Design and Implementation of Cybercrime Investigation Systems Based on Open Source

Chiu-Niang Chen and Chung-Huang Yang
Institute of Information and Computer Education, National Kaohsiung Normal University

Abstract

As the fast development of information technology, the Internet has become a principal source of information we rely on and brings human beings assorted convenience in life and at work, but on the other hand the cyberspace also grows as the hotbed of crime. Internet pornography and cybercrime are being formed rapidly day by day threatening the social security and economic development. The detection of cybercrime has become very complicated and difficult. The filtering of the hazardous key words can prevent the crime from happening and the use of the network sniffers while the crime is happening or after the crime has happened can collect evidence for the investigation. The proper preparation of creating and maintaining the digital evidence can help the follow-up cybercrime forensics. Based on the open source OWNS software, this research extends its functions and combined with database to provide remote query on crime characteristics. Our distributed cybercrime investigation system come with Chinese interface which realizes cybercrime inspection, digital evidence collection and computer forensics.

Keywords: Cybercrime, Computer forensics, Digital evidence, Network sniffer, Network traffic.

1. Introduction

According to the report "Internet Broadband Usage in Taiwan" by Taiwan Network Information Center released in July 2006, it is estimated that the Internet user in Taiwan, by the end of June 2006, was approximately 15.38 million, and 12.25 million of them were broadband users. 94.63% of the broadband users get online at home with the main purpose of browsing webpage and using email [14]. The impact of technology has permeated our personal life and work. It is so convenient to use email and group chat that social interaction has been changed in some respects. These kinds of changes have also occurred in business and economic activities. The real-time and no-boundary features of the Internet and the availabilities of commodities being displayed for sale through the Internet make trades between providers and consumers accomplished in an instant and efficient way, thus subverting the traditional trading manner. Accompanying the growing Internet population, the activities in the virtual world of the Internet include proper usage as well as wrong doings. Because of the relatively low cost and low risk due to the concealing characteristics of Internet, the cyberspace has become a breeding ground for various crimes. Illegal acts conducted by using Internet are gradually growing and extended from computer crime to many other forms of social problems. Illegal incidents such as pornography, information theft, hacker intrusion are rapidly growing; cybercrime cases are emerging in an endless stream.

Governments, especially in the US, have started to realize that in terms of academic research, shopping and entertainment, the Internet is a powerful tool.

Nevertheless, it has also become a tool for criminals and terrorists to use and allows child pornography, swindlers and financial hackers to find a hiding place on the concealed WWW [3]. As a result, the FBI in the US has developed the Carnivore (DCS1000) system, which is a network packet filtering and monitoring system used for assisting with crime investigation [13] to cope with the increasing development of cybercrime and terrorist activities. Although it used to cause legal and privacy disputes, after the tragic occurrence of the 911 terrorist attack, the US Congress and government departments have granted many new authorization to law enforcement and terrorist activities related crime prevention [7]. Digital evidence collection under the concrete and comprehensive laws and regulations should be able to avoid doubts from the public.

Criminals may use computers in one of two ways in support of their actions. Either as the repository for information relating to their criminal activity or as a tool in actually committing a crime [4]. To effectively and comprehensively prevent cybercrime incidents from happening and provide crime evidence collection and analysis, forensic theories and technologies for information communication security have become increasingly important. There are three major areas in research for capacity planning of information communication security forensic technologies: information policies and security management, criminal science, computer science. Criminal science comprises the categories of information forensics, computer forensics, digital evidence and cybercriminology which has increasing importance and necessity due to the rapid growth of cybercrime. The survey results of the future needs for computer forensics have also shown that 3 of the top 5 needs are related to forensic technologies and tools [9]. Therefore, more efforts to put into the work and research related to future cyberspace information communication security forensic theories and technologies are necessary.

