LAW ENFORCEMENT TOOLS AND TECHNOLOGIES

FOR

INVESTIGATING CYBER ATTACKS

A NATIONAL NEEDS ASSESSMENT

INSTITUTE FOR SECURITY TECHNOLOGY STUDIES
AT DARTMOUTH COLLEGE

June 2002

Michael A. Vatis
Director
45 Lyme Road
Hanover, NH 03755
(603) 646-0700
www.ists.dartmouth.edu

FOREWORD

Cyber attacks on corporate, governmental, academic, and critical infrastructure networks are increasing in number, sophistication, and severity.[1] The tools and technologies that law enforcement investigators use to investigate these attacks are not keeping pace with the instruments employed by attackers. Commercial research is largely focused on security products likely to yield near-term profits, and therefore may not adequately address the needs of the relatively small law enforcement market. What is sorely needed is mid- to long-term research into the technologies required by law enforcement. The Institute for Security Technology Studies' (ISTS) *Law Enforcement Tools and Technologies for Investigating Cyber Attacks: A National Needs Assessment* is the result of a comprehensive examination of the technological impediments law enforcement encounters during cyber-attack investigations.

ISTS's mission is to serve as a center for counterterrorism and cyber-security technology research, development, and assessment. Researching and developing tools and technology that address the needs of law enforcement are an important part of the Institute's mission, given the critical role of law enforcement in preventing and responding to physical terrorism or cyber attacks. Providing solutions for the problems identified in this needs assessment, however, is an enormous undertaking, far too large for any single institution to assume. Accordingly, the goal of this needs assessment is to set forth in detail the technology impediments encountered by cyber-attack investigators and to present law enforcement's technological requirements. The next step is to identify available technological solutions that address the requirements outlined in the needs assessment. The third step is to produce a gap-analysis report identifying the critical areas, or gaps, where scientific research should be focused. The ultimate goal is a national agenda for research and development of law enforcement tools and technologies which scientists across the country can use as a guide in determining where to devote their research efforts.

Cyber attacks are a significant threat to the United States' national and economic security in the 21st century. Research and development will play a crucial role in ensuring that law enforcement has adequate tools and technologies to respond to cyber attacks. It is our challenge, in the scientific community, to examine the problems detailed in this document, undertake the necessary research, and deliver solutions to empower law enforcement.

Michael A. Vatis

Director, Institute for Security Technology Studies

[1] For example, see the *2002 Computer Crime and Security Survey* conducted by the Computer Security Institute. This document is available at <www.gocsi.com/forms/fbi/pdf.html>.

EXECUTIVE SUMMARY

The technological impediments facing cyber-attack investigators demand immediate attention. The opportunity exists to address these needs and empower law enforcement by the application of science to operational requirements. The *Law Enforcement Tools and Technologies for Investigating Cyber Attacks: A National Needs Assessment*, conducted by the Institute for Security Technology Studies, provides insight into the technological obstacles facing law enforcement during cyber-attack investigations.

This needs assessment is a first step toward developing a national research and development (R&D) agenda for cyber-attack investigative technology. The next step is the collection and evaluation of existing tools and technologies. By comparing law enforcement requirements with existing solutions, we can identify the gaps where technological R&D is necessary. The R&D community can then prioritize its research to fill the gaps identified and give our public servants the tools they need to address one of the critical public safety and national security issues of the 21st century.

The first dedicated inquiry of its kind, the needs assessment was conducted over six months in late 2001 and early 2002. Primary data was collected through three approaches. First, a web-based survey of federal, state, and local law enforcement was conducted over four months under ISTS auspices by the RAND Corporation. Second, ISTS researchers visited law enforcement agencies in seven states and the District of Columbia to conduct in-depth interviews with cyber-attack investigators. Third, during a two-day workshop held December 3-4, 2001, ISTS and RAND presented the data collected from the survey to a select group of current and former cyber-attack investigators and prosecutors for validation and additional data collection. The resulting needs assessment reports the technological impediments faced by, and the needs of, cyber-attack investigators as they perceive them. It is important to note that in some cases the issues discovered during the course of this study may be addressed by existing technologies — a topic that will be addressed in the next phase of our study.

The Investigative Process: Preliminary Investigation and Data Collection (Page 14)

Several of the obstacles encountered during initial stages of a cyber-attack investigation can be addressed by technological solutions.

- Although multiple versions of the Windows operating system are most commonly encountered, cyber-attack investigators need to have at least a minimum understanding of other operating systems. Investigators also need ***tools to automate the collection of data files from multiple operating systems*** in victims' networks.
- Investigators said it would be useful to have a map of the network topology in the course of their investigations. ***Solutions to automate the process of developing the map of the network*** and to detect settings, recognize hardware, and graphically represent the results would be valuable to cyber-attack investigators.
- Study participants also expressed a desire for ***cyber-attack-specific data-recovery tools*** to automate the data recovery process. Currently many law enforcement organizations use specialized forensic examiners to retrieve necessary data from computers compromised by cyber attacks, who then must personally examine the

compromised machine(s) or network, which can be very burdensome and time intensive.

