Measures of retaining digital evidence to prosecute computer-based cyber-crimes

Shiuh-Jeng Wang *

Department of Information Management, Central Police University, Taoyuan 333, Taiwan

Received 4 October 2005; accepted 25 March 2006
Available online 19 May 2006

Abstract

With the rapid growth of computer and network systems in recent years, there has also been a corresponding increase in cyber-crime. Cyber-crime takes many forms and has garnered much attention in the media, making information security a more urgent and important priority. In order to fight cyber-crime, criminal evidence must be gathered from these computer-based systems. This is quite different from the collection of conventional criminal evidence and can confuse investigators attempting to deal with the forensics of cyber-crime, highlighting the importance of computer forensics. In this paper, we offer solutions to guard against cyber-crime through the implementation of software toolkits for computer-based systems. In this way, those who engage in criminal acts in cyber-space can be more easily apprehended.

© 2006 Elsevier B.V. All rights reserved.

Keywords: Digital evidence; Investigation; Computer forensics; Cyber-crime

1. Introduction

The rapid development in computer and Internet technology has brought forth a tremendous increase in computer and Internet crime. The most commonly seen crimes involve hacking into computer systems and computer viruses. Now these terrorists also threaten and cause damage via the Internet. Conventional crimes, such as fraud and money laundering, have been creatively changed into new forms by the rising popularity of the Internet. If we choose to ignore cyber-crime, the attackers’ ambition and greed will be encouraged and we will face more serious criminal behaviors. What we must do is fight these cyber-crimes by tracking the instigators through the proper procedures. This is the best way to protect our networks and keep cyberspace secure.

However, network interaction is based on principles of reliability and trust between both users and administrators: users “believe” Internet networks will not reveal their secrets, while network administrators also “believe” users will obey the rules when using the services provided by the website. Criminal behavior can damage this agreement of trust between both parties. Not only can lack of trust damage the development of Internet trade, but it can also affect the network services provided by network technology. When users are unable to trust a communication network, websites involved in network communication, connection and trading are simply nominal. Hence, when seeking secure data and website maintenance, we must discuss how legal methods and procedures can be applied to investigate those who engage in cyber and network crime.

In the past, when prosecutors and police investigated a crime, the investigators who analyzed the evidence also presented that evidence to assist the judge in making the correct decision. Criminal investigation training courses always include some forensics in order to understand what prosecutors and judges require in regard to evidence. The focus is on the collection and preservation of effective evidence. In other words, at a computer-based crime scene, the highest attention must be given to specifying digital evidence. The major feature which distinguishes cyber-crime from conventional crime is that the evidence at the crime scene is represented in electronic form. This also makes it easier for the criminal to store, conceal, propagate and remove the information and why it is more difficult to seize the suspects who commit cyber-crime.
When considering cyber-crime cases, the main digital evidence can be classified into four categories: electromagnetic records (in particular for printout information in a computer-based system); express delivery invoices; bank balance statements; and other beneficial objects. In the literature [1], digital evidence refers to any digital data that can provide a significant link between the perpetrator and the victim. In short, digital evidence is data in computer storage that can be used to prove criminal behavior. The physical characteristics of digital evidence are laid out as follows [1]:

1. It is easily copied and modified, but not easily kept in its original state: an electromagnetic record is stored in a computer system in the binary form—0 or 1. The copied object is exactly the same as the original one, but it is also convenient to proceed with user modifications. As a result, it is difficult to retain digital evidence in its original status. Confirmation of the original digital source is, therefore, susceptible to doubt.

2. Its source and integrity is not easy to prove: it is very easy to produce an electromagnetic record, so it is also very easy for it to be copied or modified. This makes it very difficult to directly infer the relationship between the evidence obtained and the suspects. That is to say, it is almost impossible to achieve “individualization”, unlike the highly efficient methods of fingerprinting or deoxyribonucleic acid (DNA), used to authenticate evidence. Accordingly, it is very difficult to prove whether the evidence has been changed, based on the observation of easy modification of electromagnetic records.

