law enforcement, ISP's, or network professionals that are involved in a hacking investigation. These utilities make it easier to pinpoint where the hack came from and can save valuable time in a real investigation.

Network Forensics Tools Assessment

Network forensic tools and methodologies can perform the following tasks:

- The distributed network collection of all network traffic.
- The availability of robust database tools for evidence analysis.
- The automated analysis of specific types of network traffic.
- The automated correlation of data across an entire enterprise.

They are a rich source of metadata, and they help investigators to see potential evidence links through advanced visualization.

However there are gaps in current Network Forensic tools and methodologies. Limited data reduction capabilities make it difficult, if not impossible, for investigators to sift through enormous amounts of data. The absence of trusted time stamping of events makes time lining and evidence correlation difficult and hard to prove in court. Weak evidence preservation may allow valuable evidence to be destroyed or tainted.

The first major hurdle of network forensics involves the collection and imaging of the digital evidence. If a computer is connected to a network, it may not be shutdown or removed without severely affecting the entire network. This is especially true for servers, where much of the digital evidence is likely to be stored. Under most circumstances the investigator must work with the system administrator or other company officials to plan the evidence collection and imaging when it will cause the least disruption to the system.

The second major problem is the sheer volume of data that can be stored on a network. The investigator will generally need to sift through a lot of insignificant data to find the valuable evidence. There are tools available to help in this task, but most are still relatively unsophisticated and require the investigator to do most of the evidence extraction and evidence searching manually. Even a small network of a dozen computers can contain several hundred gigabytes of data that need to be imaged, extracted, and searched.

A third difficulty is the transitory nature of network data. The amount of data traffic causes lots of disk writes, which can result in the overwriting of potential evidence. There is also the problem of log files being automatically deleted after a certain time period or when they reach a certain size. This is especially a problem where disk space is tight. Any delay in data collection could mean the loss of critical evidence.

The final major hurdle in network forensics is the lack of trust. This lack of trust falls into two major categories: trustworthiness of logged data and internal time clocks.

There are many privacy concerns regarding the wholesale collection of all network traffic, e.g. Carnivore. Limited evidence reasoning tools require the investigator to have a high level of technical savvy and experience with networks. Automated session reconstruction tools are crude and often do not work for many applications and types of traffic.

As the use and misuse of networks increases, development of tools to address these gaps and limitations becomes crucial.

Honeypots

A honeypot can be either a program or a computer configured to simulate a legitimate computer or network service(s), but its actual purpose is to capture details of a potential attack. Honeypots are known also as decoys or booby traps. The more realistic the system appears, the longer the attacker will stay and more will be disclosed about their techniques.

By reviewing the order, sequence, time stamps and type of packets used during a honeypot attack, the analyst may identify the tools and methodology being used by the attacker, their skill level, and their intentions (vandalism, data theft, remote launch point search, etc.).[171]

In order to make honeypot systems more forensic-friendly, they must be upgraded to include additional security measures, such as an intrusion detection system that triggers an alarm whenever an attacker breaches security on one of the networked computers, a keystroke logger that watches everything the intruder types, from commands to e-mails to chat sessions, and a separate firewall that cuts the machines off from the Internet whenever an intruder tries to attack another system from the honeypot. The output needs to be collected and stored in a way that ensures its integrity, preferably a digital signature implementation that timestamps the logs.

Honeypot Assessment

[171] Even, Loras. "What is a Honeypot?" (Available at http://www.sans.org/newlook/resources/IDFAQ/honeypot3.htm).

Even though honeypots are purported to be one of the best ways to analyze the activity of a hacker, they also have problems associated with them, and thus their usefulness as a forensic tool is limited. The problems range from complicated implementation, to issues pertaining to security of the network they reside upon. The following is a list of problems associated with honeypots:

- Difficult to emulate services that will trick hackers
- Only capable of collecting a limited amount of information
- Could provide hacker with unexpected access to a system
- May placate hackers
- Providing administration to overlook honeypot
- Limited or no evidentiary value

Difficult to Emulate Services

It may be hard to emulate a type of service on a honeypot well enough to fool an experienced hacker. A hacker may try a variety of e-mail addresses, and check for expected responses, as well as try a number of control commands. Unless the honeypot is able to pass these various tests, the hacker might realize that the honeypot exists, and decide to leave.

Collects a Limited Amount of Data

Another problem with honeypots is the fact that they are only capable of collecting a limited amount of information. An observer of a honeypot can see the initial attack, but cannot see anything else. *Honeypots are only useful for examining the hacker's methods, rather than figuring out their identity, their location, or even what tools they are using.* A honeypot is basically used to view the hacker's style of attacking systems. While this may provide a security administrator insight into what systems may be weak and in need of attention, or what the latest attack may be, it does little to aid an investigator. "Because the honeypot does not allow the hacker access to the machine, additional forensic information cannot be collected."[172]

Could Provide Unexpected Access to System

A honeypot could also allow a hacker to access a system. If an administrator records all log entries on the honeypot itself, and a hacker finds a way to break the honeypot, the hacker may remove all evidence of his attack. In addition, a hacker could use the honeypot as a tool to attack other systems.

Placate Hackers

In some cases, a honeypot could aid the attacker in his ultimate goal. "If the hacker's intent was to install a Trojan horse program, the user has given them what they've wanted

[172] Brenton, Chris. "Honeynets". Dartmouth College Institute for Security Technology Studies (ISTS).

by allowing them to install the Trojan in their honeypot."[173] The honeypot is an open system, and therefore vulnerable to this type of compromise. Honeypot users would not want to find that their honeypot played a role in an attack on another system.

Providing Administration

One of the biggest problems with a honeypot is providing staff members that are willing to give up their time and effort to the project. "It takes time and diligence on the part of the security administrator to ensure that any penetrations into the honeypot are quickly identified, and control of the system is not lost to a perpetrator."[174] Administrators need to spend several hours every day reviewing honeypot activity. As honeypots do not provide security against hackers accessing a system, someone is needed to watch the tool at all times to make sure that the intruders do not gain access to the system.