- Investigators and supervisors need new ***solutions for retrieving, storing, and analyzing very large media storage devices***. Many state and local law enforcement agencies are concerned about their long-term capacity to meet the regulations set by their state legal systems pertaining to the storage of digital evidence.

The Investigative Process: Log Analysis (Page 22)

The inadequacy of existing data and log analysis tools is perhaps the most serious technological impediment for cyber-attack investigators. Difficulties filtering out irrelevant information, managing large sets of log files, and synthesizing information from multiple log files were all cited as significant obstacles. Investigators spend a significant amount of time — on average 23% in a typical investigation — interpreting and analyzing log files. Investigators need new approaches to expedite this process, since critical data may be available for only a short period following a cyber attack.

- Law enforcement needs ***solutions that recognize and import logs from multiple platforms across a network, correct for errors in time stamping and place events into an organized timeline, perform analysis on the logs, and have the ability to link with other services***, such as exploit signature databases.
- The development and adoption of common data-sharing and communications protocols would greatly improve the ability to share digital information across jurisdictions. ***Tools with the ability to organize log data sets into a common format, easily transferable to other law enforcement agencies or to other software***, would assist cyber-attack investigators significantly.

The Investigative Process: IP Tracing and Real-time Interception (Page 28)

Internet protocol (IP) tracing and real-time interception is critical for tracking cyber attackers.

- Numerous online tools are available to help investigators trace IP addresses. In general, investigators expressed satisfaction with these tools. But IP spoofing presents significant problems for certain types of cyber attacks (for example, distributed denial of service attacks or DDoS) since the origin and location of the attacker remain hidden. ***Non-technical issues such as underemployed technologies to counter attacks utilizing spoofing and lack of record keeping by Internet service providers hamper the tracing of IP addresses.***
- Available technologies for conducting real-time interception of data communications were reported to be too complex to implement. ***Participants expressed the need for the development of real-time interception solutions that could be used by law enforcement officials with a moderate level of technical expertise and training***.

Emerging Technologies Requiring Research and Development (Page 34)

New technologies hamper law enforcements' ability to conduct successful cyber-attack investigations.

- Among the emerging investigative issues addressed during the needs assessment, the ability to defeat encryption was a leading priority. ***Techniques to circumvent encryption are necessary to assist law enforcement***, since current brute-force techniques to crack encryption are usually ineffective.
- Technological advances in wireless communication present new challenges to law enforcement as mobile attacks are difficult to locate. Investigators expressed a need for the development and implementation of ***technologies that can recognize unauthorized wireless network access and detect physical locations***.
- Steganography also presents immediate and long-term challenges for law enforcement. The study participants expressed the ***need for a clearinghouse of digital steganographic programs and signatures that could be consulted during forensic analysis*** of a seized computer, to flag the possible use of this data-hiding technique, as well as additional long-term research into breakthrough technologies for steganography detection.

National Information Sharing (Page 40)

Several topic areas throughout this study revealed a recurring theme: the need for multiple information-sharing services pertaining to cyber-attack issues.

- The national discussion, development, and adoption of common data-sharing and communications protocols would greatly improve the ability to share information across jurisdictions. A ***tightly integrated data-sharing approach, engineered into the next generation of investigative solutions,*** would provide the foundation for national cyber-attack information databases.
- Workshop participants, after discussing the survey feedback on this data-sharing topic, pointed out the ***need for a database of cyber-crime investigators in each jurisdiction, including their experience, training, and contact information*** to facilitate coordination and collaboration on both investigations and non-investigations matters.

Law-Enforcement-Specific Development Issues (Page 46)

Law enforcement requires different levels of tool complexity to accommodate different investigators' skill levels.

- Solutions must be robust and have ***built-in capability to match the skill level of the user*** with appropriate functionality.
- Clear and unambiguous ***help files that are specifically written for multiple law enforcement users*** will have a significant positive impact on the success and usefulness of new solutions.

Training (Page 47)

Training is critical to law enforcement's ability to understand emerging technologies and successfully investigate cyber attacks.

- Study participants articulated the need to have ***training programs that fit law enforcement needs***.
- National initiatives should be undertaken to: ***benchmark current training programs and identify gaps; create a national training strategy for cyber-attack investigators, prosecutors, and the judiciary***; and benchmark distance-learning solutions for in-service training.

