Research on the Application of E-data Evidence-taking Technology

Dawei Xiang [1] , Yanbo Wu [2,*]

[1] Department of Electronics Engineering，Hubei University of Police,P.R. China

[2] Hubei Police College Experiment Center, Hubei University of Police, P.R. China

[*]Corresponding author Email: 240325743@qq.com

Abstract:E-data evidence-taking technology can provide a powerful legal basis for cracking down on illegal activities. As high attention is aroused from the units of national public security, procuratorate and court, more investment has been put into the researches on legislation and technology, and into evidence-taking practice, pushing the development of e-data evidence-taking sector. Based on the prior research results of experts and scholars, and in light of the author's personal experience in e-data judicial testimony, the thesis has made an in-depth study on the e-data evidence-taking approach and the related technology.

Key words: digital evidence-taking, evidence-taking approach, key technology

With the further development of computer and internet technologies, people are becoming increasingly inseparable from the e-equipment in their life and work. At the same time, we can also see the growing number of illegal activities by means of the internet, computer and other mobile terminals. To crack down on such crimes, public security department has carried out a special research with e-data as the core to analyze and evidence the e-data in computer and the related equipment. Therefore, it is particularly important to do well in e-data analysis, extraction and maintenance.

1 The research on the basic approach of computer evidence-taking.

The basic approaches for e-data include acquisition approach, analysis approach, security approach and the submitted report[1].

1.1 Acquisition approach

Preparations for acquisition: Formally before e-data evidence taking, package and seal the original test materials, take records of their appearances, take photos and videos. When it is necessary to open or close the package of the test materials, it takes at least more than two workers to work on it, and it is best to have both sides of the test materials submitter and recipient on the site. The process of test requires a sound record as well.

The cloning of e-data: After acquiring the original evidence carrier, such as the main engine hard disc, it is not allowed to make direct analysis on the data in principle. First, it is necessary to clone the data bit by bit; then close up the original evidence carrier and make sound maintenance; next, digital evidence-taking personnel can begin the analysis on the copy of the cloned original evidence. Cloning refers to copying the original data bit by bit. Not only is the data-contained documents copied in this way, and also the relative positions of the documents, the data state and the deleted information will also be duplicated into the copy. The cloned data is exactly the same as the original data, even the hash calculation result of the whole original evidence and the copy.

The acquisition of the disc array data: Disc array data is different from the common disc data in terms of reading mechanism and storage. Based on different array types (raid 1 and raid 5), disc array generally has different amounts of discs to form a reading & writing system. Taking the commonly used raid 5 array for example, (at least three physical discs), a piece of data is likely to be stored in different hard discs by different segments. In case of the data with single disc, the data would be incomplete and hardly recognized by computers in correct way, unless all the data are acquired for all the arrays. There are generally two practices: respectively clone each of the hard discs and reorganize

the disc array; by special disc array evidence-taking equipment, clone the multiple array data into single disc data.

The acquisition of real-time data: To acquire the network data flow or the real-time data in the internal storage memory, some kinds of professional software are needed, such as ENCASE. They can generate network data flow and the image files of the real-time data, and have them stored in physical storage equipment.

Data restoration: Before taking evidence, it is possible to restore the data in the cloned copy, because the important evidence is likely to be deliberately or unconsciously deleted by suspects, and need to be acquired after having the data restored.

1.2 Analysis method

There is no fixed mode or approach for analyzing data evidence, and it depends on the type of the cases and the acquired evidence and needs to use different analysis approaches and steps. For example, in the cases of cybersex and network fraud, it shall focus on analyzing website codes, back-end data base, system logs and office e-documents; while for the case of computer virus dissemination, it requires analyzing the data if the computer virus source code, virus finished log (executable files, editing environment, system operating log and internet browsing record; for the cases of network attack and damage to computer information system, it shall focus on Trojan virus files, internet browsing records, security logs, applied logs, fire wall logs, registration sheet, system accounts, open terminal and application service[2].

With the evidence analysis for Trojan virus cases as example, the following steps can be implemented:

1) Plug out the attacked computer line and power (not the best option sometimes in the course of attack, for example).
2) Clone the original evidence carrier, keep the copy under good care and use the it for analysis.
3) Use data restoration tool to restore all the deleted data.
4) In case of known Trojan program, directly use key words to search for the files related to Trojan with the help of the special software. Because the search is done in the whole domain, it shall include the hidden files, system files, encrypting files, files protected by passwords, temporary files, and the deleted files possible for restoration. If it is unclear about what kind of Trojan it is, get clear the types and major characteristics of the virus, and analyze the relation between the data above to confirm if it can reflect the crime facts.
5) Check the system accounts, registration sheet, website codes, website browsing records and various logs to confirm if they can reflect crime facts.
6) Monitor the service and the opened port to find their directional relation with the applied program.
7) Comprehensively analyze the above evidence to sort out the thread and get the analysis conclusion.