Limited or No Evidentiary Value

If the information gathered from a honeypot system is used as evidence during a criminal prosecution, it may not be admissible in court. There are questions as to whether or not courts will accept honeypot data as evidence, or if non-technical juries will be able to comprehend its evidentiary value.[175]

Trusted Time Stamping

> "In many investigative situations, absolute and trusted time is a crucial aspect of information operations. In numerous critical systems, time is used to distinguish and identify when encryption keys are changed, the exact sequence of network events that led to an intrusion, when the preservation of digital evidence has taken place, how the recommended course of action was derived, whether a digital certificate expired, or whether a user has access to our information systems. In all of these situations, the exact and correct time is a critical factor.
> Increasingly, the demand for accurate, trusted, non-forgeable, verifiable and non-repudiating time is becoming a crucial aspect of security trust models. As information becomes more time dependent, whether that need is for the coordination of digital and non-digital evidence, the protection of the national infrastructure, or the integrity of e-commerce operations, the problem remains the same."[176]

The following sections relate the trusted time inadequacies to real problems faced by information investigators today.

---

[173] TISC Insight. "Honeypots: Sweet Idea, Sticky Business". Volume 3, Issue 2. January 26, 2001.
[174] TISC Insight.
[175] Even, Loras. et. al.
[176] "WetStone to Apply Trusted Network Time Stamping Expertise to Defensive Information Warfare Applications." April 19, 2000. (Available at http://www.wetstonetech.com/pr005e.htm).

Access Control Decisions

Insiders with access accounts can make subtle temporary adjustments to computer clocks in order to provide unauthorized access. For example, access control systems can configure policies for user access during certain limited access hours. Since the access control logic uses the local system clock to determine the current time, undesired access control decisions can be caused by inaccurate, accidentally changed, or maliciously forged clock settings.

Digital Certificates Expiration

Digital Certificates and Public Key Infrastructures (PKI) are used widely today for authentication, non-repudiation and access control. Each digital certificate has an expiration date that determines the viability of the certificate. These certificates are used to control access to subscription services, to allow users access to information within a short time frame, or to issue temporal access privileges. The certificates themselves are secured with digital signatures to ensure their integrity. Unfortunately, the certificate expiration decision is based on a comparison of the expiration date/ time in the certificate with the current local time. If the current local time can be forged, users previously holding access can forge their way into systems to which they no longer have legal access.

Replay Attacks

Many distributed network management systems, such as Kerberos, use temporary tickets to validate user privileges. The key element of the Kerberos scheme is the addition of a time stamp that proves that the client "recently" obtained the ticket. Kerberos tickets can be captured from normal network traffic quite easily; once the ticket has been captured, the attacker must break the encryption and then attempt to impersonate the client. The attack would normally fail, because the time stamp submitted would fall outside the valid time range of the ticket. However, the time stamp check is compared against the untrusted local time or network time of the network under attack.

Statistical IDS Decision Thresholds

Many Intrusion Detection Systems (IDS) use statistical thresholds to detect system, user, application, or service behavior that falls outside of the normal for that particular system. For example, during normal business hours (e.g. 9 a.m. – 5 p.m.) it might be statistically normal for 10,000 e-mail messages to be sent from inside a network to outside Internet recipients on an hourly basis. However, it would be unusual to see this volume at 2 a.m. Since IDS systems use untrusted local clocks for decision-making and statistical profiling, if the time is changed from 2 a.m. to 2 p.m., the IDS will miss the anomaly, and thus report no suspicious behavior or perform additional logs/audits of the event.

Digital Evidence Preservation

Digital evidence is preserved and protected today using similar methods. With proper handling, all digital evidence data is run through a one-way hash algorithm (e.g. NIST's SHA-1[177] or RSA's MD5[178]). The resulting hash value (a 120 – 160 bit number) is then stored.[179] Later the hash can be recalculated and compared to the stored or recorded value. There are several significant trust problems with this current "best practice" approach.

First, since the hash algorithm is known, should modifications be made to the evidence, a simple recalculation of the hash along with replacement of the originally stored hash value circumvents the protection afforded by the hash. Since there is no binding of the time when the hash was first calculated, it can be modified and recalculated as many times as needed.

As an example, digital cameras are being used to record evidence at a crime scene, as well as during criminal activities. Due to the inability to properly demonstrate how the digital evidence (such as digital photographs) can be amply protected from undetectable alteration, this significant technology advancement has become all but banned in criminal prosecutions.

Event Correlation and Decision Support

Within the domain of intrusion events, audit logs, anomaly reports, system outages, denial of service events, user activities, and system events must be correlated across local, wide area and global networks. Correlation of the data to piece together the accurate sequence of events, tracing of the originators, detecting collaborative attacks, and reconstructing systems that have been compromised all rely on time as the synchronizing element.

Having a trusted source of time is essential to many investigations. If the cyber forensics investigator cannot trust the timing of events on computers and networks, it is difficult to correlate digital evidence to non-digital evidence in a case, and create an accurate timeline of events.

Time Stamping Assessment

Computer clocks are notoriously inaccurate, which makes relying on them for accurate time purposes risky. In addition, many clocks are easily modified and, as a result, information warfare attacks may specifically target computer clocks to attempt to fool the system, its authentication methods, access control mechanisms and digital certificates.

---

[177] Computer Security Division National Institute of Standards and Technology. "Digital Signature Standard (DSS)." (Available at http://csrc.nist.gov/cryptval/dss.htm).

[178] "What are MD2, MD4, and MD5?" (Available at http://www.rsasecurity.com/rsalabs/faq/3-6-6.html).

[179] On disk or paper

Event Correlation and Decision Support

Today, the synchronizing, compensating and correcting of inaccurate, or forged, time is extremely difficult. For all practical purposes these events cannot be accurately correlated even under controlled circumstances. This results in poor decision support, false correlations, and significant staff hours to attempt to conjure up the correct sequence of events manually.

Concluding Remarks

In this task, we have identified the current types of technologies available to the investigative community, and described the techniques employed by the various tools. The types of data available to the investigator have increased significantly from the early days of deleted file recovery.

While forensic technology has matured in many ways, it has not been able to stay abreast of advances in computing technology, and the way the technology is exploited. Computers and networks are used in ways that the fathers of the technologies never imagined. New ways of computing continue to evolve, and become quickly available to end-users. Unfortunately, insecurities in the software and its implementations create new avenues of exploitation for cyber criminals.

The task of investigating non-traditional crimes remains labor-intensive. There are very few dedicated forensic tools that automate the process. The following section describes this and other gaps in the currently available technologies. These are the areas that are in desperate need of research and development.