Conclusion (Page 52)

The next step in developing a national R&D agenda is a comprehensive national evaluation of existing tools and technologies. Standards will have to be developed to benchmark tools against law-enforcement-specific requirements. Existing solutions will need to be collected and evaluated. Analysis will then determine the gaps left by existing solutions, and thereby identify research priorities for the national R&D community.

CONTENTS

1. INTRODUCTION

This paper, the *Law Enforcement Tools and Technologies for Investigating Cyber Attacks: A National Needs Assessment*, conducted by the Institute for Security Technology Studies[2], provides insight into the technological obstacles facing law enforcement during cyber-attack investigations.[3] Identifying the challenges facing law enforcement in this critical area is the first step in a process to develop a requirements-based national research and development agenda for cyber-attack investigative technology. A national evaluation of existing tools and technologies will follow. Analysis will then provide insight into the gaps left by existing technology solutions and thereby identify research priorities for the national R&D community in this critical area.

1.1 NATURE OF THE PROBLEM

Technological advances have sparked unprecedented cultural changes over the past 20 years. Owing largely to the development and proliferation of personal computers, individuals now have access to vast amounts of information and advanced technologies through complex, interconnected networks. Modern nation-states rely on these networks for many purposes, including communications, commerce, and education. Moreover, the United States' critical infrastructure — vital services such as telecommunications, energy, banking, and government operations — depend on computer networks for their basic operations. Many of these infrastructures are privately owned, with limited governmental oversight.

Yet, many of these networks, and hence the infrastructure that relies on them, contain inherent vulnerabilities. Historically, secure software has not been a top priority for vendors. Increasingly short turnaround times for new product releases often leave little time for security testing. Additional features or functionality may provide vulnerabilities that cyber attackers use to exploit systems in short order.

As reliance on the nation's computer-dependent infrastructure continues to increase, the number of people who are able to exploit networked systems' weaknesses rises correspondingly. This phenomenon has had a significant effect on technologically advanced countries such as the United States. Interconnectivity, afforded by the broad adoption of common networking protocols[4], gives cyber attackers unprecedented access to critical national infrastructures.

The computer security market gives businesses enormous incentives to innovate. Much work is currently underway in the private sector to develop new virus[5] detection software,

[2] Additional information is available from the Institute's web site at <www.ists.dartmouth.edu>.

[3] The Institute for Information Infrastructure Protection, or I3P, is initiating a needs assessment regarding the broader topic of the nation's cyber-security and information infrastructure R&D priorities. More information is available at <www.thei3p.org>.

[4] Protocol: An agreed-upon format for transmitting data between two devices.

[5] Virus: A program that infects other programs by modifying them to include a copy of itself. A worm is a program or algorithm that replicates itself over a computer network and usually performs malicious actions.

firewalls[6], and other computer security technologies. But this research is largely focused on existing threats and near-term profits. This study illustrates the need for research that looks at the mid- and long-term requirements of cyber-attack investigators.

Cyber attacks are defined for this study as computer attacks that can undermine the confidentiality, integrity, or availability of a computer or information resident on it. When a cyber attack occurs it is not clear who is behind it, what the purpose is, what the scope of the attack is (whether it affects one computer or many), how the attack occurred, or where it came from. Law enforcement agencies and system administrators must respond to the attack without any of the initial information often available during a "physical" act of terrorism or other crime. Government agencies responsible for investigating and/or responding to a computer attack must treat many cases as potentially serious and possibly the act of someone acting for political ends (whether an individual, nation-state, or non-state actor) unless and until information is gathered that suggests otherwise.

Laws, regulations, treaties, and other policy instruments have not evolved to match the new realities facing cyber-attack investigators. Budget and training limitations are significant impediments in a discipline where cutting-edge skills and equipment are essential. The rate of scientific advances shows no sign of slowing. Therefore, the struggle to stay technologically up-to-date promises to become a permanent feature of the law enforcement landscape.

1.2 OBJECT OF THE STUDY

This study addresses the following question:

> ***What are the technological impediments facing law enforcement when investigating and responding to cyber attacks, for which research and development might provide solutions?***

Data was collected from federal, state, and local (FSL) law enforcement organizations in the United States. The research targeted supervisory and operational law enforcement practitioners in investigative, forensic, prosecutorial, and training capacities.

1.3 STUDY METHODOLOGY

Multiple data-collection methods were employed to ensure the reliability of the sample population and the validity of the data, and to determine that sufficient information was available for analysis. For the purposes of this study, a "needs assessment" is defined as the process of identifying and evaluating requirements of a defined population. The "identification of requirements" is defined as a process of capturing, analyzing, and describing a target population's problems in performing its assigned mission(s). The ISTS law enforcement needs assessment was conducted over six months in five major phases:

[6] Firewall: A system designed to prevent unauthorized access to or from a private network. Firewalls can be implemented in hardware, software, or a combination of both.

