Research on the Way of Obtaining Electronic Physical Evidence Based on Python—Take Windows Log Document as an Example

Shancheng Lin*, Qianhao Chen

Collage for Investigation, People's Public Security University of China, Beijing, China

*Corresponding author, e-mail: lsc18783978379@gmail.com

DOI: 10.37420/j.mlr.2020.007

Abstract: *With the popularization of computer technology such as smart phone, cloud computing and so on, high-tech crime is becoming more and more common. However, the number of Internet users in China is large, the number of cybercrime cases is large, the electronic data that need to be collected is many and complex, and the research of electronic data forensics in China is mainly based on technical framework and model establishment, and the hardware development is mainly introduced and cooperated. So that China's electronic data forensics technology cannot meet the needs. Electronic data evidence and related legal issues, network forensics, mobile intelligent terminal forensics analysis technology, malicious code forensics analysis, cross-platform forensics, data recovery, intelligent association analysis, data depth mining, password cracking and other technical aspects need to be continuously combined.*

Accordingly, this project aims to carry on the analysis and the application research to the electronic data forensics technology of the network crime case, take the Windows log file as an example, through the reading and analysis of the user system log, we can understand the details of the user computer operating system, the behavior of the application program, the behavior of the user itself and the abnormal events in the system. Then it rebuilds the computer operation scene, monitors the computer system resources, audits the user's related behavior, carries on the alarm to the suspicious behavior, looks up and determines the scope of the intrusion behavior and so on, provides the evidence source and the key clue for the fight against the computer network crime.

Key words: *Electronic data; Computer; Electronic physical evidence; Cybercrime; Windows logs*

Introduction

Purpose and Background of the Study

In recent years, with the development of national economy and the rapid development of computer network, the wide application of computer and network technology is changing people's way of life. At the same time, we also suffer the crime of computer network crime which is bred by the development of Internet.

With the arrival of the 5 G era, more and more criminal activities against or using computer systems and interconnection, the malicious attacks of network hackers are increasing, and the ways of attacks are changing with each passing day. The resulting losses are shocking and make cybercrime a new type of criminal activity.

Cybercrime refers to the behavior of the subject who intentionally endangers the security of the computer network and violates the relevant laws and regulations by taking the computer or computer network as the criminal tool or the object of attack. Cybercrime can be divided into six categories: interference with the legitimate use of computers, such as DOS attacks, viruses, worms, other malware, cyber-destruction, cyberterrorism, spam; information theft and copyright infringement, such as industrial espionage, identity theft, identity fraud; the spread of contraband or offensive material, such as pornographic material, cyber gambling, treason or racist material; and communications threats, such as extortion, cyber-tracking, cyber-harassment, cyberbullying; Fraud, such as auction fraud, credit card fraud, theft services, stock price manipulation, and incidental crimes, such as money laundering, conspiracy, etc., attempt to gain benefits from it and escape criminal behavior.

With the development of computer and network technology, the dependence of network crime on computer technology is increasing. Because of the high degree of association and the large scope of influence, it has become the main means for criminals to use high-tech crime. Cybercrime has the advantages of low cost, rapid spread, wide spread, high interactivity and concealment, difficult to obtain evidence, and seriously endangers social security. Because the traditional forensics technology is not suitable for the investigation of network crime cases, it is urgent to combine the new forensics technology with it to play a greater value in the detection of network crime cases.

In order to prevent and crack down on computer and network crime, electronic forensics technology came into being.

Electronic data forensics technology according to scientific methods and means, with the help of special forensics tools, the discovery, recording, extraction, analysis of electronic data stored in various storage media, and in accordance with relevant provisions to the court to prove the facts of the crime. Electronic data collection technology has quickly become a powerful weapon for public security organs to solve network crime cases. The purpose of this project is to analyze and apply the electronic data forensics technology in cybercrime cases. This paper summarizes the difficulties existing in the present stage of electronic physical evidence, and finally puts forward the relevant solutions and suggestions.