Task 3: Gaps between Existing Cyber Crime Technologies and Current and Future Law Enforcement Needs

Introduction

Historically, there has been a significant gap between the cyber tools used by criminals and those available to law enforcement to prevent, detect, and investigate cyber crimes. Over the last few years, this gap has been closing, but still poses a serious threat. The existing tools available to law enforcement have had limited capabilities, and have become less effective as the size of hard drives have increased dramatically and the need to analyze information from networks increases.

Based on the research in this report and the survey[180] conducted, this section will identify where gaps currently exist, as well as indicate where future development must occur to eliminate the gaps. The survey asks practitioners to assess current computer forensic tools and to identify current and future technology needs.

These findings can provide a roadmap for future development of cyber forensics tools, as well as directing research funding for solving some of the more difficult problems.

Identified Gaps

The gaps will be addressed as they relate to specific digital forensics processes. These areas are:

- Evidence Collection and Preservation
  - Collection Methods
  - Imaging Methods
- Evidence Extraction
  - Evidence Viewing
  - Hidden Data Detection
  - Hidden Data Recovery
  - Evidence Searching
- Evidence Examination and Analysis
  - Evidence Correlation
  - Evidence Time Lining
  - Evidence Mining
- Network Forensics

[180] We conducted a survey entitled "Questionnaire for the Computer Forensic Practitioner". This survey was sent out to practitioners in order to get feedback on current computer forensic tools. Refer to Appendix E for survey results and comments.

- Evidence Correlation and Case Management

Evidence Collection and Preservation

The cyber forensics specialist must have the know-how and the proper tools to make an exact copy of the digital evidence without altering in any way the original. This step is essential - especially if the digital evidence is to be used in any legal proceedings. The investigator must be able to prove that any evidence retrieved has not been tampered with, or in any way altered, thus, is the exact duplicate of the original. With the exception of using hard drives to image hard drives, investigators are faced with a host of problems given the choices of media available to them.[181]

Although there are many cyber forensics tools available for retrieving and preserving evidence from a suspect computer, there are limitations to their effectiveness. New technologies need to be developed to fill the gaps in the collection procedure. Today's off-the-shelf software backs up evidence slowly, and is prone to errors that require that the duplication process be repeated. It requires significant technical savvy, and in some cases, expensive specialized training that can be obtained only from the vendors of the software. The backup is not always reliable, and often requires the investigator to use separate software programs to ensure authenticity with cryptographic certainty. Backing up large hard drives or multiple computers are difficult at best, and very time consuming, as the data cannot always be confined to a single backup device.

The gaps in the this area are exacerbated by the increase in the volume of data as a result of increasing hard drives and the growing number of cases where network forensics is required.

Gap: Tools That Verify Data During Acquisition Process

Several software tools are available that meet the criteria of "forensically sound" data acquisition tools. Current practice involves the creation of a hash value of the *entire* target media that will be imaged, imaging that media, and then comparing a hash value taken of the copy to that of the original.

While this approach has been used effectively to date, the ever-increasing size, variety, and complexity of storage media will make this simple approach less effective.

The imaging process is currently very time consuming. And the fact that the error checking does not occur until the imaging process is completed can prove to be a waste of time if the hash values prove not to match. This means that the imaging process must be started again, and be re-done as many times at is required to get the hash values to match.

---

[181] Refer to Appendix F for a discussion of image media.

The process would be more efficient if a tool implemented error checking periodically throughout the imaging process. Currently, there is only one tool[182] that will stop and provide the operator with an error message.

Gap: Lack of Tools that Collect Volatile Evidence

The 'working memory' of most operating systems only holds data while the machine is still in operation, or while a power source remains attached to the machine. Once the machine is powered down or disconnected, the data in these areas is lost. Of most vital importance is the current working memory, especially if a search warrant is served on a suspect who actively is engaged at his/her system.

While most collection and preservation methods are conducted 'post-mortem'[183], there are times where it would be beneficial to the investigation to gather volatile, or active, data. Where volatile evidence is involved, it is clear that it need be collected on site, while the system is still in an active mode. But tools are lacking that would accomplish this task in a manner that can be proven to not alter other components of the system that may be of evidentiary value.

Additionally, techniques need to be developed that would link this volatile evidence to other subsequent evidence that is later collected from the target media, because, owing to the size of much of the target media, and its expected growth in the future, it will be impractical to conduct on-site collection of electronic evidence.

Gap: Lack of Tools that Collect Data from Active Systems

While this may be closely related to the discussion in the previous section, it is of a somewhat different aspect of an active system. There may be times when, due to the specific purpose of an evidentiary system, it is not possible to take that system offline, or otherwise render it inactive, in order to conduct the collection of the data. It may be ill advised, if not impossible, to take down the system, and conduct a full forensics analysis, either due to operational (mission-critical systems) or financial (cost) factors.

Evidence Extraction

It is evident from the research and comments from practitioners that there are many gaps in the evidence extraction arena. There is limited multi-platform support (i.e. Windows®, Solaris®, Linux, Mobile, and Network extraction) within a common technology. Other needs and limitations in the evidence extraction arena include large media and work group restrictions, in scope of warrant extraction, inculpatory and exculpatory balance, network extraction, isolation and detection of malicious code, expanded universal file type identification, expanded file viewing capabilities, and methods to scale legal and privacy issues. Furthermore, there are a limited number of qualified personnel who are skilled in using these tools. Clearly, more tools need to be

[182] ByteBack

[183] The suspect computer has been powered down, and brought to a laboratory for analysis.

developed to meet these needs and more individuals need to be trained in using both the tools that currently exist and those that are being or will be developed.

The current investigative approach is to treat the "suspect" computer as a pseudo file cabinet. Search and analysis tools and techniques focus first on the extraction and organization of documentary data (files and images), followed by a search of this data for "clues" to the prescribed crime (keyword search). Additionally, these files are analyzed for their "metadata" information, this being the information contained in their file header (dates/times related to the creation, last access and modification of the file).

While this information is invaluable for any investigation, and the tools that do this type of analysis have been found to be effective, tools that may be used to do further analysis of the suspect system are few, difficult to locate, and typically can only be used by the most technically oriented investigator.

Several of the most notable absences of available tools are discussed in the paragraphs that follow.

Gap: Current Forensic Tools Remain Labor Intensive

The majority of tools used by law enforcement do not have easy-to-use graphic interfaces, and those that do may not always be the best tools for the job. The documentation that accompanies the command line tools is often as difficult to read and interpret as the tool itself, and requires the investigator to have an intimate knowledge of operating system commands and directory structure.