Phase 1. Survey Development

- During the formative stage of this study, a literature review was undertaken to identify relevant studies and reports. No similar current research was discovered during the review of the resulting document collection. ISTS staff and an independent statistician designed the survey mechanism in close consultation with experienced current and former cyber-attack investigators. The RAND Corporation was selected to conduct the survey. The survey was reviewed and edited by the RAND Survey Research Group and pilot-tested by cyber-attack investigators across the country.

Phase 2. National Survey

- Primary data was collected through a FSL law enforcement survey conducted under ISTS auspices by RAND over four months. A web-based survey was used as the primary collection device. Invitations to participate and follow-ups were conducted via United States mail, FedEx, e-mail, and telephone calls. Invitations were sent to all levels of law enforcement including, but not limited to, the National Infrastructure Protection Center, Federal Bureau of Investigation, United States Secret Service, United States Attorneys Offices, Department of Defense Criminal Investigative Service, Office of Inspector General at the National Aeronautics and Space Administration, and numerous state and local law enforcement agencies.
- Out of the 311 individuals validated to participate in the survey, 151 completed it — a response rate of 48.5%. Unless otherwise noted, the number of respondents for the survey data presented in this report is 151. On average, the respondents had worked on 81.2 computer-related cases in the last three years. Many respondents are part of high-tech crime[7] units and worked on several types of cases including fraud, extortion, and child pornography, which are not considered cyber attacks for the purposes of this report. During the survey, participants were directed to limit their responses to their experience investigating cyber attacks.
- On average, respondents investigated 15 cyber-attack cases in the last three years. A majority of the population had one to four years of cyber-attack investigative experience. An additional 25% had five or more years of experience, while 23% had less than one year of experience. On average, 50% of respondents indicated they were in a supervisory role.

[7] High-tech crime, computer crime, cyber crime, electronic crime: These terms are often used interchangeably to describe crimes including but not limited to fraud, theft, forgery, child pornography or exploitation, stalking, traditional white-collar crimes, privacy violations, illegal drug transactions, espionage, computer intrusions, or any other offenses that occur in an electronic environment for the express purpose of economic gain or with the intent to destroy or otherwise inflict harm on another person or institution. This definition is found in the National Institute of Justice report titled *Electronic Crime Needs Assessment for State and Local Law Enforcement* available at <www.ojp.usdoj.gov/nij/pubs-sum/186276.htm>.

Almost all survey participants (93%) received training for cyber-attack investigations.

Phase 3. Law Enforcement Site Visits

- ISTS researchers visited 12 FSL law enforcement agencies in seven states and the District of Columbia to conduct in-depth interviews with cyber-attack investigators. One additional set of interviews was conducted via telephone. In total, ISTS staff interviewed 39 investigators and prosecutors during this stage of the study.

Phase 4. Workshop

- During a two-day workshop, ISTS and RAND presented the data collected from the survey to a select group of 23 present and former cyber-attack investigators and prosecutors for validation, and to collect further data for analysis and prioritization.

Phase 5. Final Report Production

- ISTS staff created the final report by synthesizing and analyzing the data collected in Phases 2 through 4. A draft copy of the report was made available to a broad array of law enforcement and industry cyber-attack experts for review and comment. The feedback was reviewed and integrated into the final version of the study.

1.4 ORGANIZATION OF THE REPORT

The *Law Enforcement Tools and Technologies for Investigating Cyber Attacks: A National Needs Assessment* was designed as an inquiry into the technological impediments law enforcement encounters when investigating cyber attacks. The data presented in this study reports the needs of cyber-attack investigators as they perceive them. It is important to note that in some cases the issues discovered during the course of this study may be addressed by existing technologies.[8]

Throughout the study participants were presented question sets for several topic areas pertaining to cyber-attack investigations and training. In addition, study participants were asked to provide details on emerging investigative issues and law-enforcement-specific requirements for tools and technologies developed as a result of their input.

[8] For example, research staff for this report discovered that many agencies were using home-grown tools to overcome specific technological impediments. These tools are not widely disseminated. Individuals and organizations with tools and technologies to solve the issues reported in this document are encouraged to report their solutions at <www.ists.dartmouth.edu/lep>.

This document presents study data and analysis in the following topic areas:

1. Investigative Process: Preliminary Investigation and Data Collection
2. Investigative Process: Log Analysis
3. Investigative Process: IP Tracing and Real-time Interception
4. Emerging Technologies Requiring Research and Development
5. National Data and Information Sharing
6. Law-Enforcement-Specific Development Issues
7. Training

Each of the major topic areas is introduced with a *Background* narrative. This section is intended as a high-level overview of the key issues facing law enforcement. The heading *Findings and Analysis* indicates the presentation and discussion of both quantitative and qualitative data collected during all phases of the study. Figures and tables where total percentages exceed 100% are a result of survey participants' ability to choose multiple answers.