3. The presentation of digital information cannot be well perceived by human senses. This is because the electronic record has been electromagnetically recorded and stored inside the computer system. It is therefore impossible to perceive its content without the help of a suitable toolkit.

Nevertheless, the major problem arising from cyber- and Internet criminal crime is how to obtain digital evidence. Up to now, digital evidence has played a subsidiary role; it has been used to strengthen the evidential facts of a crime committed by a suspect. The courts have been concerned with whether the digital evidence is identical to the original record, whether it has been purposefully modified or whether there are other factors which have changed the original content. Stated simply, the most critical factor is being able to satisfy points of evidence and probative force. Following this, it is possible for the evidence to be categorized into “Hearsay” or “Direct Evidence” according to the representation of content [4]. The difference between digital evidence and physical evidence is that each examining action would be treated as “Access” in the collection of digital evidence, as it is easy to change the content of digital evidence and so disqualify the criminal facts. The only way around this is to “Freeze” the staff related to computer-based devices at the original crime scene. Computer-based (or Internet environment) criminal investigation procedures and principles are, in fact, similar to those of traditional investigations. The one thing worth noticing is how to obtain the “evidence” from these new technological products, required by the court, to prove whether the suspect is guilty or not. Thus, in order to retain the effective clues from computer-based systems, in this paper, we have developed a method of using forensic toolkits to facilitate the collection of powerful digital evidence. In this way, crucial tracks left at the scene of the cyber-crime can provide persuasive testimony to convict the perpetrators.

The remainder of the paper is organized as follows. In Section 2, we address the crime investigation and briefly review the relevant regulations. Strategies in computer forensics are presented in Section 3. Utilization of the toolkit associated with the forensic procedure is revealed in Section 4. Discussion around the efficient application of forensic toolkits is set out in Section 5. Section 6 concludes the paper.

2. Criminal investigation and relevant regulations

Significant criminal investigation research is based on modern scientific principles, which implies that today’s criminal investigations can prove not only if a suspect is guilty, but also if a suspect is innocent. The objective of the investigation is to unveil crimes and collect evidence, in the hope of discovering the truth. All of the evidence must be collected, whether favorable or adverse, in order to offer justice to the suspect facing trial.

Basic cyber-crime investigative models and procedures are as follows: crime information collection and case study, tracking of criminal suspects, and collect/search/seize/authenticate computer crime evidence. Once wrongdoing has been confirmed and reported to the authorities, the case moves into the investigation stage. In the initial stages of cyber-criminal investigations, all of the information related to the case is collected; then according to this information, a decision is made to investigate the crime, at the same time understanding the occupation of the suspects, their motivation and methods used.

The solving of both conventional and Internet crimes will benefit from immediate investigation of the crime scene to obtain the main evidence for solving the case. The key features associated with providing valuable evidence are generally categorized in the following: provision of case clues, study of the modus operandi, determination of the relationship between the suspects and the victims, determination of the relationship between the suspects and the crime scene, proof of testimony,
interrogation of suspects and provision of investigation direction. These factors can provide the relationship between evidence at the scene of the crime, the victims and the suspects, and becomes the basic principle of four dimensions (four relationships) connection method [7] as shown in Fig. 1.

If digital evidence against Internet criminals is collected without focus, and not enough attention is paid to collection procedures and methods, it can increase the burden on human resources and allow criminals to go free. According to the literature [4], Internet criminal evidence must meet the following requirements to maintain its reliability:

1. It must be produced, maintained, and used in a normal environment.
2. It must be professionally authenticated. In other words, the report from the forensic experts is reliable.
3. It must meet the “best evidence rule.” This means that what is produced must be the best evidence available and not a substitute for the evidence offered.