Without the help of appropriate software tools, identification and collection of digital evidence cannot be visually identified and analyzed directly. Therefore, it is necessary to establish a set of scientific methods for evidence identification and collection. Applying suitable software tools to identify and collect digital evidence is the main requirement for cybercrime investigation. For example, Encase and FTK are tools often used by law enforcement and investigator. Tools applied in the areas related to the computer forensic science have distinguishing features respectively [5]. The computer forensic science has been extensively researched for many years in Taiwan, yet it tends to aim for the establishment of policies and standards, management and operating procedure principles. The discussion of forensic tools is limited to forensic procedures of software tools from overseas. For supporting forensic personnel to identify, collect and examine cybercrime evidence with friendly user interface in today's increasingly rampant cybercrime environment, it is still deficient in solutions that are low in cost, high in performance, easy for deployment and management, capable of multi-mode analysis with filtering function and supportive of flexible scaling frameworks. In this paper we present a solution in attempt to solve those shortages.

2. Related Work

This research is aimed at cybercrime forensic investigation, therefore computer related forensic science, the network monitoring technologies and the crime investigation technologies used by the FBI of the United States will be discussed.

2.1 Computer Forensic Science

Computer forensic science is the branch of digital forensic science. The term "computer forensics science" was first addressed by IACIS in 1991. It is mainly targeted at the research of relevant digital evidence collection methods for the on-the-scene evidence left in the cybercriminal activities. “Digital forensic science” is the use of scientifically derived and proven methods toward the preservation, collection, validation, identification, analysis, interpretation, documentation and presentation of digital evidence derived from digital sources [6]. “Computer forensic science” is the use of an expert to preserve, analyze, and produce data from variable and invariable media storage [10]. In short, the purpose of computer forensic science is to ensure invariability and integrity of digital evidence during the collection process to allow collected evidence to have the evidence capability to achieve the ultimate purpose of court acceptance.

2.2 Network Sniffer Technologies

Sniffer, utilizing the "sniffing" technology, is a mechanism most often used to investigate cybercriminal activities. Many network intrusion detection software and network sniffer software are developed using this technology. For investigators who detect cybercrime and collect digital evidence, the use of Sniffer tools is a must.

Packet Sniffing is a method which can capture and log network traffic. This is the technique a user can use to sniff out other people's information. Packet sniffers can be hardware or software which connected to network of monitoring devices. When data streams are transmitted back and forth over the network, the packet sniffers can capture each packet and decode and analyzes its contents. Network managers often use packet sniffers to manage and monitor networks as well as diagnose and test network problems. However, packet sniffers are also used often by people who intend to be engaged in illegal activities.

Normally, a network interface card (NIC) only accepts incoming packets which are specifically for it. When a NIC is placed in promiscuous mode, it can accept all incoming packets that it sees, regardless of their intended destinations. Packet sniffers generally work by placing the NIC in promiscuous mode; the user then configures the sniffer to capture all packets or only those with particular characteristics. This promiscuous mode can only be operated in Ethernet that use hubs as connection equipment; packets under such environment are transmitted using a broadcasting method. Another approach to monitor traffic on a switch is to utilize a feature called Switched Port Analyzer (SPAN) in which the packets are transmitted using a point-to-point method. A SPANned port (mirrored port) enables eavesdropping by copying network traffic from one port on the switch to another [2]. A computer installed with packet sniffer is connected to the mirror port to capture network traffic to achieve the purpose of network sniffer.

2.3 Discussion of the Carnivore System

The Carnivore system [13] is the third generation of online investigation forensic tool developed by the FBI (Federal Bureau of Investigation) of the USA in 2000. It can be used to capture and filter suspicious criminal activity information on the network.