Significance of the Study

The commonly used electronic data forensics techniques include Windows forensics, Mac OS forensics, digital cryptographic forensics, UNIX/Linux forensics, mobile terminal forensics, cryptographic forensics, Office document forensics, digital image forensics, database forensics, system environment simulation forensics and so on. Through the analysis and study of these electronic data forensics techniques, the application of the above technical methods to the detection of cybercrime cases can solve the problem of difficulty in obtaining evidence in cybercrime cases and improve the detection rate of cybercrime cases.

Maintain the security of cyberspace.

The fragility of electronic evidence determines its limitation as evidence. Only "source" data can prove the facts of the case, through legal approval. At present, most of the research on electronic data forensics technology is focused on the field of computer forensics.

Therefore, analyzing and studying the platform and method of establishing electronic data forensics in cybercrime cases, and solving the practical problems of less evidence sources and difficulty in obtaining evidence in cybercrime cases, has become an important direction of the development of electronic data forensics technology in cybercrime cases.

Research Methods

This group intends to adopt the literature inquiry method, the experiment research method, the induction method carries on the research. Through a large number of Chinese knowledge network, Wanfang data and library literature, to understand the relevant knowledge of Python application methods in Windows. Then use the Python to read the Windows log, through the way of reverse repair of the Windows log, scene reconstruction, induction and summary, and then find a feasible method of electronic physical evidence forensics.

Study on Electronic Exploitation for Explosive Expresses at the Period

Overview of Evidence Collection in Electronic Evidence

To study electronic physical evidence, we must first understand what electronic physical evidence is. It can be clearly defined in such legal documents as the Regulations on the Administration of Electronic Data Interchange for International Economic and Trade issued by Shanghai, the Regulations on Electronic Transactions issued by Guangdong Province, the Regulations on Electronic Data Interchange issued by Hainan Province and the Trial Measures on the Administration of Digital Certificates issued by the State Council. Electronic physical evidence is a variety of information materials that can prove the true situation of the case in various forms of data on various computer memory, such as soft, hard disk, memory bar, CD-ROM and so on. Evidence is mainly based on its relevant legal procedures, namely: First, whether in the whole process of obtaining evidence has the subject examination qualification, whether the illegal operation; The second is to ensure the effective integrity of physical evidence information;Third, extract and organize the effective part of physical evidence to ensure the accuracy of evidence results.

Current Situation of Research on Electronic Evidence Evidence in China

China's electronic data forensics technology has developed relatively late. In May 2000, the Ministry of Public Security formulated the "Key Ideas for Combating Computer Crime Technology ", which marked the entry of electronic data collection into China. The initial stage of the development of electronic data forensics is mainly the application of foreign forensics tools and the construction of relevant laws and regulations. In

2005, the Computer Evidence Expert Committee of the China Electronic Society and the Computer Evidence Technology Research Group were established in Beijing, which greatly promoted the development of electronic data forensics technology in China. By the end of 2016, the China Computer Forensic Technology Summit had successfully held 12 sessions. The conference covered legal issues related to electronic data evidence, network forensics, mobile intelligent terminal forensics analysis technology, malicious code forensics analysis, cross-platform forensics, data recovery, intelligent association analysis, data depth mining, password cracking and other technical aspects. In January 2018, the first high-end forum on electronic data foresight technology was held in Nansha, Guangzhou. The conference mainly discussed the legal puzzles and technical problems in the field of electronic data forensics, and opened a new era of electronic data forensics.

Academic, electronic data forensics, analysis technology and other related topics were included in the national 863 information security emergency plan. Some research achievements have been made in electromagnetic trace forensics, data recovery, intrusion decoy and memory forensics. For example, Guo Linlin elaborated the method based on Windows security log forensics. Tang Zhenhua put forward the Windows registry forensics analysis technology and formed the Windows registry forensics analysis system. The system is divided into three modules: evidence collection module, evidence analysis module and evidence representation module. However, the system is a static forensics analysis system after the event, which can not fully meet the current needs of dynamic forensics analysis. Deng Jincheng designed and implemented a fast search method for supporting TXT, PDF, Office03/07 series of files on the basis of predecessors. Yan Xi et al. put forward a computer dynamic forensics method based on honeypot technology, which transfers intrusion to a virtual environment through honeypot technology to obtain electronic evidence. The system has the characteristics of high detection rate, low false alarm rate and strong forensics ability. Guo Mu and Wang Lianhai proposed a Windows physical memory analysis method based on KPCR (Processor Control region) structure to obtain physical memory and analyze and obtain evidence, which extended the extension of memory forensics technology.