Most commercially available tools are of a proprietary nature, and come only in their precompiled form. Modifications for a particular investigator's needs may be difficult if not impossible. Even those that do allow for a modicum of customization do so in the form of scripting languages that are unique only to that tool, requiring the investigator to learn a new programming language to make full use of the tool. Evidence files produced by these tools are also found to be in a proprietary format, and not easily exported to other programs for further examination.

Gap: The Lack of Tools for Operation/Analysis on Alternate Operating Systems

Most current computer forensic tools only support the examination of Windows® operating systems. Linux and Unix operating systems are now getting more attention, and forensic tools that support these systems are beginning to emerge, but have yet to be proven. However, systems like Macintosh® and Sun® have virtually no tools that support them. This is not to mention the many other operating systems in existence today.

Some tools claim to have the capability of text searching, but because of the file system schemes used by these systems, all indication of file structure is lost. By the same token, the file attributes may not be recoverable. It is these file attributes, such as time of

creation, modification, and deletion that enable a piece of evidence to be linked directly to a suspect.

In response to our questionnaire, practitioners expressed the need for tools that handle more operating systems. Multi-platform support was a gap that they felt needed more attention to assist them in their daily investigations.

Evidence Viewing

Gap: Lack of Tools Capable of Viewing Obscure File Formats

Criminals also tend to use various file formats to confuse investigators. When writing an important document, the criminal might decide to use an obscure format to confuse the investigator, and fool the computer forensic tool into recognizing the document as some other file. Using obscure file formats is a popular technique with child pornography offenders. These criminals tend to use various formats to store their pornographic image in so investigators will not identify them.

At the present time there are more than 30 different digital image file formats.[184] This vast number can cause certain problems to those investigating child pornography. There are currently no computer forensic tools on the market that will recognize all of these formats. Programs need to be able to recognize all of the image formats that are available to the criminals. Either new tools must be developed that have these characteristics, or the current tools must be updated to perform these tasks.

Also, if the tools will not recognize different image formats, there is no way that the investigators could recognize the presence of steganography. Not only will pornographic images get past the investigators, but hidden information as well.

Respondents to the questionnaire identified the need for more updated graphic viewers to support image formats, especially AOL's .art files. AOL's .art files give the investigators trouble when they are trying to view them. These files are not compatible with most programs other than AOL.

The default on AOL software is set to use compressed graphics when viewing image files on the web. By using these compressed image formats, web pages will load much faster when an AOL user is searching the web.[185]

The problem the investigator has is the fact that AOL changes the formats of these image files from .gif, and .jpeg files to .art files. As mentioned before, few programs other than

---

[184] Berg, E. "Digital Enhancement and Transmission of Latent Prints Who Will Set The Standards?" *THE PRINT*, Volume 12 (4), July/August 1996, pp 6-9.

[185] "AOL's 'Art' Format: An Evil Plot To Make The Web Look Ugly." (Available at http://members.aol.com/_ht_a/prncssvi/aol/artht.html).

AOL are compatible with these image file formats, and investigators are having a hard time trying to view these files.[186]

Gap: Inability to Mount and Examine a Variety of Operating Systems Within the Same Environment

While closely related to the issue of operating system-specific tools, survey respondents singled out this inadequacy of current tools.

The individual forensic tool operates on a narrow set of platforms, forcing the user to use multiple tools to investigate disparate operating systems (a single storage media may use multiple file and/or operating systems).

When a single investigation involves the analysis of more than one suspect computer, it would be advantageous for the investigator to have the ability to mount these different systems within the same investigative environment, allowing them the ability to correlate the evidence in a central evidence file.

Gap: Ability to Preview HTML Pages That Are Extracted From Unallocated Space

Unallocated space involves those areas of the storage media that were previously referenced by the file allocation table, but the user has deleted these references, and the area on the disk is now free to be overwritten.

Tools exist that will collect data from unallocated space, and possible search this space for key terms or patterns of data (e.g. email addresses or URLs), but further analysis is labor intensive.

Survey respondents identified the lack of tools that find, bookmark, identify, extract and display specific file formats from unallocated hard drive space. HTML pages, as well as image file formats (JPGs, GIFs, etc.) as an area that needs to be further developed.

Hidden Data Detection

Gap: Lack of Tools for Discovery of "Hidden Information"

Steganography, the practice of invisibly embedding one form of digital media within another, using these 'carrier' files as a way of disguising the presence of the hidden data, is an area in which criminals have a significant edge over law enforcement. This capability is extremely difficult to detect, much less interpret with available techniques. Many of the steganographic tools also employ encryption, which would then render the hidden data virtually undecipherable. The far-reaching nature of the Internet is leveraged, along with the easily obtained software, in order to make the technology perhaps the easiest, most powerful way ever devised to distribute concealed information.

---

[186] AOL's, et. al.

The real dangers of steganography have been revealed by numerous news reports alleging that the planers of the September 11, 2001 terrorist attacks used steganography to conceal their communications. One such example is from a Prime Time Thursday report on October 4, 2001.

"The terrorists responsible for the Sept. 11 attacks may have communicated over the Internet using a computer version of invisible ink that allows secret messages to be concealed in image and music files. Western intelligence officials say they have learned that instructors at Osama bin Laden's camps in remote Afghanistan train his followers in the high-tech secret-messaging technique."[187]

Software packages that perform steganography are widely available, and extremely easy to use. Most packages can be downloaded off the Internet for free. This makes it very easy to download the software, and send proprietary or illicit information over the Internet, with no one being the wiser.

"Steganalysis is the relatively new science of discovering, decoding, and/or rendering useless, covert messages hidden in a carrier file."[188] This is an area in which there is still a significant amount of research being conducted, with relatively little information filtering down to the law enforcement community. Many investigators may not even be aware of this practice or capability.

Until a steganography detection technique is developed, law enforcement will be unable to identify, much less extract information from, files containing hidden documents or images of child pornography.[189]

Hidden Data Recovery

Gap: Lack of Tools/Techniques that Address the Use of Encryption

Criminals are increasingly using encryption to hide their criminal activities. A recent report contains startling estimates of 50 to 100 percent annual increase in the use of encryption by criminals.[190]

---

[187] Ross, B. "A Secret Language: Hijackers May Have Used Secret Internet Messaging Technique." (Available at http://www.abcnews.go.com/sections/primetime/DailyNews/PRIMETIME_011004_steganography.html).