1.3 Fixed evidence and submitted report

The fixation of evidence refers to sorting out and storing the key evidence after the case is basically clear, the crime evidence is sound, all the data evidence has been extracted and analyzed. After fixation, the data evidence will serve as the sworn evidence for court.

2 Research on e-data evidence-taking technology

There can be different categories about the key data evidence-taking technologies based on different research angles. From the angle of operation process, it can be divided into data-obtaining technology, data analysis technology and other technologies.

2.1 Technologies for main engine data evidence maintenance, restoration and analysis

1）Site e-evidence analysis and record technology

It mainly studies the related evidence-taking technologies after computer crimes, such as how to handle the target computer system, how to avoid any change, hurt, data destroy or virus infection, and how to maintain correct extraction of e-data from the storage equipment. It includes field survey, hard

disc check and clone, special data acquisition, file property analysis, file data extract analysis, log analysis, encoding and decoding, password acquisition, information searching and filter, reverse engineering, e-data appraisal and standard system.

2）Storage equipment read-only interface technology

It mainly aims to research the connection approaches and tools for multiple different interfaces, passage technology, data read-only realization, etc.

3）Computer hard disc high-speed clone and proofing technology

It mainly aims to research the related agreement to reading & writing hard disc data, high-speed interface, data fault-tolerance, CRC-32 signature proofing.

4）Data restoration technology

It mainly researches computer data restoration technology, such as how to restore the special files or data modules completely or as many as possible to make them effectively applicable for data evidence. It includes the restoration of the residual data (slack space, DUMP internal storage and SWAP file）, common file system（FAT，NTFS, Ext2 and Ext3）and magnetic force microscope (MFM for short), etc.

5）Document fragmentation interpretation technology

It mainly researches how to defer the possible format of the fragments file control modules (record modules)； based on the wording, grammar and writing style of the acquired files or data, defer the author’s possible habit characteristics and computer level.

6）Hidden file recognition and extraction technology

It mainly researches file camouflage, data hiding, data watermark and decompilation, etc.

7）Network application index technology

Analyze Slack disc space, the space of the undistributed disc and the information contained in the free space, research the swapping files, cache files, temporary files and network flowing data, so as to discover the evidence for E-mail and Internet browsing, file uploading and downloading that have ever occurred in the system； extract the contents related to key words (such as terrorist attack, extreme speeches and sex crimes).

8）Index analysis technology for website code, procedure code and data base record

9）Credible hash proofing algorithm (such as CRC and MD5 proofing, etc）.

10）Evidence-taking analysis by the external equipment.

2. 2 Capturing and analysis of the network data flow

The research scope mainly covers how to capture the network information dataflow by real time, and how to locate the important information such as the attack source by data capture and analysis[3].

1) Research on the service agreement for file transmission, e-mail, online short message, BBS, P2P. The research over e-mail service protocol mainly focuses on the acquisition and analysis of the important information in the process of e-mail transmission, such as intercepting the user name and password of the e-mail, the e-mail and the attachment sent or received for restoration decoding and the analysis on the e-mail head, etc. The research over the network short-message service in mainly concentrated on analyzing CMPP and SMPP, extracting and displaying the short messages transmitted by use of the short messages. The research over BBS is built on the basis of BBS agreement analysis to analyze the data transmitted between the client and the server to restore the users’ IP.

2) Capturing and analysis on wireless network IP

The most used WLAN agreement recently is IEEE802.11 based on the WLAN of IEEE802.11 which uses confliction-avoiding carrier to monitor multiple visiting agreements (CSMA/CA) to visit medium. There is a promiscuous mode in wireless network card. Besides receiving data package, the wireless network card in promiscuous mode can also send data package at the same time. Most of the network cards have another RF monitor mode besides the normal operating mode and promiscuous mode. The mode only allows the wireless network cards to receive data, instead of sending data. When the wireless network cards are in RF monitor mode, it can capture all the data package in the

basic service set in the air. Therefore, it can realize data package capturing and analysis based on the physical structure of WLAN.

2.3 Positive evidence-taking technology

1）Network attack technology

Such as SQL infusion, boundary access (buffer zone overflowing), automatic footnote execution, keyboard hook, and the scanning of weak configuration and password.

2）Honeyport technology

Network attach and trick system can monitor every move of the suspects in the system, and even the record of every key stroke; the hidden monitor technology can make it undiscovered by the intruders or suspects; make sure the logs of the main engine trick not be revised or deleted, and the intruders or suspects cannot pose threats to other main engines by use of the trick main engine after they get access to the system. Therefore, it needs to take into consideration such key technologies as seduction, camouflage, records and evidence, analysis and assessment, response warning and automatic protection to realize network attack and trick system.

