Volume 2, Issue 1, January 2015 ISSN: 2348-8565 (Online)

International Journal of Modern Engineering and Research Technology

Website: http://www.ijmert.org Email: editor.ijmert@gmail.com

Forensics Analysis-Technique for Cybercrime

Ravindra Kumar Gupta
Research Scholar
Department of Computer Science & Engineering
Vindhya Institute of Technology and Science
Jabalpur (M.P.), [INDIA]
Email: ravindra.gupta38@gmail.com

Sanjay Gupta
Head of the Department
Department of Computer Science & Engineering
Vindhya Institute of Technology and Science
Jabalpur (M.P.), [INDIA]
Email: Sanjiit@rediffmail.com

ABSTRACT

Forensic analysis techniques provide a systematic approach of investigation. This helps to ensure the overall integrity and survivability of network infrastructure. If you consider computer forensics as a new basic element i.e. "defence-in-depth" approach to network and computer security then you can help in an organization's data integrity. The computer forensics must be practiced responsibly otherwise, there is a risk of destroying vital evidences or forensic evidence ruled inadmissible in the court of law. Also, you or your organization may run afoul of new laws that mandate regulatory compliance and assign liability if certain types of data are not adequately protected. Identification, collection, presentation and analysis of data such that integrity of evidence collected is preserved and can be presented effectively in the court of law.

Keywords:—*Cyber crime, Forensic Models, Forensic Analysis.*

I. INTRODUCTION

There are two typical aspects of a computer forensic investigation. Firstly, the investigators have to understand the kind of potential evidence they are looking for such that they may identify the areas to be searched. Digital crimes have a large spectrum of variation in criminal activities from child pornography to theft of personal data to destruction of intellectual property. Secondly, the tools to be used for investigation are given proper importance. Recovery of detailed damaged or encrypted files is required so investigator must be familiar with all such approaches. The collection of ephemeral data is one of the important tasks that are required to be performed by an investigator. The data which is present in hard disk and secondary storage is persistent but the data present in RAM, registers, cache is volatile. For doing investigation, manual process takes long time and energy, it is not possible to complete the process with in short spell of time. Cyberspace has no specific jurisdiction; therefore, criminals can commit crime from any location through computer in the world leaving no evidence to control[1]. When someone—steals data from cyber space or uses information for unintended purposes, it is called cyber crime. With the increase usage of computer technology, cyber crime is on the rise. Like any crime, cyber Crime should be investigated and prosecuted where necessary. Computer forensics describes the practice of retrieving evidence in the form of data from a computer that relates to a crime in a manner that meets the requirements of the given legal system. Computer forensics evidence needs to be handled with the same care that physical evidence requires. However, there is added complexity due to the technical nature of computer based technology and has added another dimension with digital evidence. As greater emphasis is placed on digital evidence, it becomes increasingly critical that the evidence be handled and examined properly.

II. Forensic Techniques

Computer forensics refers to the legal processes, rules of evidence, court procedures, and forensic practices used to investigate e-Crimes. Specifically, computer forensics is the application of scientific, forensically sound procedures in the collection, analysis, and presentation of electronic data. For computer evidence to be accepted in a court of law, the forensic investigation process must identify, preserve, examine, and document any computer evidence retrieved. Computer evidence is entirely different. It cannot be seen, touched or smelled and it often lasts for only very short periods of time. Computers typically store data in three ways, magnetic, semiconductor, and optical. Other less common data storage methods include magneto-optical disk storage, optical jukebox storage and ultra-density optical disk storage. Potentially significant new developments in technology suggest that techniques like phase-change.