[188] Johnson, N.F., J. Giordano, and J. Sushil. "Steganography and Computer Forensics: The Investigation of Hidden Information." George Mason University, Center for Secure Information Systems, Technical Report: CSIS-TR-99-10-NFJ, Oct 99.

[189] Astrowsky, B. H. "Steganography: Hidden Images, A New Challenge in the Fight Against Child Porn." Update, Vol 13, No. 2, 2000. (Available at http://www.survivorship.org/html/update_stack_7.html).

[190] Leahy, P., Statement at Hearing of Senate Judiciary Subcommittee on Technology, Terrorism, and Government Information on "The Encryption Debate: Criminals, Terrorists, And the Security Needs of Business and Industry", September 3, 1997.

Criminals use encryption for the same reasons most individuals use it, to keep their personal files and communications private. Most operating systems, applications, and Internet communication channels provide a means for the user to encrypt their files. Many systems, such as PGP, are freely available, and many systems employ strong encryption, rendering it next to impossible to break. And with the Digital Millennium Copyright Act, the hands of researchers that would assist law enforcement in their efforts to find a solution are tied. This will make it extremely difficult for law enforcement to investigate any suspect that employs cryptography.

By using the encryption, criminals can decrease the effectiveness of search warrants and wiretaps. There is no way to compel a suspect to give up his/her cryptographic key. In most cases in which an investigator has encountered and defeated encryption, the suspect had left a copy of or provided clues to the key.

Based on answers from our computer forensic practitioner survey, practitioners acknowledged that there is a need for tools that would analyze encrypted information. Currently few tools exist that will detect encryption; much less actually try to break the codes.

With the advancement of encryption technology it is all but impossible to break the encryption with the current tools available, and the development of new tools is not forthcoming in the near future.

Evidence Searching

Gap: Lack of Tools That Conduct "In Context" Searches

There are several tools that allow the investigator to input keywords, terms, and lists of the same. The tool will then return all occurrences of that specific word or term, and mark all of these occurrences, allowing the investigator to further inspect these 'hits'.

While this ability is very useful during any investigation, the subsequent analysis of the returned items is typically quite labor intensive. *All* occurrences of the term are returned, whether relevant or not. If the term happens to be a commonly occurring one within the given environment, the investigator is forced to manually examine and eliminate irrelevant usages.

Respondents felt that there was a need for more intuitive tools that give the investigator the option of seeing the keyword or term in context of its use, either providing the name of the source document (a cached web page containing the term may not be as relevant to the case as a Word document), or providing the surrounding text, thus giving the investigator a quick preview of the usage of the term.

Evidence Correlation

Gap: Lack of Tools that Assist in the Correlation of Intrusion Detection Data

Current Intrusion Detection Systems (IDS) that purport to implement forensic procedures operate by a process known as "signature matching." This involves searching through the available data, looking for known attack patterns. When collected data matches one of the signatures, an alarm indicates the need for review by an analyst.[191]

In most cases, a forensic investigator will start by analyzing the alarms generated by the system, and then will continue to analyze other data such as system logs, and system information. In order to acknowledge the presence of an attack, the investigator must combine knowledge of the system, knowledge of various attack techniques, output from tools used, and information from the system.[192]

Manual analysis of a system takes a considerable amount of time, and requires a highly savvy individual. The skills needed by these individuals include savvy in hardware systems (mostly disk technology), the internals of operating systems, and network protocols.[193]

To further complicate the analyst's task, the effectiveness and technology behind automatic signature matching tools varies a great deal. Some of the signatures accurately identify an attack, while others are too general and generate many false alarms. The profile of the attack can also determine the effectiveness of the signature. There can be the obvious use of publicly available scripts, or there can be well-hidden custom attacks developed by a highly savvy individual. These signature systems also fail on new attacks that have not yet been analyzed, and which do not have signatures developed for them. IDS systems are like virus scanners and forensic tools because they are dependent on updates. On top of that, signatures are not available until after an attack has already been completed.[194]

Evidence Examination and Analysis

Once the digital evidence has been imaged and extracted to a central location, the cyber forensic investigator must examine what has been collected. The challenge is to sift through and eliminate the extraneous data as quickly as possible, leaving only the pertinent digital evidence. Most of this should be completed during the extraction phase through the use of techniques such as known file filtering. During the examination phase the investigator uses the available tools to further refine his searches to target specific digital evidence.

---

[191] O'Boyle, T., and B. Hill. "Cyberspace Detectives Employ Intrusion Detection Systems and Forensics." (Available at http://www.mitre.org/pubs/edge/february_01/oboyle.htm).
[192] O'Boyle, et. al.
[193] O'Boyle, et. al.
[194] O'Boyle, et. al.

Many of the same search type tools that are used during the extraction phase are used during this phase. The difference is that the investigator has refined his search at this point so that it focuses on specific pieces of relevant data or information. The variety of tasks required of the investigator pose a number of gaps that need to be addressed.

1. In order to view all available information, an investigator may need access to virtually every application currently used to create the data. A universal information viewer is needed.
2. At this time, there are no standard taxonomies of words, phrases, data formats, or data organization that can be applied to specific crimes under investigation and used to search the data. Without them, data searching is rudimentary and very time consuming.
3. There is also no tool available that allows investigators to identify possible authors based upon their known writings, i.e. vocabulary, grammar, or style.
4. While an investigator is under obligation not to examine privileged information, there is currently no way for an investigator to know that what he is about to read is privileged, without actually reading it.
5. Investigators are also in need of forensic tools that can assist in decrypting data, breaking passwords, or accessing protected information contained in electronic organizers, which are becoming more popular.
6. If there is any back door access to devices, the investigators currently must procure it from the manufacturers and software vendors. A comprehensive database for investigators would expedite the process.

As the use of digital evidence becomes more prevalent in court cases, it will be increasingly necessary to develop cyber forensics tools to meet these needs.

Evidence Time Lining

Gap: Lack of Source of Trusted Time

There is an overwhelming need by law enforcement, as well as all users of networked computers, to have access to a source of trusted, reliable, verifiable time.

Whenever the investigation of an incident involves more than one system, whether this is a system that has been attacked, the system analyzing the attack, or any other combination of more than one computer system, having trustworthy timestamps for data and events is critical.

It is of the utmost importance to have a common time reference across each network device and host involved in handling the incident. This is particularly true for investigators who travel to sites to analyze an incident, as their portable systems need to be synchronized with all of the local systems on the network under examination. Although this may sound trivial, it can be of vital importance in diagnosing a highly technical network issue, not to mention demonstrating an incident chronology in court.