2. THE INVESTIGATIVE PROCESS: PRELIMINARY INVESTIGATION AND DATA COLLECTION

Background

The preliminary stages of a cyber-attack investigation are in many ways the most critical, since the timely collection and preservation of data may ultimately determine the success of the investigation. Cyber-attack investigations often involve multiple victims. The process for tracing an attacker's activities in a network is often repeated at multiple locations.

A typical investigation begins when a victim files a complaint with a law enforcement agency. After the incident is reported, the investigator(s) assigned to the case will contact the victim and usually speak directly to the parties responsible for computer security. The individuals responsible for assisting investigators are often the network system administrator(s)[9], in-house investigator(s), or outside security consultant(s).

Like any crime scene, a compromised network may be full of potential evidence. Following the discovery of an attack, a prepared victim is more likely to preserve evidence, while an unprepared victim may unwittingly destroy it. When basic computer security practices — such as enabling and maintaining adequate log files — are not employed, the case investigator may have limited options available to discover and pursue the attacker. Significant proactive efforts are underway to educate network administrators and provide them with the tools to report cyber attacks and preserve evidence.[10] In general, however, law enforcement can only influence what happens after an investigation begins, placing increased importance on technologies that ameliorate problems encountered during the investigation.

Computer attacks rarely focus on a single computer. Instead, attackers will use unauthorized access to one machine to test and exploit the vulnerabilities in other computers. By exploiting weaknesses specific to particular operating systems[11] or software, attackers can hop from one machine to another, elevating their access levels from a simple user to a super user with extensive network privileges. An investigator will collect information on which machines were compromised, in what way they were compromised, and how the attacker gained initial access. Investigators often build a catalogue of all operating systems, network services, and protocols in use to determine the extent and sequence of the attack. Investigators develop and use a map of the network

[9] System administrator: The person(s) responsible for the administration and often the security of a computer network(s).

[10] For example, *CIO Cyberthreat Response and Reporting Guidelines* jointly developed by the Federal Bureau of Investigation, the United States Secret Service, and private-sector chief information officers. This document, released by *CIO Magazine* on February 12, 2002, is available at <www.cio.com/research/security/incident_response.pdf>.

[11] Operating system: The collection of software responsible for booting a computer and providing basic libraries and tools for using the machine. Operating systems perform basic tasks, such as recognizing input from the keyboard, sending output to the display screen, keeping track of files and directories on the disk, and controlling peripheral devices such as disk drives and printers.

topology[12] to reveal patterns and links between compromised machines. Network topology maps are often created manually.

Once investigators have determined which machines have been compromised, the digital evidence recovery process begins. At a minimum, network log files from the compromised machines must be retrieved and analyzed to determine the origin of the attack. Digital evidence is provided in a number of formats such as text files, binary[13] files, and paper print-outs. This data may be transferred by e-mail, CD, magnetic tape, zip disk, jazz disk, or floppy disk.

In addition to deleting or altering system logs, an attacker may leave harmful or deceptive programs on compromised machines. Viruses and logic bombs may be set to destroy evidence or to damage the system. Root kits[14] may be deployed to disguise the presence of an unauthorized user's activity and provide utilities for exploiting other computers. Trojan horse[15] programs might be utilized to give an attacker remote surveillance or control capability.

2.1 COLLECTION OF DATA FROM MULTIPLE OPERATING SYSTEMS

Findings and Analysis

Networks are commonly constructed from computers utilizing several different operating systems to perform different tasks. Investigators often collect data from several computers to understand how a network was compromised. A basic understanding of non-Windows operating systems is required during this stage. When polled on the types of operating systems encountered in their investigations, participants indicated that the Windows operating systems dominated their caseloads (see Table 1, page 16). UNIX and Linux operating systems were encountered less frequently, but were still found in a significant number of cases. Mac OS (through version 9) and Mac OSX were seen the least during the last three years, but still on occasion by some investigators. Investigators want solutions that can automate the collection of data from multiple operating systems. Participants also articulated the need for solutions to identify and report system configurations and file locations.

Once the data is collected, investigators need tools that will help them to analyze the attack data across multiple platforms, regardless of the platform that the investigator is working on. Tools that can automate parts or all of this initial process may contribute significantly to an investigator's ability to spend investigative time focused on analysis rather than collection.

[12] Network topology: The layout and organization of a computer network. This includes the inter-connectivity of cables, routers, hubs, switches, firewalls, and computers.

[13] Binary: A number system that has just two unique digits, "0" and "1." Decimal number systems use ten unique digits, "0" through "9." All other numbers are then formed by combining these ten digits. Computers are based on the binary numbering system. All operations that are possible in the decimal system (addition, subtraction, multiplication, division) are equally possible in the binary system.