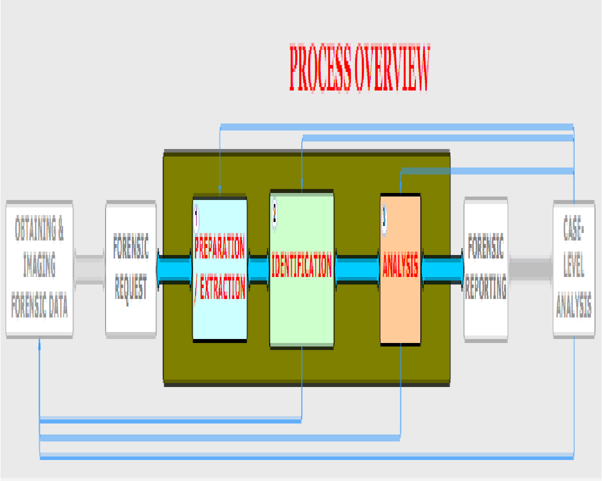

Figure 2: Process Overview

III. Forensic Duplication

Forensic duplication is the copying of the contents of a storage device completely and without alteration. The technique is sometimes known as bitwise duplication, sector copying, or physical imaging. Forensic duplication is the primary method for collecting hard disk, floppy, CD/DVD, and flash-based data for the purpose of evidence gathering.

Copying files from a suspects device using standard techniques (Windows Explorer, cutting and pasting, x copy) or imaging of logical drives (using Ghost or Drive Image) provides some of the data for an investigation but is usually insufficient for forensic imaging and may violate best evidence rules.

The failings of standard duplication techniques from a forensic standpoint are as follows:

Lack of authenticity. There is no verification of authenticity in a standard file copy. This can be addressed through the use of external tools such as MD5sum that provide this facility.

Loss of non-file data. Information stored in slack space, un-partitioned space, or free space is not copied. These locations may contain previously deleted content or other information of interest that will not be available with logical imaging.

Alteration of metadata. Depending on the file systems copied To/From, metadata associated with a file may be lost. Rights and permissions stored on a specific file system (for example, NTFS) as well as system attributes (for example, Read Only bit) may be altered or deleted when a file is copied between disparate file systems. This can include the loss of the ability to look up permissions (based on SID) even on copies to similar file systems.

Inability to provide context. A copy of the data in a logical file does not provide the same machine context as an image. Contextual data can include location in a directory tree, or duplication details, and details on other surrounding files.

Failure to copy all data streams. Alternate Data Streams, a feature of NTFS file systems, are not supported by most other file systems. By copying files to a non-NTFS file system, these streams are lost, as only the primary stream is retained.

IV. Windows Registry Analysis

There are four primary methods of analyzing the registry in a forensic analysis of a system (five if the real-time analysis provided by Regmon is included):

Perform a live system analysis graphically. This method is the easiest but the least forensically sound. Opening the registry using regedit on the target system will show the easiest to traverse view of the registry, but allows for the accidental or intentional altering of data when viewing, in addition to a very small amount of memory overwrite (or paging file) to run the program (around 250K).

Perform a live system analysis using the command line. Command line analysis allows for a lower profile (and less risky) gathering of registry information, with some loss of interactivity. The reg command can be used to gather specifics on a given set of keys determined beforehand and placed in a batch file or an entire registry section (recursively printing values) as needed.

Perform a live systems analysis remotely. Regedit provides the ability to connect to a remote registry, provided the appropriate permissions are present. This enables an administrator to examine a registry on a remote system without directly alerting the user. Likewise, the use of a Null session connection (such as that used by Superscan from Found-stone) will allow the enumeration of several registry keys, including user lists and information on current users.

Perform an offline analysis on registry files. EnCase is able to parse the raw registry files on acquired drives. This allows for the offline analysis of registry information in a completely forensically sound manner. Although this is the most sound mechanism, it is also the least fruitful. Any dynamic information is lost, the structure is more difficult to navigate, and the links are missing (for example, the HKEY_CURRENT_USER linkage).

When analyzing the registry, numerous values are relevant to a wide range of investigations. A few of the key values, and the types of investigations they are relevant in, are detailed in the following sections.

Registry Basics

The Windows registry is a hierarchical database of configuration values stored in a proprietary file format. Within the files are an organized set of hives that form the building blocks of the registry. Under each hive is a list of keys. All keys have a name and may contain multiple name/value pairs and subkeys.

General

Several registry keys are examined in numerous types of investigations. They are not necessarily specific to a given area; however, they are relevant to a number of investigations. These keys include basic system information (who used the system and what applications are installed) and more detailed information on key system areas (what hardware was installed and what drives were mounted).