(1) The Evolution of the Carnivore System

In 1997, the FBI deployed its second generation of investigation program called Omnivore. According to the FBI, Omnivore was designed to look through e-mail traffic streaming over a specific ISP and capture the e-mail from the targeted source, saving it to a tape-backup drive or printing it instantaneously. In 1999, after Omnivore was retired, a more comprehensive system DragonWare Suite was developed. This software enables FBI to reconstruct the e-mail messages, the downloaded files or even the web pages [8]. Despite little official information about DragonWare Suite and Carnivore has ever been made public, it is basically a packet sniffer application in the network. This software can sniff data stream in a certain network, capture and examine data packets that flow through it. Components of DragonWare include **Carnivore**, which captures and filters the network traffic; **Packeteer**, which is an application for reassembling packets into cohesive messages or Web pages; **Coolminer**, which is a process for inferring and analyzing data found in the messages and displays collected information using a browser.

(2) Carnivore System Architecture

The Carnivore system architecture [13] comprises:

- **One-way tap**: It is a one-way wiretap that connects Ethernet.
- **Collection computer**: It is a general purpose computer to filter and collect data. This computer does not need a keyboard and a screen installed. It is controlled by the remote monitoring software pcAnywhere, which is used by the controlling computer that the carnivore system is installed.
- **Control computer**: It is used to control the "data collection computer" and to check the data. Packeteer and CoolMiner sofware are installed in this computer.
- **Telephone link:** It is the connection between the "data collection computer" and the "control computer".
- **Carnivore software**: It is one of the components of the DragonWare Suite. The other two parts are Packeteer and Coolminer.

(3) The Operating Modes of the Carnivore System

The data collection mode of the Carnivore system can be divided into the "full mode" and the "Pen mode". The full mode can capture the complete communication content, such as the full message of an email. The Pen-register mode, take the email as an example, can only capture the email address and the IP address of senders and receivers. The other data has to be abandoned. The content allowed to be captured is completely dependent on the authorization from the court.

The Carnivore system is a passive wiretap and will not interrupt the normal communication of the network. The operation procedures are that when the FBI is reasonably suspicious of someone being engaged in criminal activities, they will request the court’s permit to observe online activities of the object and collaborate with the ISP to install "one-way tap", the monitoring device, at the confirmed location to capture online communication content and send it to the "data collection computer". The FBI then filters the content according to the collection range authorized by the court and saves the filtered data as the reference basis for the law enforcement and the court trial.

3. Research Background and Purposes

Computer crimes are profitable criminal activities and continue to grow rampantly as the wide-spreading of computer and high frequency of computer use increased [1][12]. The cybercrime that they are transnational, concealed, different in legal regulations and hard to obtain evidence. Therefore, it is not easy to recognize the true identity of a cybercrime in the highly concealed network environment, which makes the investigation of cybercrime very difficult. In addition, it is very easy for the digital evidence, which stored as files in computer systems or in networks, to be modified or deleted and consequently affects its proving ability as evidence and ravages the ability of decreasing computer crime and resolving Internet law disputes. As a result, law enforcement and investigation agencies are facing more and more arduous challenges in cybercrime cases.

How to detect the cybercrimes in advance and proactively prevent them through the precautionary mechanism plus retrace the crimes development after their commitment have become extremely important research issues today. A general network intrusion detection systems includes two detection modes. The first one is the single or distributed detection of network intrusion and the other one is the host intrusion detection and they both provide precaution services for intrusion incidents [15]. For cybercrime precautions and evidence collection, the criminal characteristics can be sifted and the criminal activities can be recorded for specific or randomly variant targets by applying the method similar to network intrusion detection and utilizing the network Sniffer technology. Using network traffic as a source for obtaining evidence, investigators usually use open source tool and commercial software to assist their investigations. Normally open source tools are used only for simple cases as they provide only basic functions. They are lacking functions of business software tools that are specifically designed for handling network traffic as evidence collection and the integration functions are deficient as well[5].