In Taiwan criminal code, Act 155 specifies that “the testimony of provided evidence should be optimistically evaluated by the judge.” However, with the current rapid development in scientific technology, it is very difficult to ask the judge to make judgment based on his personal knowledge. In fact, to judge and commit a criminal to justice requires the assistance of a computer forensics expert to authenticate the collection of evidence and the operating process provided by the investigator. This is done under the precondition of protecting human rights and justice. On the other hand, whether a computer record is suitable for adaptation to the hearsay principle is also dependent on the forensic result being reliable. Thus, in the criminal investigation and evidence collection, as long as it does not violate the criminal code, Act 156: “The defendant’s confession must not be obtained under violence, threat, interest, fraud, illegal custody or other illegal methods, which are the truth and can be used as evidence.” The defendant’s confession cannot be used as the only evidence in a guilty verdict; there must be other evidence to support the final verdict and it must be investigated to determine if it matches the fact. It cannot be concluded that a defendant is guilty without confession or evidence, and a verdict cannot be reached solely because a defendant refuses to confess or remains silent. In the criminal code, Act 159, the statement of “testimony given by a witness outside the trial should not be used as evidence, other than by the prescribed person” is regulated, and the criminal code, Act 160 further states that “personal opinions or presumed testimony of any witness shall not be used as evidence.” Under these regulations, cyber-evidence must be carefully handled according to evidence disposal procedures. In principle, the evidence should first be subjected to a back-up process, and the comparison procedure followed to carry out individualization. The reports can be offered as supplementary evidence. At the moment, “Criminal Investigation Regulations for Police” in Taiwan has pointed out the terms that should be taken note of in the case of search evidence, computers held in custody and Internet crime cases associated with computer forensics. However, some of these regulations are very unrealistic. For example, one of the investigation regulations requests computer evidence to be printed; however, if the computer under investigation is not connected to a printer, a printer driver must be installed before this evidence can be collected. This goes against another point: “Do not install or copy any programs and documents to the uninvestigated computer” in this regulations. Moreover, once the evidence is asked to send to the computer crime investigation squad of police bureau, apparently, it would be difficult to do so in the transportation of all the evidence. The evidence may exist not only in one computer; there is a possibility that more than 10 or even 100 computers are involved, or the evidence may be connected to an Internet facility which includes a server, which might need to be turned off before transport. In cases such as this, evidence transit and maintenance becomes very difficult and it is easy to further damage or interrupt the operations of the victim companies involved. Thus, if law-enforcement officials can effectively make use of forensic toolkits and follow proper evidence collection procedures, relevant information could not only be obtained before the computer is shut down, but could also reduce number of evidence items requiring seizure.

3. Enabled strategies to computer forensics

Computer forensics involves the preservation, identification, extraction, documentation and interpretation of computer data [6,14,16], in which computer forensics is defined as “an art of science using sophisticated methods and procedures to preserve, identify, extract, document, examine, analyze and interpret digital evidence.” This methodology and basic principles are briefly stated as follows:

1. Acquire the evidence without altering or damaging the original.
2. Authenticate the recovered evidence as being the same as the originally seized data.
3. Analyze the data without modifying it.

In compliance with the above, the computer forensic operating strategies itemized below provide effective investigation procedures for cyber-crime cases.

1. Preserve the evidence: The most important thing after entering the scene is to preserve the evidence, which is especially important for digital evidence in an Internet crime case. Digital evidence is, by its very nature, fragile [11]. Digital evidence can be changed at any time by striking the keyboard or clicking the mouse. If the evidence is not handled properly, it could result in evidence damage, inaccessible data, inability to prove that the suspect committed the crime, or even a possibility of having no evidence. Therefore, the first step upon arrival at the scene of the cyber-crime is to properly control the scene and begin to record the time, carry out the investigation and collect all significant digital evidence. According to the chain of custody [6], the following key
features are inspection requirements for valuable data analysis:
1. Who collected it?
2. How and where?
3. Who took possession of it?
4. How was it stored and protected in storage?
5. Who took it out of storage and why?

Hereby, there is one thing that needs to be paid attention to: that the computer should not be immediately shut down when arriving the scene of a cyber-crime, as to do so, the program and data being held in the memory would inevitably be lost and could destroy the evidence [2]. For example, during investigation of a suspect proceeding with a denial of service attack, to shut down the power suddenly would result in a program disconnection from the Internet. Even if related programs are discovered on a suspect's hard disk, it would still reduce probative force and give the suspect a chance to deny committing the crime. Thus, before dealing with any information, photographs or a video of the scene must be taken. Copying evidence stored inside the computer requires a special tool and must be carried out bit by bit; this is to say, the data must be copied using a bit-stream-copy method, which provides copied information in exactly the same format as the original data. By using this bit-stream-copy method, the data obtained are given more probative force. The most important thing is that the evidence is not amended or changed during the copying process.