[14] Root kit: A collection of utilities that an attacker may download to a compromised computer to further or hide malicious activity.

[15] Trojan, Trojan horse: A malicious program that arrives at a host computer disguised as, or hidden inside, a benign program or file.

Table 1: How frequently have you seen each of the following operating systems in your investigations?[†]

	Very often	Often	Occasionally	Seldom	Never	Can't remember
UNIX (any flavor)	11%	20%	28%	18%	21%	2%
Linux (any distribution)	9%	15%	25%	24%	24%	3%
Windows 95/98/ME	42%	22%	14%	11%	10%	1%
Windows NT/2000	25%	37%	25%	5%	5%	3%
Mac (through OS version 9)	1%	0%	14%	32%	47%	6%
Mac OSX	0%	0%	7%	17%	67%	9%
Other	0%	1%	2%	5%	87%	4%

Note: Values are rounded to the nearest integer.
[†] In the last three years.

2.2 MAPPING NETWORK TOPOLOGY

Findings and Analysis

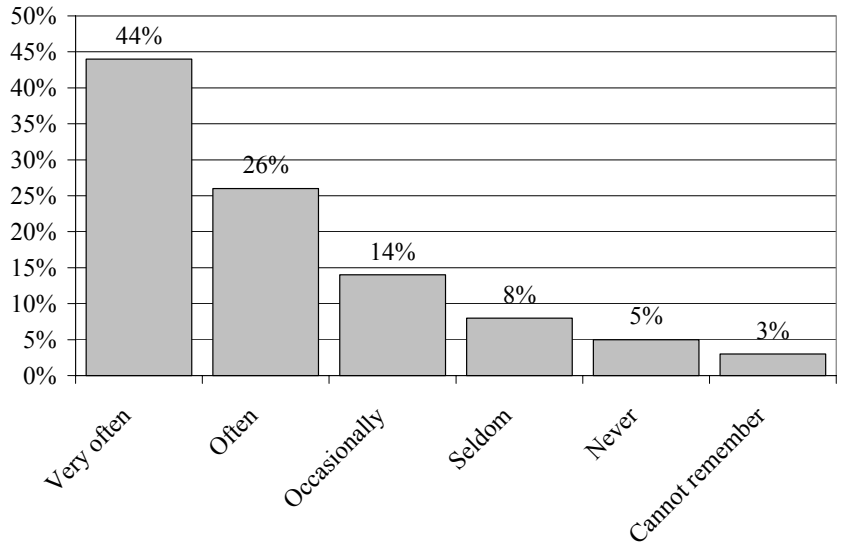
Investigators need to quickly and accurately map the victim's network during the preliminary stage of a cyber-attack investigation to assess the extent of the attack. Currently, investigators are often forced to perform this task manually. This point was articulated by a workshop participant, who remarked:

> I would say we probably had to [map the network topology] a good 90 to 95% of the time. … We ask the network administrator, "Hey, can you give us a site map of the topology of the department?" … He basically got out his pen and a piece of paper, drew two boxes that said Computer 1, Computer 2, drew a line, and said, "Here's your network diagram." I said, "I don't think that's going to do it."

Figure 1 (page 17) indicates that the majority of investigators would have found it useful to have a map of the local network topology available during cyber-attack investigations in the last three years. Investigators need network mapping software that can detect settings, recognize hardware, and graphically represent the results. Investigators commented that automated network mapping would be extremely valuable for understanding how the attack unfolded, in addition to speeding up the investigative process. Investigators related that information about firewalls, routers[16], and network addressing would significantly enhance their understanding of the victim's network topology. Automated collection of topology data such as the location and settings of switches or hubs would also assist in determining the future placement of legally authorized data-collection mechanisms. The ability to graphically represent network mapping results would enhance investigators' understanding of the complex relationships in the victim's network.

[16] Router: A device that determines the next network point to which a packet should be forwarded toward its destination.

Figure 1: How often would you have found it useful to have available a map of the local network topology?[†]



Note: Values are rounded to the nearest integer.
[†] In the last three years.