The research and development work for areas related to computer forensic science in Taiwan needed to made more efforts to include integration and value extension of functions of relevant computer forensic tools as well as localization of them. This should be the research focus in the future. However, for supporting forensic personnel to identify, collect and examine cybercrime evidence with friendly user interface in today's increasingly rampant cybercrime environment, it is still deficient in solutions that are low in cost, high in performance, easy for deployment and management, capable of multi-mode analysis with filtering function and supportive of flexible scaling frameworks. In this paper we will present a solution in attempt to solve those shortages.

4. Design and Implementation of Distributed System Framework for Cybercrime Investigation Systems and Its Localization

4.1 System Design

To improve the integration and the value extension of functions of the forensic tools used in cybercrime detection and prevention and to better localize them, we focus on the following objectives in order to obtain efficient and effective solutions:

- **Localization of complete open source systems:**
  It can efficiently reduce the total cost and increase the maintainability of the system. The localized operating interface and the interactive messages can effectively help common users and forensic personnel to work with the system because of the Chinese processing capability.

This meets the essential requirements for users in Taiwan.

- **Implementation of distributive and layered multi-tier dynamic architecture:**

  The establishment of a distributed architecture is easy for integrating sniffing operations cross different networks and the system built through the multi-tier framework will be able to improve effectively the deployment, expansion and the usability of the system.

- **Operation on cross-platform:**

  Propose the idea of providing independent or integrated operating mechanism for working on cross-platform systems to current mainstream operating system company such as Linux-Base or Microsoft Windows.

- **Multiple data filtering and searching including realtime pre-processing and offline post-processing:**

  Condition filtering can be performed on specific keywords and the IP addresses of sources and destinations in real-time running to increase the effectiveness of the data collection and process. The offline full-text-retrieval under combined condition filtering through flexible modules can be operated on visualized data provided so as to be able to thoroughly strengthen any insufficiency of the online network sniffer.

- **Providing with complete monitoring database:**

  Saving keyword-related monitoring records by connecting to the database and providing searching functions of compound to monitoring records enables the scalability of criminal characteristic analysis performed by OLAP.

- **Providing with file-based original evidence data:**

  To avoid the unidentifiable of raw data recorded by some sniffer service, the system will reassemble the raw data into specific classes and save them into individual files, hence advantage the restoration of evidence and improve the related application of searching and analysis.

- **Applying cryptography technology for data integration:**

  While forensic monitoring index data is recorded to the database, its hash value calculated using hash algorithm is also saved to ensure the data integration.

- **Providing with Web-based management mechanism:**

  Provide Web-based remote management service will make efficient independent operation possible and will be able to flexibly establish the integration with other applications using XML through Web Services.

4.2 System Architecture

To implement the objectives of the system design mentioned in the previous section, a flexible system framework is presented in Figure 1.

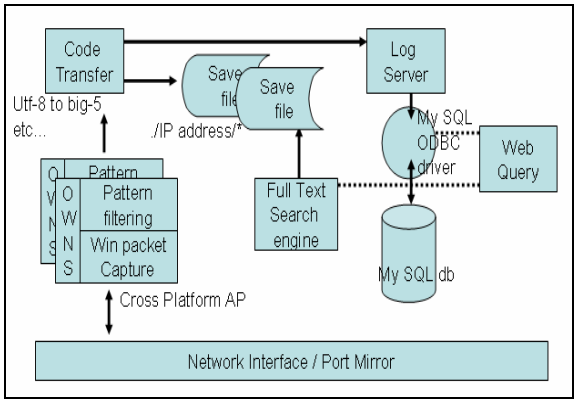

Figure 1. System Architecture

We use Borland Delphi/Kylix to develop a cross-platform-service application system based on OWNS[11], the open source system, and mySQL database. Service functions are described as follows:

■ **Distributed Agent-based packet Sniffing**

We extend the original OWNS (One Way Network Sniffer) service functions to create distributed agent-based tools that can be operated across platforms and across subnets concurrently for on-line packet sniffing, analysis and filtering. The system will send the characteristic data that matches the filtering conditions to the database and calculate and save the hash value for the raw sniffing data to LogServer database. This flexibly expands the collection range of digital evidence for the subsequent forensic analysis work. See Figure 2 for the front-end tools that are function-enhanced and localized.