(2) Examine the evidence: After obtaining the evidence from the scene, the next step is to analyze it. Common computer documents, pictures and sounds can be examined by many different software programs. However, the biggest problem is deleted documents, sometimes having been deleted by the suspect, which could be the most important evidence of all. Thus, slack space in the hard drive must also be scanned; this is one of the main reasons for using the bit-stream-copy method. A software tool must be used here to do string searching and document rebuilding.

(3) Evidence analysis: After examination and analysis of the evidence, the relationship between the forensic results and the suspects must be investigated. Through evidence classification, comparison and individualization, examination of forensic results can be connected to a suspect's behavior; for example, if a suspect often visits a certain website or often dials a particular phone number.

(4) Presentation of evidence: Results must be clearly presented. When discussing the source of evidence and cause and relationship with suspects, all other possible explanations must be eliminated to prove the final conclusion, thus confirming the guilty or not guilty hypothesis. In court, any suspicion relating to the evidence or the cause could affect whether or not that evidence is accepted. Moreover, a significant representation could provide clues for investigators to crack other crime cases.

4. Utilization of computer forensic toolkits

Although some people will try to write their own computer forensics program, this is not always a good idea. A personally written program is very difficult to sell and convince people that it is effective and meets all requirements; it could also affect the probative force of the evidence gathered. Hence, in many cases, ready-to-wear computer forensics software is more stable and reliable and has been tested and proved. The relevant toolkit requirements are as follows:

(1) Disk drive backup software: In order to retain the original evidence, data backup is the first consideration when proceeding with the work of computer forensics. A reliable backup software tool must comply with the requirements of the National Institute of Standards and Technology (NIST) [10], listed below.
1. The tool shall duplicate a bit-stream or an image of an original disk or section, where this so-called image refers to saving the content and related storage information as a document.
2. The tool shall not alter the original disk, i.e., the program cannot make changes to the original evidence media.
3. The tool shall be able to verify the integrity of a disk image file.
4. The tool shall log I/O errors; i.e., this program must offer a resolution to fix I/O error messages.
5. The output of the recorded documentation shall be correct in the wake of software operation.

To solve an Internet crime case, computer forensics is used for the collection of evidence, which must be analyzed and presented to the court, to prove illegal activity. It is imperative, during any computer forensics procedure, to allow no tampering with the data, which could induce viruses or cause damage; only complete evidence is qualified to be entered into the justice procedure [3]. An intensive analysis consists of, first of all, carrying out the safe investigation of the computer forensics program in order to secure the integrity of evidence and the safety of the information. The best way of doing this is to use a hardware image tool. Choosing and using the right tool in computer forensics is critical. Hardware imaging is also a backup method, backing up not only a copy of the document, but also “Ambient Data,” where a great deal of important evidence has been discovered. An example of a Microsoft Windows (MS-Windows) system is as
follows: Ambient data always exists in the swap file of the MS-Windows system, which does not take up any file space or slack space. Thus, if a normal copy method has been used, ambient data will not be copied into the new storage media. This process is shown in Fig. 2.

In short, all the backup data must be the same as the original data, not only in content; the condition of the drive must also be the same. This not only proves that the two are identical, but that it can also be used to search for any important evidence that may have been deleted in the drive. According to the test records [5] of the forensic toolkits, nd in the Linux/Unix system, Safeback and SnapBack DatArrest (http://www.snapback.com/) under an MS-DOS/MS-Windows system (http://www.forensics-intl.com/) are the best products to precisely copy a hard drive. Currently, for most agencies, Safeback and nd are used to carry out the backup task [11]. There must be at least two copies of each backup, to avoid a suspect declaring the inability to make a living because of a computer being held in custody.