On more than one occasion, having inadequate timestamps on systems and their associated files has disqualified evidence. No existing tools actively support this issue. Confidence in timestamps will remain a point of contention for some time to come until a solution is implemented that will meet the rigorous criteria of the courts.

Changing the element of time on a computer or network is a simple task for an experienced user. "It is possible that a malicious originator of data could "backdate" the time of transmission to make it appear as if it took place prior to certificate expiration or to the placement of their certificate."[195] This can create many problems for evidence correlation and time lining of evidence. The absence of the trusted time stamping function makes time lining and evidence correlation difficult, and hard to prove the link between 'activity and evidence' in court. In order for digital and network forensics to take the next step forward, a trusted time stamping protocol must be developed.

Although most of the practitioners from the survey claimed that trusted time has not been a problem, they did indicate that it would be in the near future. One of the practitioners is quoted as saying, "It hasn't been a problem yet, but I can see it as an issue in later cases, trial, and testimony."

The problem could involve the courts challenging the integrity of evidence that has been extracted. Time lining the evidence, and establishing a chain of custody could be issues that could determine the viability of certain evidence in a court of law.

Evidence Mining

Gap: Lack of Tools/Techniques for Analysis of Distributed Systems

When investigating a crime that involves a network (e.g. a Local Area Network), many problems arise. First, the investigator must determine the topography of the entire network. This involves determining which computers are connected to each other, what peripherals are available to these hosts, which hub, switch, or router they pass through, and so on. Then they must document the operating systems that are being used on the different computers, and what file system scheme each of the operating systems are employing.

When networks are involved, it is possible that information could be stored on any one of a number of machines; or, when the source of a particular activity is being investigated, it may have originated from any of the computers in the network. The information being sought is not always on the computer that the suspect sits in front of. This means that the investigators may have to search all of the computers in the network to try and find clues and leads.

---

[195] PKI PMO Public Key Infrastructure, United States Department of Defense. (Available at http://www.c3i.osd.mil/org/sio/ia/pki/faq.html).

Investigative tools are needed that assist in the mapping and display of the entire layout of a network. It would be of great value if the tool could show the route taken by network traffic through the different hosts, and possibly indicate the operating system running on each, or even indicate the purpose or owner of that particular system. There are utilities that are available that perform these individual tasks, but none specifically designed for or used by law enforcement.

Network Forensics

There are serious gaps in current network forensic tools and methodologies. Limited data reduction capabilities make it difficult, if not impossible, for investigators to sift through enormous amounts of data. The absence of trusted time stamping of events makes time lining and evidence correlation difficult. Weak evidence preservation may allow valuable evidence to be destroyed or tainted. There are many privacy concerns regarding the wholesale collection of all network traffic, e.g. Carnivore. Limited evidence reasoning tools require the investigator to have a high level of technical savvy and experience with networks. Automated session reconstruction tools are crude and often do not work for many applications and types of traffic. As the use and misuse of networks increases, development of tools to address these gaps and limitations becomes crucial.

Gap: Lack of Tools to Identify Users of Chat Networks (IRC, ICQ, IM)

Chat networks are often used in the identification of potential perpetrators of child pornography crimes. Law enforcement officials will commonly enter chat rooms disguised as a child, and try to lure pedophiles into sending child pornography to them.

A problem the officials have, is tracking the pedophile down once they have enough evidence to arrest them. This is where a tool used to identify a user of a chat network would come into use.

Many of the respondents stated that there was a need for a tool to complete this task. This would enable the investigators to quickly and accurately track down the location of the perpetrator so an arrest can be made.

Evidence Organization/Case Management

While there are several methodologies currently being used to organize and manage digital evidence, the available tools have limitations that can curtail the effectiveness and efficiency of an investigation. As of yet, there are no tools available that automatically correlate non-digital evidence with digital evidence (including phone records, credit card receipts, eye witness testimony, Internet Service Provider (ISP) records, or other forensic evidence). The programs that are being used have limited capabilities to correlate the evidence from computer break-ins. There is no tool that can effectively correlate computer information from the same computer or case, or make associations among cases or evidence files. The manual completion of such links is extremely labor intensive and

time consuming. In order to have the most effective and efficient investigations possible, new tools must be developed to meet these needs.

Gap: Enhanced Reporting Capabilities

Current forensic tools lack the ability to correlate the collected evidence with the set of steps taken to collect and extract that evidence.

Reports produced from current forensic applications (those that provide reporting capabilities) display the ultimate findings, in the form of case or evidence files that document the occurrence of evidence on the system. But, they do not allow the inclusion of the processing steps involved in the forensic analysis.

The inability to include this information in a report could be a serious point of contention when the investigator is called upon to testify to the process taken to achieve the results. It must be shown that the steps were logical and repeatable; the investigator must be able to establish that further analysis utilizing the same steps will yield the same results.

Future Tools for Cyber Crime Prevention

This section identifies future tool requirements based on the research analysis and the survey response. Recommended strategies and resources to assist law enforcement in closing the gaps of current tools are included.

Automated and "Intelligent" Tools

More tools need to be automated tools to improve performance. These tools would automatically perform a series of functions, allowing the investigator to focus on more important tasks. "Automated tools can be developed to help investigators meet standards of performance, speed, completion of tasks, and reduce the need for costly training."[196] A search of an employee's work for a few months on an automated tool can take a few hours to a day or two at most, while the same search on a manual tool can take weeks.[197]

The lack off automated tools is even more of a problem in the Unix environment than in the Windows environments. Most of the time, investigators working in Unix environments have to manually type in commands in order to get functions like file extension searches to operate. This makes the investigations very cumbersome and time consuming.

Automated hacking tools have made it very simple for criminals to find targets and perform malicious acts. A cyber criminal does not need a lot of skill and knowledge in order to perform these attacks. Instead they rely on the manufacturer's knowledge built into the tools that they use. The power and expertise incorporated in these tools increases the number of possible attackers, and makes them even more sophisticated.[198]

In order for law enforcement to keep pace with these criminals, there needs to be automated investigation tools developed to counter the automatic hacking tools. An early example of an automated investigation tool in the Unix environment is the Coroners Toolkit (TCT), which speeds up and standardizes the process of making a digital-forensic examination.

However, automating tools is just the first step. Intelligent tools need to be developed that not only filter out unimportant information but key in on the information that is critical to the investigation.