Current Situation of Foreign Electronic Evidence Evidence Collection Research

Because forensics comes along with computer crime, so it was first called computer forensics (Computer Forensic). The earliest computer forensics technology began in 1984, Beginning with the establishment of the Computer Analysis and Response Team by the FBI (FBI), The British Police then set up a computer forensics department. In 1991, the International Conference of Computer Experts formally proposed "Computer forensics ". From the early 1990s to the early 21st century, the research focuses on the development and utilization of forensics tools.

Windows Log Forensics

Windows Logs

When the computer system runs, it will produce a large number of log files. When the intruder enters the computer attack system, these log files can record the information and play the role of fixed evidence.
By Windows logging data, we can extract the information elements we need to obtain evidence i.e.: time,

type, user, computer, event ID, source, category, description, data, etc.

Windows Log Format and Storage Path

From the beginning of the Windows 2000, Windows system logs are stored binary. By default, their location is systemroot%\system32\config,% And 512 kb, in size Can not exceed. Type C: WindowsSystem32winevtLogs in Windows Explorer to enter the location of the windows log directly. "System", "Setup", "Application", "Security", can be found in them Corresponding Windows common system logs, installation logs, application logs and security logs, Among them, The security log doesn't open by default, Need system administrator to open. In addition to the above system log files, According to the configuration and application of the system, There may also be file directory service logs, DNS logs, PC custom log files, FTP logs, WWW logs, and Internet firewall (ICF) logs. These log files, Are protected by EventLog, can not be deleted.

Extraction of Windows Logs

For some log files stored in text, such as WWW, FTP log and planned task log, access is relatively simple, can be extracted directly. Other protected system logs or applications are stored binary, so they can only be obtained by the system or the software itself. For binary log files, we can operate and extract such log files through the corresponding functions provided by Microsoft.

Table 1. For some log files stored in text

EventLog.Exists	Determines whether the specified log exists
Event Log.clear	Clears all items in the event log
EventLog.CreateEventSource	Establish - an event source that can write an event to a log
EventLog.WriteEntry	Write related event records to the log
EventLog.SourceExists	Determines the existence of an event source
EventLog.GetEventLogs	Gets an array of logs for an event

For text class log files, we can go through Microsoft. NET the FileSystemWatcher data on the platform to monitor and obtain the corresponding electronic data information in real time.

Table 2. For text class log files

Method	Function
FileSystem Watcher	Class FileSystemWatcher initialized to monitor the specified file type of the specified directory
OnChanged	A Changed event is raised that calls the method when the directory file changes
OnRenamed	Causes a Renamed event and calls the method when the file directory or file is named

On Deleted	Causes a Deleted event to be called when a file directory or file is deleted
WaitForChanged	Returns a structure that contains specific information about changes that occur
	• • •

We can also view Windows logs through our own event viewer in the Windows system. Open event viewer method: start -> run - quickly open the tool > input eventvwr-> enter.

Forensic Analysis of Windows Logs

(1) According to the characteristics of log files and the challenges in log file processing, we propose a forensics analysis method based on Windows host log. The main elements are as follows:In order to solve the problem that log records are easily damaged and can not be obtained, we study the method of log repair, and design the method of log repair according to the update strategy of flag bit, offset copy and log file.