(2) Authentication software: In order to prove that evidence has not been changed during the investigation, programs like MD5 or SHA-1 are required. The hash code in cryptography, produced by MD5, is one of the hash methods used to examine whether the original drive and the copy are the same. MD5 adopts an implied mathematical calculation method to save the data to the drive or document. The hash value produced by this mathematical equation can select any one of the data and it is in low possibility to have the same hash code for different data input; to find two identical hash codes would require approximately $2^{64}$ times calculations. In addition, to recover through hash code operation requires approximately $2^{128}$ times calculation [13]. This means that we can use the hash code to compare the original data to its copy. An example of this hash operation, the results of the hash of MD5 for a ‘Central police university’ message, is shown in Table 1.

(3) Decryption tool: When data from documents, lists or a hard drive are obtained for evidence, they could possible have been encrypted by the suspects to cover the evidence. If access to a document requires a password, which cannot be obtained from the suspect through conventional investigation methods, the following methods can be used to deal with the problem:

1. Guess the password. The majority of computers use a one-way function to protect passwords. The most commonly used methods, brute-force and dictionary attack could work here. If the code still cannot be decrypted then trust a professional to decrypt, using a high-speed computer.

2. Make use of the system vent to obtain the manager’s authority. For example, one of MS-Windows XP vents, dealing with the ‘universal plug and play NOTIFY’ code, can produce buffer overflow; this vent allows an intruder to transfer the instruction of NOTIFY to a system with a vent, allowing the intruder to perform any program he wants and have complete authority to run any system with any event [9]. Also, this type of vent allows an investigator to obtain system control authority when facing a suspect’s encrypted system, thus having a chance to solve the crime.

In addition, the most commonly seen encryption/decryption toolkits and methods, nowadays, are organized as follows:

A. BIOS universal password:

A.1 Use the BIOS password: the BIOS universal password is the backdoor to the top BIOS programs and is convenient for engineering manufacturing staff to use; therefore, no matter what password you have set up, the BIOS password can enter BIOS and reset the program. BIOS passwords, set up by the manufacturers, are different at different stages of manufacture. The most commonly seen BIOS passwords are shown in Table 2.

A.2 Clear all the CMOS setups: Usually, there will be a Jump on the top of the main board, near the battery, which could be used to clear the CMOS setup (usually known as Clear CMOS); this method will usually clear all the CMOS set ups, including passwords. If Jump cannot be found, the battery can be temporarily removed and put back after a period of time.

B. Decryption program for screensaver password: At the scene, the computer may have entered screensaver mode, with a password in

<table>
<thead>
<tr>
<th>Table 2</th>
<th>BIOS: commonly used passwords</th>
</tr>
</thead>
<tbody>
<tr>
<td>BIOS maker</td>
<td>Commonly used</td>
</tr>
<tr>
<td>AWARD</td>
<td>AWARD?SW, J62, HLT, Spyx, 589589, AWARD?SW, J262, J256, BIOSTAR, wod3, amphire, aLLy, biostar, ejnukl, lkwPETER, lkwPETER, SHER, SKY...FOX</td>
</tr>
<tr>
<td>AMI</td>
<td>AMI, BIOS, PASSWORD, HEWITT RAND, AMI?SW</td>
</tr>
<tr>
<td>PHOENIX</td>
<td>Phoenix</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 3</th>
<th>CD content to decrypt screensaver password</th>
</tr>
</thead>
<tbody>
<tr>
<td>Name</td>
<td>Usage</td>
</tr>
<tr>
<td>Autorun. inf</td>
<td>CD automatically play Batch file</td>
</tr>
<tr>
<td>Sssp</td>
<td>Decrypting program for screensaver password</td>
</tr>
</tbody>
</table>

Table 1
An example of MD5 hash for varying “Central police university” of message

<table>
<thead>
<tr>
<th>Message</th>
<th>Hash of MD5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Central police university</td>
<td>cd3e0617731f7e81a250220e00676d44</td>
</tr>
<tr>
<td>Central polices university</td>
<td>8096b72365ae9e6d0051902ef008ad3</td>
</tr>
<tr>
<td>Center police university</td>
<td>9a7a6922657c1e3d88a9d175d9ee665b</td>
</tr>
</tbody>
</table>
place on the screensaver program. In this case, the computer must be restarted; however, this will stop the programs processing and may cause other unpredictable conditions to occur during the restart process. Here, the vent of the CD auto-run in Microsoft Win95/98/Me and sshp (http://www.nt.unets.ru:8101/john/) can be used to solve this problem. We can burn the following document onto a CD, as shown in Table 3. Then, the burned CD is placed into the CD-ROM and the screensaver password can be decrypted.