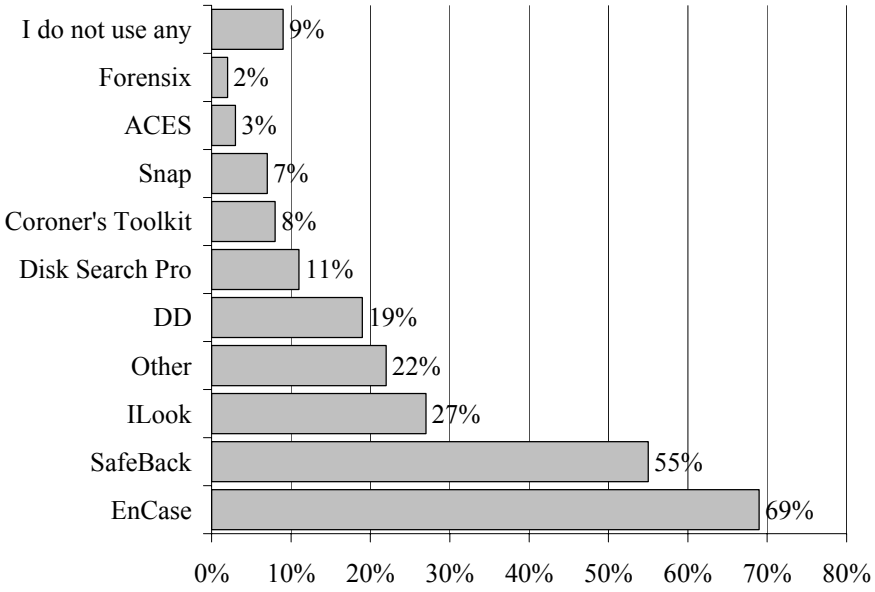
2.3 DIGITAL EVIDENCE RECOVERY

Findings and Analysis

Respondents are generally dissatisfied with existing data recovery tools for cyber-attack investigators. Analysis of the survey responses indicates a lack of essential features — or a lack of tools altogether — as significant impediments. The cyber-attack investigators surveyed were required to personally examine a compromised machine or network in almost all of their cases during the last three years (90%). Survey participants were asked to indicate the tools they used. As exhibited in Figure 2 (page 18), the majority (69%) used EnCase. SafeBack was also commonly employed, with 55% of the respondents indicating its use. Coroner's Toolkit, ACES, and Snap had each been used by less than 10% of the respondents.

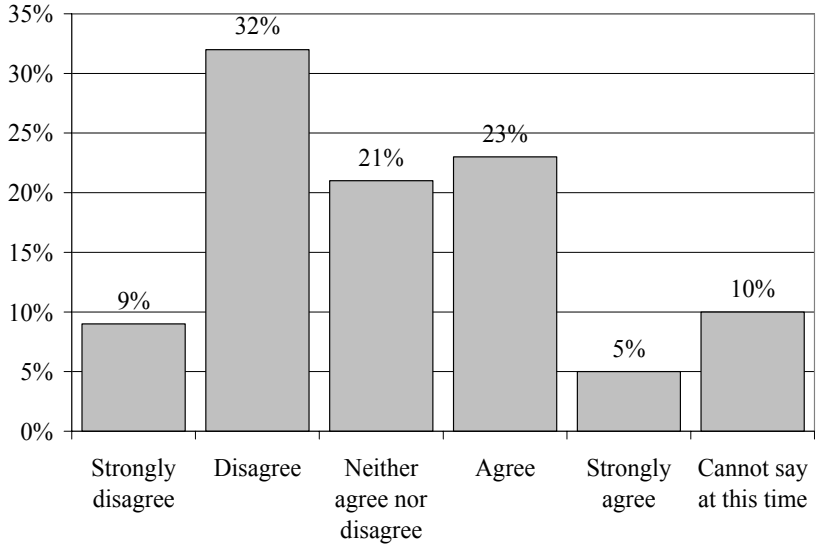
The survey probed whether or not respondents were satisfied with the examination tools available to them. Figure 3 (page 18) reveals that 41% of the respondents were dissatisfied with the tools at their disposal. As seen in Figure 4 (page 19), investigators who expressed dissatisfaction indicated that available tools are not specific to law enforcement (22%), lack essential features (37%), are too expensive (49%), require too much training (26%), or simply that no tools are available (30%). ISTS researchers queried investigators further about these issues in the site visits and workshop phases of the needs assessment. A significant problem was uncovered that may be addressed by technological solutions.

Figure 2: Which of the following software tools have you used for the task of examining a compromised machine or network?†