Folder Locations

The key folders on Microsoft systems are the most likely location for files of interest in an investigation. These folders include the My Documents folder (and My Music/My Pictures folders), the Startup folder, the Recent folder, and the Internet folders (Cache, History, Favorites and Cookies).

By confirming the locations of these folders, an investigation can be directed at target locations for initial analysis.

Determining what files, folders, or applications were most recently used is a key task in investigations. Showing that an individual opened a file, saved a file, or searched for a file can prove the suspect knew the file existed (or even created the file, in the case of Save As lists). Sometimes a suspect will delete a file after viewing it. Unless explicitly cleared, the file name may still appear in the Most Recently Used (MRU) registry keys.

In addition to the Microsoft-specific MRU lists, installed applications may have their own most recently used keys. The most likely location for these is under the HKEY_CURRENT_USER\Software\ *AppName* hierarchy.

Start-up Items

Spyware, viruses, and other malicious code will frequently continue to infect a computer after a reboot. To accomplish this, the code needs to be run automatically unless it is associated with a file the user is expected to reopen frequently, such as a mail file or common executable. The Windows Registry contains numerous locations from which code can automatically be run; Anything suspicious found in these keys may require further investigation, but many legitimate programs also make use of these keys.

Intelliforms

Microsoft Internet Explorer 5.0 introduced the Auto complete feature to allow users to easily store and automatically fill out form entries online. Auto complete uses a Microsoft technology called Intelliforms. The technology itself matches the name on a form input field to a group of values stored in the registry. These registry entries are then queried when a website with similar fields is opened.

Intelliforms can store all information typed into forms on a web browser. This includes credit card numbers, passwords, addresses, and other pieces of information critical to an investigation. To protect this information, Microsoft encrypts any registry entries stored for Intelliforms use and places them under the registry key HKEY_CURRENT_ USER\Software\Microsoft\ Protected Storage System\ Provider\. This key is accessible only to the SYSTEM account by default, and the data is not viewable except in encrypted form even with the proper permissions granted.

V. FILE SYSTEM ANALYSIS

Searching, validating, recovering, and analyzing the contents of an imaged drive are the most common forensic tasks performed by an examiner. Since the largest portion of evidence generally resides on a hard disk, this is where the most effort is spent and the most rewards are found in a large number of forensic scenarios.

File system analysis covers the examination of Windows-compatible file systems (NTFS, FAT, CDFS, and so on), as well as non-file space on a drive (slack space and unallocated space). Many of the listed techniques may also be applied to non-Windows file systems and un-partitioned areas. These are noted where appropriate.

Searching

The most common forensic activity is searching a hard disk for strings of data. The searching can be file-based or slack-space–based, and there are even searches of unallocated space. As in other forensic tasks, the context of the investigation determines the search type used. There are two primary search methods: index-based searching and bitwise searching.

Index-based searching generates a keyword index on the first pass through a series of files. Bitwise searching performs a full, regular expression–based search on the raw data, file-specific or not. An index-based search may be used to provide quick, repeated searches with new terms on files copied from a shared drive. Conversely, a full bitwise search may be more relevant if a hard disk is being searched for deleted files or residual fragments of their contents.

Index-based Searching

Index-based searches rely on the creation of an index of keywords based on the contents of files. A search tool generally opens all files on a drive/share/image/partition, searches them for repeating strings of printable characters, and creates a table of the repeated strings with pointers to the original content. The initial indexing can take hours or days. However, when completed, searching the index can be done in near-instant time.

Bitwise Searching

Bitwise searches look for simple text strings or regular expression matches in any sectors on a drive including both sectors that are currently unallocated and those residing in OS slack space. The ability to do regular expression searches enables the examiner to search for complex text terms as well as non-text (binary) values such as file headers. This precludes indexing, and as a result the entire contents of the drive must be searched for each term, making bitwise searches significantly slower than index searches.