C. Examine the notation of ‘*’ password program: In MS-Windows, there are many password setup and display methods using star ‘*’ to represent characters, the purpose of which is to provide protection. However, this can be decrypted, using free software like CopiX (http://toget.pchome.com.tw).

D. Decrypt the Microsoft Office document: with the popularity of Microsoft Office, Word, Excel and Access are used to create many computer documents. Documents produced by this set of software can be protected by the ‘read only’ or ‘password for access’ functions. However, Passware Kit (http://www.lostpassword.com) can be used to decrypt these safeguards.

E. Decrypt the Acrobat PDF password: Adobe, the portable data format (PDF) is the most widely used current electronic file production open standard. The common document format of this PDF can save the document in any font, format, color or diagram, regardless of the software or desktop used to create it. It also may have a password to prevent access or printing by others. In this case the Advanced PDF Password Recovery Pro (http://www.elcomsoft.com) can be used to carry out decryption.

F. Decrypt the compress file password: To save document space, many users will compress the files into one single document and may also add a password to protect the compressed files. Common compress programs are WINZIP, ARJ, and RAR. Here we use WINZIP and Advanced Zip Password Recovery (http://www.elcomsoft.com) as an example of password decryption.

G. Decryption program for Internet neighbor password: Sometimes, in order to have access to other computer resources, directories will be set up as share resources, with a password to prevent unauthorized access. However, MS-Windows 95/98/Me is not highly secure; thus, NetBrute Scanner (http://www.rawlogic.com) can be used to detect which directory has been set up for network sharing. Once its IP address and the title of the share directory are known, pqwak can be used to proceed with decryption, revealing the password.

H. Decrypt the MS-Windows NT/2000 password: Normally, MS-Windows NT/2000 requires a password before entering the system, with the Administrator having the highest authority; if there is no password, no one can enter the system. However, we can use a floppy disk to start the computer, and then use the Passware Kit (http://www.lostpassword.com) tool to change the password.

(4) Disk data editing and searching tool: when data and documents have been saved on the hard drive, even after it has been deleted, the actual data may still exist. An identity code is added to the document title to represent the fact that it has been deleted. Thus, if a deleted file can be found, it may be possible to discover important evidence. If all data in the storage is done using the manner of the data recovery, this will take a long time and the evidence discovered is probably not satisfied with the realistic needs. In the criminal code, if remarkable evidence cannot be discovered within 24 h, it will be very hard to continue with the investigation. Hence, software like WINHEX EDITOR (http://www.winhex.com) can be used on the drive analysis and string search, to determine if there are any data worth recovering before running the recovery program.

(5) Daily audit log file: Normally, a large computer host or intrusion detection system (IDS) has a daily audit file to record important activity, including the times of user login and logout, website connection and document access and even a record of users’ commands. By investigating these records, the initial source of the attacker can be identified [12]. When identifying suspects, information relating to the account, passwords, and program used, should be taken note of, as these items will be help track down the attacker. As a matter of fact, information provided by an audit file may not be entirely correct; we should take other ways to find out more complete data sources to support the criminal facts. Records of the suspect’s website visits can be used, however, to prove whether the suspect was involved in the crime. Using MS-Windows as an example, Cookies analysis can uncover the websites or accounts that have been accessed by the suspect.

(6) Track the attack point address: the IP address is required when investigating a cyber-criminal. Implementation problems are listed below:

1. The IP correspondence address can be obtained by using the PING command to discover the domain name. Then, using the WHOIS and traceroute commands, other relevant information can be gathered.