Figure 2. Localized Front-end Tools

■ **Code Conversion**

Convert code for improving searching and filtering of data from different aspects.

■ **Original Network Sniffing Data Storage and Full-text retrieval**

After reassemble the original network packet sniffing data, the system stores it as a file in the directory name composed by its source and destination IP addresses in order to use as the evidence in the future and provide the full-text searching tool with related data to process extensive characteristic retrieval.

■ **Log Server and Database**

Establish an independent Log Server for the integration of database. The system provides the ability to receive filtered characteristic data delivered by several distributed agent-based tools concurrently. The characteristic data, such as IP addresses of source and destination, keywords, save file path names after reassembling, time and hash values of data file, can be distinguished by IP address of front-end tools for its coming source ,enable processed and analyzed in the future.

This research uses Borland Delphi/Kylix to create Log Server, which is combined with a database to store characteristic data. See Figure 3 and 4 for the demo of Log Server collaboration with OWNS clients.

Figure 3. The Operating Screen of Log Server

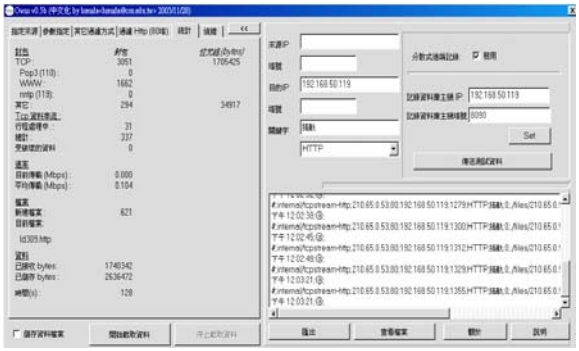

Figure 4. The Operating Screen of OWNS Clients

■ **Web Inquiry Service**

Use web-based technology to create an inquiry service with high usability. Remote data can be retrieved using source IP addresses, destination IP addresses and keywords as query conditions.

5. System Presentation

The completion of the system implementation fully satisfies the above described research objectives and system design requirements for cybercrime investigation systems. The major results of this research are demonstrated with the main features of the system operation:

(1) Designate sniffing source target on the localized agent-based tool interface.

Designate a packet sniffing source target on the agent-based tool interface and perform packet monitoring for a specific network interface to capture packets that is copied or transmitted through the NIC by Port Mirror, the network application as seen in Figure 5 or directly load the files stored of network packets that are acquired through network sniffer of the ethereal tool for capturing packet data and reassembling classified archives.

Fig 5. The agent-based tool to designate a sniffing source target (NIC).

(2) Designate a complete file storage path and a classification method.

Designate a location and a type for complete file storage path from the agent-base sniffing tool after the monitored data is reassembled and classify and save the files through a specific directory structure for the application of investigation search as seen in Figure 6.

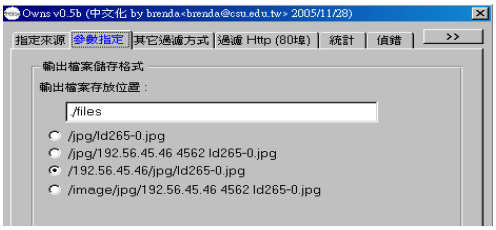

Figure 6. Settings of the File Storage Path and the Classification Method

(3) Transit the filtered keyword data to Log Server after online condition filtering, file reassembly and storage.

Designate online filtering conditions on the agent-base tool to start performing packet sniffing as seen in Figure 4. The client program returns the filtered characteristic data to the designated Log Server, saves them in the database, automatically classifies the captured data based on the selected archive type and saves it to the local hard disk as seen in Figure 7. The archived content (the MSN messaging content) is shown in Figure 8.