Search Methodology

Many investigations call for a combination of searching techniques, and the methodology applied to a particular case needs to be context specific. Factors affecting the choice and order of searching include:

Awareness of suspect. If the suspect is aware that he is under investigation, file-based content may have been deleted, which leans toward bitwise searching. If the content is likely to still be present on the drive intact, index-based searching may be more effective.
Likely data format. If there is a chance that the content resides in PDF, XLS, zipped, gzipped, or Windows-compressed files, index -based searching will be more thorough. A preliminary bitwise search for the header bytes from these file types and subsequent recovery of deleted files before the index-based search will combine both techniques for the maximum effectiveness.

Time constraints. For a single keyword or group of keywords, a bitwise search will be slightly faster in most circumstances due to the overhead created by indexing. If subsequent or real-time searches are expected, which is common when performing an investigation interactively with non-IT investigators or Subject Matter Experts, index -based searches will be faster overall.

Search complexity. When searching for very complex regular expressions (for example, looking for all strings that match a credit card number or phone number), bitwise search tools will have more success. Searches based on synonyms of words, phonetically similar words, and fuzzy spellings of words will be more successful with index-based searches.

VI. Log File Analysis

The two areas of Microsoft logging of most interested to the forensic examiner are the standard log repository for system, application, and security events (Event Log), and the key Internet server–based log files (HTTP, FTP, and SMTP). The Event Log details provide insight into what a given user was doing on a machine or what the machine itself was doing. The Internet server log files show what remote activity was attempted or successful on a given system.

Event Logs

Microsoft's answer to Syslog, event logs retain key log event details on Windows systems. The event logs are broken into three areas: application logs, which store information on individual applications; system logs, which maintain operating system event details; and security logs, which hold data on logins/logouts and other security functions.

The information that is stored in logs is very useful, particularly when it comes to gathering forensic evidence related to intrusive actions, fraudulent behaviour or malicious attacks. Accordingly, log-files are important sources of forensic information because they usually connect a certain event to a particular point in time. So, this time we will delve a bit more deeply, and investigate Microsoft Windows Event logs.

The Windows operating system comes with a complex architecture with which to handle events such as logging on requires proper security measures. It is also possible to trigger the writing of an action to the log when a specific type of event occurs. Firstly, the System and Application logs can be used in a number of ways by both applications and the Windows operating system. However, there are conventions about how to write specific events to the log. Secondly, there is a special type of logging: security logging, during which security related events are written to the separate security log. This log can only be directly written to by the Local Security Authority Subsystem Service, or LSASS.

Internet Logs

Windows servers log access information on individual requests that arrive through their respective services. Connections to and activity on FTP, HTTP, and SMTP (the main Internet Information Server services) are frequently used in forensic examinations of server use or compromise. The logging for these is turned on by default when the service is started and can provide a wealth of information on Internet-based system activity.

HTTP Logs

Internet Information Server (IIS), Microsoft's built-in web server, is one of the most commonly used on the Internet (second to Apache) and even more common on intranets.

The web logs for IIS websites are stored by default under %SYSTEM-ROOT%\System32 \Logfiles\W3SVC, and each date is provided its own log file (although dates with no accesses will have no log file). Administrators may change the directory for performance and back-up reasons as well as the file cycling frequency.

The default log file settings contain several key fields that are indicated at the start of the file (shown under the #Fields metadata for the example file that follows). The key fields of interest in an examination are:

Date and time (date time). The date and time, listed in GMT unless an offset is included, is the current time on the server when a request was made. If a web site defacement occurs, look at the first request after the defacement. Many times an attacker will use a shell account from a compromised machine to launch an attack, and then view their handiwork from their home machine. Referrer (cs-Referer). The Referrer is not included by default with several log formats but should be added to all web logs for forensic use. The field shows the previous site visited before the server in question, giving a possible homepage location or additional clue to the identity of an attacker. The referring site can include a query string as well, useful if the previous site was a search engine.

VII. Internet Usage Analysis

The most common investigation in most corporate settings is inappropriate usage, with inappropriate web surfing being the most prevalent form. Depending on policy, inappropriate use may be defined as any usage of computing resources for personal use. Additionally, the viewing of certain types of content may be considered inappropriate. Most corporations frown on employees browsing pornographic materials, hacking websites, or playing on gambling sites using company resources. Finally, storing or exchanging copyrighted material through peer -to-peer services is growing in popularity and is becoming a serious issue in corporations.