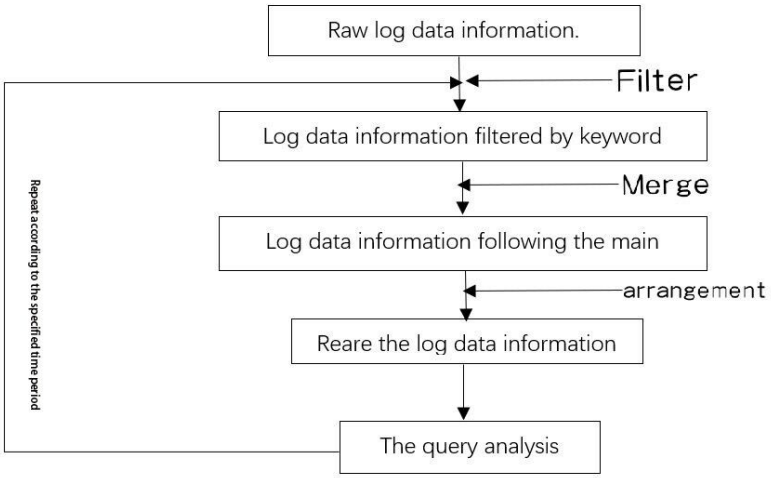


Figure 1. Log data analysis process

(2) After generating rules based on frequent item sets, we envisage a format rule matching method based on simulation attacks. First, the rules related to the attack are excavated by simulated attack, then the formatting rules are extracted according to the rules. Finally, the attribute prefix and formatting rules of the rules to be matched are matched to extract the suspicious rules related to intrusion in the rules to be matched.

(3) We envisage scenario reconstruction based on rule base and attribute tracking. The method is based on the rule base. According to the intrusion step, the corresponding rules and records are found out, and combined with the registration changes, the scene reconstruction is realized by the intrusion step according to the attribute correlation between the intrusion steps. The experimental part shows the process of the method in detail through a specific case. Finally, the feasibility and effectiveness of the method are verified by comparing and differentiating the experimental results.

Python Technological Inquiry

Python Technical Summary

As we all know, Python is a programming language that explains variables, comprehensive objects, dynamic analysis, data research types. It is easy to operate and can be programmed automatically. It has strong stability. It is mostly used in simulation calculation, data processing analysis, software development and web crawler. At the same time, in network technology, Python have a complete and powerful variety of standard libraries and third-party libraries. These include (1) Urlib library (2) Beautiful Soup library (3) Threading library;

Based on this, we intend to crawl the relevant data of the target through the rule base and attribute tracking, collect and store the structured data to the database, and other unstructured data are converted into a specified format by binary. Store in a local log file. Multiple simulation mining, and the attack on the target into the crawler reading, collection, processing, storage, and then find the invasion related suspicious rules.

Python Network Mapping Principle

Taking Python3.7 as an example, under the multi-type basic library and the tripartite library, we can effectively detect and interact the related network through the IPV6 Baotou data structure, and make an effective analysis of this.

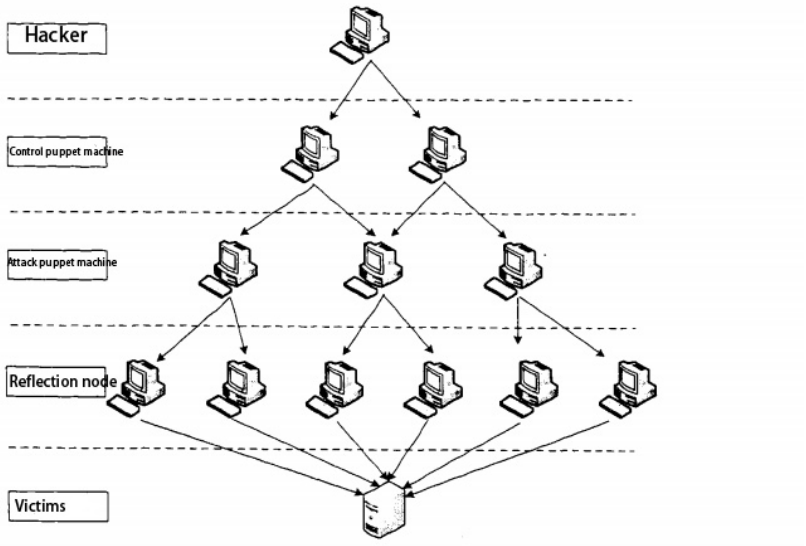


Figure 2. DDOS reflection attack architecture

In the reflection attack, the intruder does not perform puppet program operation, which makes the physical position of the reflection node more dispersed. moreover, in a reflection attack, intrusion packets are actually normal' P packets with legitimate P source addresses and packet types. therefore, packet filtering techniques based on address spoofing and routing-based DoS detection mechanisms are difficult to function.