3）Dynamic monitoring（IDS）and log maintenance technology [4]。

In light of the dynamic nature and possibility for loss of the network intrusion status, it is necessary design intrusion evidence-taking system, aiming to completely record all the flow rates on the network. At the same time, transfer and protect the main engine logs by real time, and encode all the records, so that it is possible to keep all the original, complete and unchangeable event records on the network.

3 The application of data restoration technology

In daily life, there can be many reasons for data loss, such as vicious delete by virus, improper operation and aging equipment. In judicial field, illegal criminals will not touch the files about important cases which are stored in electronic devices, such as computer and mobile phone and wait for us to test personnel to take evidence. They may choose to delete files, format the disc or even reinstall operating system to destroy evidences or clues. Among, file deletion under Windows operating system is the most common practice the case detectors could meet. Therefore, the test personnel has to master the related storage structure and delete principle of FAT and NTFS files to restore as many as data files with the help of related software, so that more valuable evidence or information can be provided for the case detectors.

First, judge the category of the file system. WinHex software can be used to open the test U disc to see that the empty main partition sheet leads the operating system unable to recognize the disc. Based on the file system structure, we can know the starting sector of the first partition is generally 63 and skips to #63 sector. The relative shifting value at 0x03~0x06 is “4E 54 46 53”，that’s ASCII “NTFS”. Therefore, we can have such a basic judge that the U disc uses NTFS file system.

Second, acquire boot sector information. It can be seen that #63 sector is the boot sector in this partition. The two bytes “3F 00” of the relative shifting 0x18~0x19 mean the number of sectors on each magnetic track is 3FH（63D）; The two bytes “FF 00” of the relative shifting 0x1A~0x1B mean the number of recording heads is（255D）; The four bytes “20 33 3C 00” of the relative shifting 0x28~0x2B mean the total number of the sectors in NTFS partition is 3C3320H（3945248D）, of which the capacity can be obtained by the following formula:

$$3945248 \times 512 \div 1073741824 = 1.88\text{GB}$$

It can be judged from the capacity of the partitions that the disc only has one partition, or it can be judged by checking the total number of the sectors of the driving disc if there is an expanded partition.

Due to the 225 magnetic heads in the boot sector and 63 sectors on each magnetic track, LBA adddr3ess of the starting #63 sector converts to CHA address with #00 cylinder and #01 magnetic head and #01 sector. So we can restructure the main partition sheet of U disc (as shown in table 1)

Note that in FAT12/16 file system, the total number of sectors is stored at the relative shifting 0x13 ~ 0x14 of the boot sector; in FAT32 file system, the total number of sectors is stored at the relative

shifting 0x20~0x23 of the boot sector. If there is an extended partition in the driving disc, it needs to continue analyzing the boot sector of the extended partition, and acquire such data as the starting sector, file system and the total number of the sectors.

Table 1 Data restructure table

Shifting	Length	Meaning	Value
00H	1 Byte	Partition guide sign and the real case is in the unguided non-activity partition indicated by 0x00	00
01H~03H	3 Byte	Partition starts from CHS address，while the case starts from #1 CHS recording head, #1 section, #0 cylinder which is already out of use.	01 01 00
04H	1 Byte	Partition type with the case as NTFS indicated by 0x07	07
05H~07H	3 Byte	Partition ends at CHS which shows as an arbitrary value out of use	
08H~0BH	4 Byte	Partition starts from LBA, and the case starts from #63 sector of LBA which can convert to hexadecimal 0x3F	3F 00 00 00
0CH~0FH	4 Byte	The total number of partition sector and the case sector is 3945248， and the total number in the partition table shall add 1 and convert to hexadecimal 0x3C3321.	21 33 3C 00

Third, restructure the data in partition table. Use WinHex software to change the data in table 1. After storage, reload test U disc which can be recognized by operating system, and the files in storage can also automatically restore.

Either common delete or complete delete, either FAT file system or NTFS file system or HFS+ file system, it is possible for restoration as long as there is such a predicate that the real data part remains there.

The formatted restoration is more difficult and complex than delete restoration. The effect of the formatted restoration will be under the influence of multiple factors such as file system, operating system and formatted parameters.

Acknowledgement

It is a major research project of Hubei Police Institute （2014GZ022，2012AT006，2015ZD009），and an approved topic at the provincial humanities social sciences base of the institute（2016-3）.

References

[1] Wu Yanbo， Xiang Dawei, The Application of E-data Technology in Falun Case (a kind of anti-human superstitious doctrine) [J] .Police Technology，04/2015

[2] Xiang Dawei, Wu Yanbo, Evidence Extraction from the Connection Sign of Windows 7 Mobile Storage Equipment [J] .Information Security Research，03/2016

[3] Mai Yonghao, Sun Guozi, Xu Rongsheng, Computer Evidence-taking and Judical Appraisal (Edition II, Beijing: Tsinghua Press,03/2014)

[4] Yang Jia, Mai Yonghao, E-data log Evidence Analysis and Case[J], ComputerScience, 10/2014