In addition to inappropriate usage, an individual's Internet activities may be used to prove or disprove other crimes. Showing that an individual was on the Hotmail website the same time as a harassing email was sent, finding evidence of a POST to a message board of a questionable nature, or noting numerous visits to target website may all be of value to a general investigative response.

Web Activity

Most corporate web activity now takes place using Microsoft Internet Explorer, which holds approximately 92 percent of the browser market share. Until recently, Internet Explorer dominated the browser market almost exclusively, but a recent push by Firefox has reignited the Mozilla-based client interest. Capturing almost 6 percent of the market and growing, Firefox is positioned to play more dominant role in future forensic investigations.

VIII. Email Investigations

The contents of email messages have incriminated many individuals (and lead to some embarrassing disclosures). When an email is created, the sense of permanence associated with penning a letter is not always at the front of the mind. As such, individuals convey things in emails that they would never put to paper and ink. The medium is treated more like a telephone conversation. Unfortunately for the individuals that treat it as such, this is not the case, and email communications differ greatly from phone communications. Emails are rarely transient. They are stored on the sender and recipient(s) machines as well as any mail servers used, at least until viewed and securely erased. Emails are visible in transit unless explicitly encrypted. They can be routed accidentally to the wrong persons, and they and can be forwarded beyond the intended, original recipients.

Email is a store-and-forward protocol. Copies of messages are actually stored to disk in most cases, on all mail relays between the sender's SMTP server or internal mail server (for example, Exchange), and on the recipient's home POP3, IMAP, or internal mail server.

Email investigations cover the gamut of computer crimes and policy violations. Inappropriate material may be transmitted through email, it can be used as a harassment tool, an impersonation tool (for example, phishing scams), or a communications tool for other activities. Forensic analysis of email can be as simple as performing a pen-register analysis (who sent a message to whom and when) or as complex as analyzing and reconstructing message chains from numerous mail files.

Three primary email products exist in the corporate setting: Outlook, Outlook Express, and Lotus Notes. Web-based mail clients are also encountered frequently (for example, Hotmail, Yahoo! Mail, Gmail), but these do not have a client-side component to analyze unless POP3 or IMAP has been enabled, and other clients on Windows platforms are either less prevalent than before (for example, the excellent Eudora email client) or just burgeoning (for example, the equally excellent Thunderbird from Mozilla).

XI. Conclusion

Effective collection and analysis of digital evidence is a tedious task, which needs continuous analysis of data because Reliability, Security and Formality of collecting data as evidence is important legal basis or social consensus that recognizes legitimacy. In recent years the investigator had to use many forensics tools to perform investigation task. Integration of all types of forensics tools is a major challenge. The computer profiling models present a structure for development of automated computer forensic investigation. This collected evidence needs to be correlated so that, the correlation of events decides the success of forensic study and crime investigation in windows operating system environment.

X. Acknowledgment

The research presented in this paper would not have been possible without our college, at VITS, Jabalpur. The Authors wish to express our gratitude to all the people who helped turn the World-Wide Web into the useful and popular distributed hypertext it is. We also wish to thank the anonymous reviewers for their valuable suggestions.

REFERENCES:

[1] Cyber Crimes: A New Challenge, Deputy Controller (Technology), CCA, Ministry of Information Technology, India, 2002.

[2] Danquah, P., & Longe, O. B. (2011). An Empirical Test of the Space Transition Theory of Cyber Criminality: The Case of Ghana and Beyond. African Journal of Computing & ICTs. 4(2), 37-48.

[3] Bossler, A. M., & Holt, T. J. (2010). The Effect of self- control on victimization in the cyber world. Journal of Criminal Justice, 38, 227-236.

[4] Berg, S. E. (2009). Identity theft causes, correlates, and factors: A content analysis. In F. Schmalleger & M. Pittaro (Eds.), Crimes of the Internet (pp., 225-250). Upper Saddle River, NJ: Pearson Education, Inc.

[5] http://en.wikipedia.org

[6] Anconelli M., “Introduzione al digital profiling,” www.cybercrimes.it, 2010

* * * * *