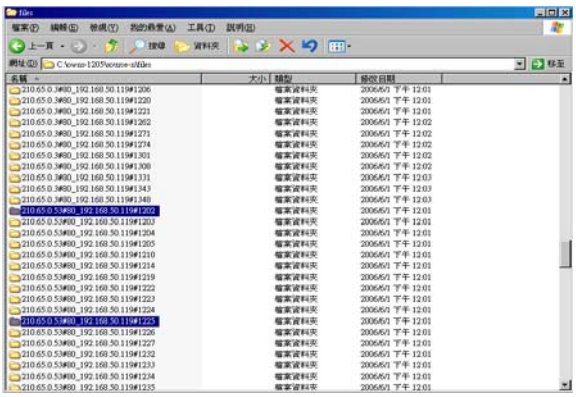

Figure 7. Data Classification and Storage after Reassembly

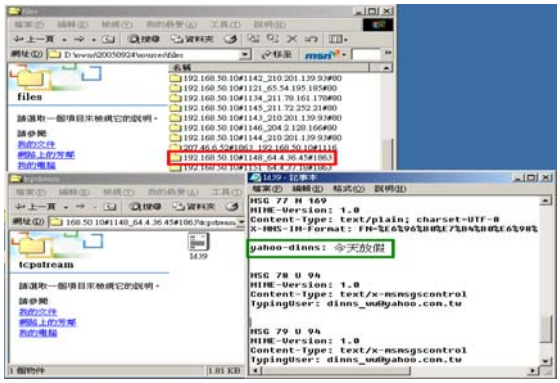

Figure 8. The Archived Content (the MSN Messaging Content)

(4) Retrieve through full-text searching service for collected data.

For data files stored on the disk, plug-in full-text searching engine modules (e.g., the Torndo System) can be called to create a pre-stored index to expand integrated inquiry service functions as seen in Figure 9.

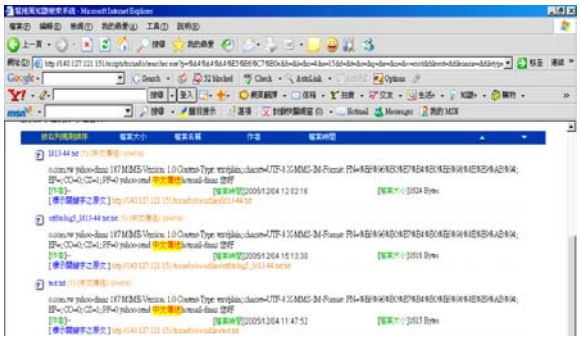

Figure 9. To saved file with full-text searching

(5) Web-based online compound condition search

Log Server integrates the database and web-based technology to provide remote online compound condition filtering, retrieval and searching services. It can simplify and reduce the procedures and range of dangerous criminal data searching. It can also reduce the searching time for criminal evidence data in original files via providing of file paths as seen in Figure 10.