2. Using ICQ: ICQ can help track those who use Internet chat rooms to carry out illegal business by using the ICQ/IP configuration function to discover the IP address.
3. Trace by e-mail: the original e-mail content can provide the IP address of the sender or server of mail forwarded. Other than this, a network packet could be intercepted by the operation of TCPDump (http://www.tcpdump.org/) in the Linux system or Windump (http://windump.polito.it/) in MS-Windows. This observation report could be offered for further analysis.

4. Data recovery: Many criminals will delete documents which could be used as criminal evidence. Thus, a data recovery tool, such as Final Data (http://www.finaldata.com), can be employed. Documents, however, are not the only target; therefore, R-Mail (http://www.r-tt.com) can be used to recover deleted e-mails.

5. Document examination: Due to the complexity of document formats, and with most criminals changing subcontract names to reach their objectives, it is no easy task to uncover evidence. Thus, Quick View Plus (http://www.stellent.com) is a tool that can be used to search and eliminate these efforts at camouflage.

5. Discussions

In the new territory of computer forensics, many problems still exist. The following is an analysis and discussion of the blind spots and cautions to be heeded when investigating Internet crime.

1. The Law: Where does network intrusion or attack take place? Who are the intruders? Where does it happen? There will be problems when applying for search and investigation warrants. Relevant information will be saved in storage media or networks in an electromagnetic format; it is impossible to confirm whether this storage media contains criminal evidence solely through the human senses, so the basis of the investigation and custody in Internet crime can lead to many disputes, such as:
1. Whether justice officials are entitled to enter another’s network domain to search for and collect evidence through an Internet connection.
2. When criminal Internet access could destroy or endanger public security, a Trojan virus can be used to track down its progress, enabling manpower and resource savings. However, the credibility of the evidence gained through this Trojan virus, implanted into a criminal suspect’s computer and used as to track his/her movements, must be determined.
3. Under what circumstances is the seizing of criminally related computer hardware reasonable? In order to protect human rights, this must be taken into account. Laws should be unambiguous and explicit to enable the checking and searching of cyber-crime suspects, while at the same time protecting human rights.

2. Identity: Identifying the real attacker is very difficult, especially when anonymous accounts or an embezzled identity have been used. Moreover, the identification of a suspect within a specific time period could be deadlocked, as computer system passwords are usually made public. Even though, when investigating Internet invasion, the MAC (Media Access Control) address can be used to identify an individual’s network card number, in the case of the Linux system, the MAC can be altered, making it useless for evidence identification [8].

3. Encryption: Most programs come with the encryption methods and most people can gain access to security applications through the Internet, especially with hardware devices, where activated keys in the storage media are embedded in chip-structures, such as a smart card or SIM card. To carry out cryptanalysis is more difficult than ever. Furthermore, if a steganography technique is applied to hide data in multimedia, it will present a new challenge in identifying evidence based on imperceptibility of human visual system.

4. Storage capacity: Presently storage capacity increases dramatically year by year. Most users often have more than 40 GB hard disk capacity. The larger the capacity, the longer it takes to make copies of evidence and the more storage media required. Consequently, more time will be required for analysis.

5. Audit records: System audit recordings are dependent on system users following proper procedures. System users may often delete audit records with the purpose of restoring the computer to avoid an invasive attack, resulting in computer recognition problems. Thus, safety management must depend on all security and management techniques being followed in order to establish a complete audit record; this is a mission to be worked on in the future.

6. Complexity of system: Common computer systems are mainly DOS, Microsoft Windows, Linux/UNIX and MAC OS (especially for Macintosh computers). Every system comes in different varieties. This results in more complexity in computer recognition within different computer systems, computer programs, hard disk equipment, etc. If the computer’s hard disk and software are too obscure, it can be difficult to collect and analyze evidence.