Therefore, we only need to sniff the data stream, track the attributes of the target, and use the IPV6 Baotou data structure to obtain the intrusion target TCP header. As shown, Baotou format string and part of the source code.

```
if protocol == PROTOCOL_TCP:
        stripTCPHeader = packet[ipHdrLength:ipHdrLength + 20]
        tcpHeaderBuffer=unpack（'!HHLLBBHHH'， stripTCPHeader）
elif protocol == PROTOCOL_UDP:
        stripUDPHeader=packet[ipHdrLength:ipHdrLength+8]
        udpHeaderBuffer=unpack（'!HHHH'， stripUDPHeader）
```

Figure 3. Baotou format string and part of the source code

Current Problems and Difficulties

Although the log file is generated by the running of the internal program of the system, it records the details of the computer operating system, the behavior of the application and the behavior of the user and the use of the system, the abnormal events in the system and so on. However, most of the current computer forensics methods are ex post evidence, which makes the traces in some logs incomplete or contaminated. There are also defects in the log itself, which makes the log analysis technology immature, and the information contained in the log file can not be thoroughly excavated and applied. We therefore summarize the reasons for this situation as follows:

(1) The operating system's protection measures for log files are not perfect, and computer users lack the habit of protecting log files, which makes log files easily damaged, easily tampered with, and can not be used as legal evidence;
(2) At home and abroad, the research on computer log forensics is not mature, and the related technology needs to be developed, and the relevant professionals are not many;
(3) Although many types of log files are included in the system, log research is still in its infancy and there is no relevant standardized method for defining valid and legitimate electronic evidence;
(4) Log files record the details of each event, which makes log files large and complex, and manual analysis is not realistic, but there is no mature and practical log analysis tool;
(5) At present, the study of log files is limited to the acquisition and preservation of log files, and the analysis of the contents of log files is less.

Conclusion

Computer network crime and illegal outside the place, for public security personnel, skilled use of information means to crack down on illegal crime is the current trend and problem of public security work. Based on the characteristics of the log files, this project explores and analyzes the common log files and the acquisition methods of special log files through computer Windows log, and summarizes the current situation of electronic physical evidence collection at home and abroad. It lays a foundation for future research and even public security work. We believe that with the arrival of the 5 G era, for the future computer crime,

recording the process of crime, extracting and preserving the traces of crime is the key to electronic physical evidence and even solve the case.

Funding

This paper was supported by the National Undergraduate Innovation and Entrepreneurship Training Project, "Research on the way of obtaining electronic physical evidence based on Python—take Windows log document as an example" (基于 python 的电子物证获取方式探究 —— 以 Windows 日志文件取证为例), People' s Public Security University of China.

References

Sahoo, R. K., Oliner, A. J., Rish, I., Gupta, M., Moreira, J. E., Ma, S., ... & Sivasubramaniam, A. (2003, August). Critical event prediction for proactive management in large-scale computer clusters. In *Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining*, 426-435.

Fu, S., & Xu, C. Z. (2007, October). Quantifying temporal and spatial correlation of failure events for proactive management. In *2007 26th IEEE International Symposium on Reliable Distributed Systems (SRDS 2007)*, 175-184. IEEE.

Salfner, F., & Tschirpke, S. (2008). Error Log Processing for Accurate Event prediction. In *USENIX workshop on the analysis of System logs (WASL)*.

Lou, J. G., Fu, Q., Wang, Y., & Li, J. (2010). Mining dependency in distributed systems through unstructured logs analysis. *ACM SIGOPS Operating Systems Review*, 44(1), 91-96.