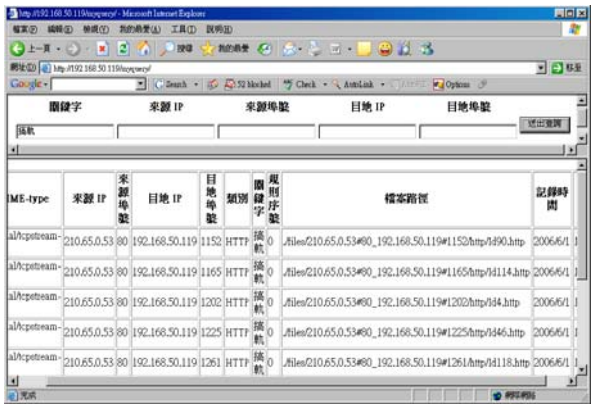

Figure 10. Web-based Online Condition Search

6. Conclusion

As today's cybercrime incidents emerge rampantly, how to take precautions and proactively prevent cybercrime from happening, or how to collect evidence data through network sniffing and correctly retain, analyze and identify criminal incidents during and after their occurrence, are the essential issues requiring resolutions urgently. This research present a cybercrime investigation system based on the open source OWNS and a distributed architecture combined with the database technology. The user interface and interactive messages are fully localized and a flexible multi-tier system framework that common sniffing tools are lacking is implemented. The installation cost is low and it is easy to further expand service functions for special needs. The distributed cross-platform and concurrent multi-source data collection function is easy to deploy and is flexible on lightweight data collection. Raw data can be reassembled in a real-time manner. It is also easy to directly carry out high-performance processing in the memory for data collection. The file-based online keyword filtering mechanism and the offline full-text searching framework are easy for data analysis with the full view of the source information, therefore can effectively avoid

processing complexity and increase overall operating performance. Dynamically creating a centralized database for criminal characteristic data is easy for combining the data mining mechanism for investigation, prevention and precaution. Remote criminal characteristic data inquiry system based on Web service is integrated and flexible, which makes it easy to be able to combine directly with other services to generate a synergy effect. A good solution, which is low in implementation costs, highly convenient in use and efficiently in operation, is very helpful for cybercrime investigation, prevention and evidence search. It is also good for maintaining overall order of cyberspace and its security.

References

[1] CERT/CC , Cert statistics 2003, Available from: http://www.cert.org/stats/, Retrieved June 1, 2003.

[2] Cisco. Configuring the catalyst switched port analyzer (SPAN) Cisco Tech Notes, Available from: http://www.cisco.com/warp/public/473/41.html ,2003 .

[3] Donald Kerr, Hearings on the Carnivore Diagnostic Tool Before the Subcomm. on the Constitution of the House Comm. on the Judiciary, 106th Cong. (2000), (statement of Donald Kerr, Assistant Director, Laboratory Division, Federal Bureau of Investigation).

[4] Dorothy A. Lunn, “Computer Forensics – An Overview,” SANS Institute 2000 – 2002, February 20,2001.

[5] Eoghan Casey, “Network traffic as a source of evidence:tool strengths, weaknesses, and future needs, ” Digital Investigation (2004) 1, 28-43.

[6] G. Palmer, “A Road Map for Digital Forensic Research,” Report from the first Digital Forensics Research Workshop, Utica, New York, August 2001.

[7] Holly E. Ventura, J. Mitchell Miller, and Mathieu Deflem, “Governmentality and the War on Terror: FBI Project Carnivore and the Diffusion of Disciplinary Power, ” Critical Criminology, Volume 13, January 2005 , pp. 55-70 .

[8] Jeff Tyson, ‘‘How Carnivore Works,’’ Available online: http://www.howthingswork.com , 2003.

[9] Marcus K. Rogers, Kate Seigfried, “The future of computer forensics: a needs analysis survey, ” Computers & Security (2004) 23, pp. 12-16.

[10] Mary Mack, “Electronic Discovery vs. Computer Forensics,” New Jersey Law Journal, 2003, page 1.

[11] OWNS, Available from: http://owns.sourceforge.net/.

[12] R. Richardson, “2003 CSI/FBI Computer Crime and Security Survey, ” Computer Security Institute, Available from: http://www.gocsi.com.

[13] Stephen P. Smith, Henry Perrit Jr., Harold Krent, Stephen Mencik, J. Allen Crider, Mengfen Shyong, Larry L. Reynolds, “Independent Technical Review of the Carnivore System, ” Final Report, IIT Research Institute, Dec. 2000, http://www.epic.org/privacy/carnivore/carniv_final.pdf .

[14] TNIC, Internet Broadband Usage in Taiwan- A Summary Report of The July Survey of 2006, 2006-8.

[15] Vern Paxson,“ Bro: A System for Detecting Network Intruders in Real-Time, ” Computer Networks, 31(23-24), Dec. 1999 , pp. 2435-2463.