---

[196] JPS: Job Performance Systems, Inc. "Automated Tools." (Available at http://www.jobperformancesystems.com/automated%20tools.html).
[197] The Journal of Public Inquiry, Inspectors General of the United States, Fall/Winter 1997.
[198] The Challenges for Law Enforcement and Revenue Agencies. (Available at http://www.austrac.gov.au/text/publications/rgec/1/word/report-part2.doc).

Advanced Preservation Tools and Media

In the computer forensic community there is a need for more advanced preservation tools and media. Currently there are only a few tools available that will successfully preserve information on a computer system. PDBlock is one of the only software tools that will successfully prevent unexpected writes to a physical disk drive.

There are more hardware write blocking tools available then there are software tools. Some of these hardware tools consist of ARS, Drive Lock, and FastBloc. The hardware tools do have problems that are associated with them.

Originally, these hardware write blocking tools were not designed for forensic purposes. These tools were designed to act as a bridge between IDE and SCSI drives. Over time, users have adapted these tools to be used in forensic investigations. Also, there is no technical support for most of these tools. The manufacturers of these tools are from overseas, and do not offer support. Most of the web sites for these tools are not even in English. These tools are also only components, and not completed products.

In the future there is going to be a need for more tools that will successfully preserve information in a computer forensic investigation. Integrity is one of the most important qualities of the information that is collected.

Multi-Format Evidence Viewers

There is a need for multi-format evidence viewers. Criminals continually use obscure file formats in their malicious acts in order to confuse law enforcement investigators and halt investigations.

There are a few multi-format viewing programs available today, including some that can handle some popular CAD data formats as well as common graphic image data formats and general office data formats, but there are not nearly enough available to law enforcement.[199]

Having multi-format viewers available integrated into forensic tools will be a very important advancement for digital evidence investigations and analysis.

Multi-Platform Support

There is an immediate need for more forensic tools that have multi-platform support. In the evidence collection and preservation category, there are only a few tools that have multi-platform support. In the evidence extraction category, The Coroner's Toolkit (TCT) is one the few tools that provides multi-platform support for the following

---

[199] Harrod, G. "IntraVISION 3D Viewer Review." Spatial Technology, Inc. (Available at http://www.cadinfo.net/reviews/intravision.htm).

operating systems: DOS, Windows 95/98/ME/NT/2000, Macintosh, SunOS, Unix, Novell, and OS/2 systems.

Within the computer forensics community, there is strong consensus that multi-platform tools can save a lot of time and money.

Steganography Detection Tools

There are only a few tools that will accurately detect steganography. These detection tools are emerging technology and some have a high level of false positives. Little work has been done in the area of recovering the hidden message once steganography has been detected.

One approach is called blind steganography detection, and was developed by WetStone Technologies, Inc.

> "The Blind Steganography Detection method was developed under a program funded by the DoD. Our research focused on the characteristics and attributes of image, audio, or video files (including color, intensity, saturation, hue, frequency, tone, noise and distortion) and developed a mathematical definition of normal for each type. Once the signature, or "normal statistical characteristics" was mathematically expressed, that expression was compared to a given image and specific deviations provided clues as to the purity of the image data."[200]

Another approach is signature based, using the know signatures of steganography hiding programs to detect their use.

George Mason University's Neil Johnson is building a stego-detector; a program he says examines hard drives "like a virus scanner" and identifies the electronic fingerprints sometimes left by steganographic applications. "Different authors have different ways to hide information to make it less perceptible," Johnson says. "The author may come up with ideas that nobody else is using. That tool may have a special signature. Once that signature is detected, it can be tied to a tool."[201]

Steganography detection tools and methods are needed now by law enforcement to investigate a wide range of cases. "The development of such software is imperative to the battle against on-line child pornography."[202]

---

[200] Hosmer, C. "Steganography Overview." (Available at http://www.wetstonetech.com/).
[201] McCullagh, D. "Secret Messages in .Wavs." (Available at http://www.wired.com/news/print/0,1294,41861,00.html).
[202] Astrowsky, et. al.

Encryption Detection and Extraction Tools

With the increasing use of cryptographic techniques, tools need to be developed to provide law enforcement specific information about how it is being used.

"These tools must have the ability to identify the presence of cryptographic data, identify the algorithms, key sizes, key management, and access control techniques, estimate the best techniques to extract information from the cryptographically protected data, and provide an estimate of the value of the potentially extractable information based upon other corroborating information."[203]

Secure Distributed Evidence Repository

Another tool that would be very beneficial to the military, law enforcement, and those who design computer forensic tools, would be a secure distributed evidence repository. This database would be the source of forensic data coming from intrusion detection programs on networks from many different locations, and from previous computer forensic investigations. This tool could also be used for multi-jurisdictional cyber investigations where investigators need to share evidence in a highly secure web based environment.

One example of such a tool is SI-FI (Synthesized Information from Forensic Investigations).[204] This tool provides a secure web based environment where data can be posted so analysts can examine, search, correlate, and graph information on various attacks that have occurred. With data mining capabilities, analysts can search for common trends among the various attacks. This data could provide analysts with answers as to how these attacks were performed, and eventually ways to stop the attacks from happening again. This would be a great test bed for forensic tools, and a good research and development tool to aid designers in figuring out how the attacks are being performed, and what tools need to be designed to combat these attacks.

Comprehensive Database of Intrusion Vulnerability and Attack Signatures

The DCFL has recently expressed a need for a database that will incorporate attack signatures from various cyber crime tools to assist them in their forensic work. This database could be used to perform post mortem string searches on a compromised system. The string searches will identify instances of attacks.[205]

If a database of the signatures were completed, investigators would not have to go through the rigorous process of writing their own. This would save them valuable time and money.

---

[203] Hosmer, C. et. al., "Advancing Crime Scene Computer Forensic Techniques." (Available at http://www.wetstonetech.com/crime.htm).

[204] Currently being developed by WetStone Technologies, Inc.

[205] Bartholomew, D., "Standardized Naming and Databases." DCFL.

Linux Based Tool Suites

With the growing popularity of Linux systems, digital forensic investigators have an immediate need for computer forensic Linux tool suites. Currently there are many suites of forensic tools available, but most of these suites are Windows based.

These suites combine all of the tools needed for an investigation. This provides for fewer steps, more possibilities for automation, thus resulting in significant time savings. Currently there are no full Linux suites available on the market.

Network Forensic Tools

There is a need for more reliable, sophisticated, and automated network forensic tools. There are currently only a handful of good network forensic tools available, and these tools have their problems and weaknesses.