In [15,17], recently, which announced the cracking of the hash, either MD5 or SHA-1, where the security complexity of MD5 is degraded from $2^{64}$ to $2^{37}$ and the security of SHA-1 is downgraded from the probability of $2^{-80}$ to $2^{-69}$. It means that they are not foolproof. So far, there are a number of forensic toolkits (or audit toolkit), such as Encase, TCT (The Coroner’s Toolkit), Sleskkit and Tripware, where the verification procedures are carried out via the hash of MD5 or SHA-1 algorithms. The hash usage in these toolkits is to compare the digest of the hash output of the data content of the file or message. It is almost impossible to trace back the original data content through the digest, even though the collision probability has been raised by the hash cracking. In practice, the salts of hash, such as HMAC-MD5 and HMAC-SHA1, are more difficult than the general hash if the cryptanalyst wishes to unveil only the original content from the digest, since the secret key is incorporated into the implementation procedure. In this way, the ability of offering evidence and proof of evidence is still safe to be used in the investigation of cyber-crime cases. But it is worth noting that the hash-based evidence should be questioned if the
current hash algorithms are not changed to a more secure hash algorithm in the near future. That is to say, it is necessary to apply a higher level of security, and a more secure evaluation in the hash use of forensic toolkits. Fortunately, except for MD5 and SHA-1, there are some substitutes proposed. In 2002, for example, NIST released a series of new algorithms including SHA-224, SHA-256, SHA-384, and SHA-512. They are more difficult to break than SHA-1 and have been considered for the candidate list.

6. Concluding remarks

In this paper, we discussed cyber-crime committed through computer or network systems. Although investigating cybercrime requires investigation of digital evidence, the basic principles, comprising four dimensions, are the same as those used in conventional crime scene investigations. Because of these four dimensions, in order to authenticate the latest digital evidence, we have also quoted the requirements of Acts 155, 159 and 160 of the criminal code in Taiwan, which extend to digital evidence left by cyber-criminals in a computer-based system. Significant progress in this field will provide important references for the investigation and solving of crimes. Furthermore, with the increasing popularity of computers and the Internet, information intrusion and illegal cyber-crime have already become new channels for those intent on breaking the law. In this regards, we have suggested some forensic toolkits which can assist in the validation of digital evidence. Finally, we offer various measures to help in the potential roadblocks which will inevitably arise in the future, in the fight against cyber-crime. In fact, the potential problems we mention here are only offered as a direction for discussion. Cyber-crime will continue to develop as criminals become more creative in the use of systems, different computer environments and changing criminal objectives. Computer forensics and recovery, including the collection of digital evidence, will also continue to evolve in response to these changes. The exploratory research contained in this paper is offered as a reference tool for the protection of computer-based systems from criminal hackers.

Acknowledgement

I am indebted to the anonymous referees for their valuable comments that greatly improved the presentation of this paper. I also thank my student H.J. Ke who is a member of ICCL (Information Cryptology & Construction Lab.) for his fruitful discussions in the course of the preparation of this manuscript.

This work was supported in part by National Science Council in ROC under Grant No. NSC-94-2213-E-015-001.

References


Shih-Jeng Wang was born in Taiwan in 1967. He received his M.S. degree in Applied Mathematics from National Chung-Hsing University, Taichung, Taiwan, in 1991. He received his PhD degree in Electrical Engineering at National Taiwan University, Taipei, Taiwan in 1996. He is currently with the Department of Information Management at Central Police University, Taoyuan, Taiwan, where he directs the Information Cryptology and Construction Laboratory (ICCL, http://hera.im.cpu.edu.tw). He was a recipient of the 5th Acer Long-Tung Master Thesis Award and the 10th Acer Long-Tung Ph.D Dissertation Award in 1991 and 1996, respectively. Dr. Wang was a visiting scholar of Computer Science Department at Florida State University (FSU), USA, in 2002 and 2004. He also was a visiting scholar of Department of Computer and Information Science and Engineering at University of Florida (UF) from Aug. 2004 to Feb. 2005. He serves as the Director of Chinese Cryptology and Information Security Association (CCISA). He is the editor-in-chief of the Communications of the CCISA in Taiwan. He is the author of three books (in Chinese versions): Information Security, Cryptography and Network Security, and State of the Art on Internet Security and Digital Forensics, published in 2003, 2004, and 2006, respectively. He is a member of the IEEE, ACM, and IACR. His current interests include cryptography, digital investigation and computer forensics, steganography, data construction and engineering.