These forensic tools can be very helpful to network administrators and law enforcement officials, but there are a few barriers that inhibit the use of some of these tools.

The first major hurdle involves the collection and imaging of the digital evidence. If a server is connected to a network, it may not be shutdown or removed without severely affecting the entire network. This is especially true for servers, where much of the digital evidence is likely to be stored. Under most circumstances the investigator must work with the system administrator or other company officials to plan the evidence collection and imaging when it will cause the least disruption to the system.

The second major problem is the sheer volume of data that can be stored on a network. Limited data reduction capabilities make it very difficult for investigators to sift through enormous amounts of data. The investigator will generally need to sift through a lot of insignificant data to find the valuable evidence. There are tools available to help in this task, but most are still relatively unsophisticated and require the investigator to do most of the evidence extraction and evidence searching manually. Even a small network of a dozen computers can contain several hundred gigabytes of data that need to be imaged, extracted, and searched.

A third difficulty is the transitory nature of network data. Weak evidence preservation may allow valuable evidence to be destroyed or tainted. The amount of data traffic causes lots of disk writes, which can result in the overwriting of potential evidence. There is also the problem of log files being automatically deleted after a certain time period or when they reach a certain size. This is especially a problem where disk space is tight. Any delay in data collection could mean the loss of critical evidence.

The final major hurdle in network forensics is the lack of trust. Many logs are easily modified or deleted. There is also the problem of the perpetrator assuming the user ID, and therefore the identity, of an innocent party. Trust in these forms of digital evidence

can be increased if logs are encrypted and/or digitally signed, and if there are strict and secure access control procedures.

Along with the four major gaps with network forensic tools, there are a few other limitations. These include privacy concerns regarding the wholesale collection of all network traffic, e.g. Carnivore, limited evidence reasoning tools require the investigator to have a high level of technical savvy and experience with networks, automated session reconstruction tools are crude and often do not work for many applications and types of traffic, and one source corporate memory is also unavailable.

As the use and misuse of networks increases, development of tools to address these gaps and limitations becomes crucial for law enforcement.

Tools to Collect Volatile Evidence

Some evidence on a computer system will not last very long, and could be lost by simple power shutoffs. This evidence relies on a power source, and is lost if the power is shutoff. All of this information is called volatile evidence. This volatile evidence could be very important in an investigation, and should receive utmost attention when an investigation begins.

In an investigation, evidence should be collected in the order of most volatile, to least volatile. An example of most volatile to least volatile is as follows: registers and cache, routing tables, arp cache, process table, kernel statistics and modules, main memory, temporary file systems, secondary memory, router configuration, network topology.[206] Volatile evidence is information like memory, active processes, active network connections, contents of a computer screen, and fingerprints.[207] Non-volatile evidence is information like physical equipment, persistent storage, printouts from various audit and monitoring logs, and recorded video surveillance. Unlike volatile evidence, this evidence will not disappear soon. After the most volatile evidence is collected, an investigator can then power down the system and continue on with their normal operations. This might include making a bit stream copy of the entire hard disk drive.

The Coroner's Toolkit is being developed to enable a user to collect both static and volatile evidence from a computer system.[208] This is one of the few tools that perform this kind of information collection.

Law enforcement investigators need more tools like The Coroner's Toolkit to combat the criminal's actions. Some criminals will install various viruses and Trojan horses that will erase volatile information on their computer when the power is turned off. In order to prevent this, the investigators need more tools to collect this information while the computer is still running.

---

[206] Braid, M. "Collecting Electronic Evidence After A System Compromise." April 17, 2001. (Available at http://rr.sans.org/incident/collect.php).
[207] Winterton, E. Incident Response Fundamentals Class, Arca Systems, Inc.
[208] Casey, E. Computer & Internet Crime FAQ, Knowledge Solutions.

Concluding Remarks

Remarks by Ronald Stevens, Director of the New York State Computer Crime Unit, articulate the problems and issues law enforcement faces.

> "Forensic examination of digital evidence is becoming more crucial in the investigation of crimes facilitated by the use of technology. A growing number of investigations involve crime where critical evidence is stored on digital media. Whether the case is criminal, civil or administrative, processing digital evidence requires technically skilled personnel with specialized training and equipment. As the volume and complexity of casework grows, it will become increasingly important for additional resources be made available in a more efficient and effective manner.
>
> An examination of current practices in the processing of digital forensic evidence find that the urgent and time-sensitive cases receive top priority; evidence in important investigations is analyzed and completed in a short period of time. However, low priority cases tend to have slower processing times due to the sheer volume of cases and available resources. At the current pace, it may require up to 18 months to clear just the pending cases at current staffing levels, if no additional cases were received.
>
> The lack of accepted standards and procedural uniformity in the area of computer forensics has prompted independent responses from a variety of law enforcement agencies at the State, Local, and Federal levels, as well as disparate training programs being offered by private interests, as well as universities. If this disorganized approach continues, defense challenges could call into question the credibility of computer forensic analysis due to the lack of standardization. The science of computer forensics must evolve to meet the same standards of evidence that have been established for other forensic disciplines such as fingerprinting, ballistics, drugs, and DNA."[209]

The challenges faced by law enforcement require new tools and methods. An August 2000 NIJ report concluded, "There is a significant and immediate need for up-to-date technological tools and equipment for State and Local law enforcement agencies to conduct electronic crime investigations."[210] The dangers and threats of the post September 11th world require that this be done with all due speed. These new tools and methods include:

---

[209] Stevens, R. R., Statement of Ronald R. Stevens, Director of the Computer Crime Unit for the New York State Police, On Cyber Crime in New York State, Before the House Committee on the Judiciary Subcommittee on Crime, May 24, 2001.

[210] Stambaugh, H., et. al, Electronic Crime Needs Assessment for State and Local Law Enforcement, National Institute of Justice Report, Washington, DC: U.S. Department of Justice, March 2001.

- Standardization and certification of tools to increase trust, integrity, and reliability;
- Integrated, automated, and intelligent suites of high performance tools that are not merely point solutions;
- Ability to review massive amounts of data to eliminate data outside the investigative realm and to quickly identify potential digital evidence relevant to the case;
- Secure and trusted time built into tools and methods;
- Advanced steganography and data hiding detection and recovery tools;
- Advanced network forensics tools that provide real time versus post-mortem analysis, and
- Adequate training on new tools and methods.