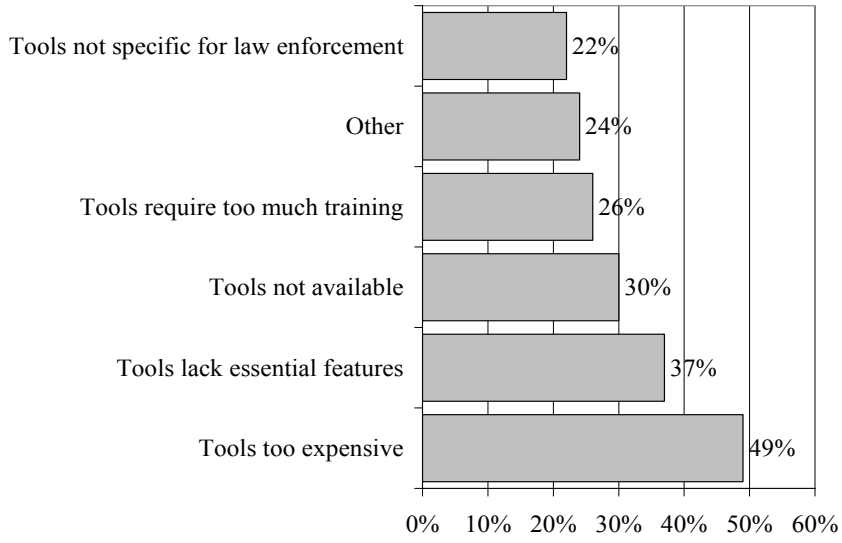
Note: Values are rounded to the nearest integer. Number of respondents = 117; excludes respondents who were not required to personally examine the compromised machine(s) or network (33) and respondents with missing data (1).
† In the last three years.

Figure 3: In general, I am completely satisfied with the software tools I have available for the task of examining a compromised machine or network.



Note: Values are rounded to the nearest integer.

Figure 4: If you are less than completely satisfied with the software tools you have available for the task of examining a compromised machine or network, please indicate the reason(s) for your dissatisfaction:



Note: Values are rounded to the nearest integer. Number of respondents = 129; excludes respondents who strongly agreed that they were completely satisfied with the software tools available for the task of examining a compromised machine or network (7) or who cannot say at this time (15).

Investigators requested technological solutions that help automate the collection of data often provided by the system administrator. In some circumstances, the system administrator may not be working in the best interest of the investigation or may not have the time or the resources to provide adequate or comprehensive information. One investigator commented at length:

> The problem is that the information they give us about those logs may not be sufficient for us to understand what they mean. For example, if a particular domain we're looking at has multiple servers[17], what server did this log come from? Or perhaps the log doesn't include a date and time. Quite often, the log files will be huge, and they don't want to print it, so they find a section that they think is important to the case [and] they'll cut and paste it someplace else and print it. A piece out of a log! We have no idea what machine, no context. So we need always to develop that information. Secondly, we need to know who captured the logs, when, what happened before, and who's going to testify to the validity of those logs in that system. It isn't just "Do we have the logs?" it's "How did we get the logs?" It's been a fairly common scenario that … someone in the

[17] Server: A system or program that provides network service such as disk storage and file transfer.

> IT department had some involvement in the intrusion or their friends may be [involved].

Investigators must quickly become familiar with a compromised network to capture relevant attack data. This knowledge transfer is often facilitated by the system administrator. Study participants noted that when they are assisted by a knowledgeable network administrator there are fewer technological problems. Conversely, investigators pointed out several problems when the system administrator provides incomplete information. For example, log files may be provided as spreadsheet documents produced after an interpretation by the system administrator. Many times the interpretations contain human transposition errors.

In other cases an insider is a suspect in the intrusion. Study participants articulated a need to alleviate their dependence on an insider altogether in this circumstance. Solutions to this obstacle would be of value to cyber-attack investigators.

2.4 CAPTURING RESIDENT MEMORY DATA

Findings and Analysis

A computer uses its random access memory (RAM)[18] to store temporary data. Passwords, encryption[19] and decryption keys, and transcripts from online chat sessions that reside in RAM are lost if the computer is rebooted or shut down. Many respondents reported that they do not have solutions to capture this data. In addition, needs assessment workshop participants discussed emerging threats that exploit a computer's RAM. For example, RAM-resident worms do not leave any traces on a local hard drive but can cause significant harm until power is removed from the system. Although the number of cyber-attack investigators conducting analysis of RAM-resident worms is probably small, the technology developed to capture resident memory data may have wider applications to computer forensics in general.

[18] Random access memory (RAM): A type of computer memory that can be accessed randomly. A single byte (a unit of storage capable of holding a single character) in RAM can be accessed without touching the preceding bytes. RAM is the most common type of memory found in computers and other devices, such as printers.

[19] Encryption: The process of changing data into an unintelligible code. Decryption is the process of changing encrypted data back to its original state. There are two main types of encryption: asymmetric encryption (also called public-key encryption) and symmetric encryption. Asymmetric encryption uses separate but related keys (sequences of digital data) to encrypt or decrypt data. Symmetric encryption uses the same key to encrypt and decrypt data.

2.5 ANALYZING EXCESSIVELY LARGE MEDIA STORAGE DEVICES

Findings and Analysis

Working with very large data sets is a concern for the law enforcement officials surveyed during this study. Table 2 presents data indicating that 87% of the survey respondents feel there is an urgent need to develop tools to address extremely large disk drives. Survey participants commented that an investigator may have to perform multiple scans and analyses of compromised machines. Large-capacity hard drives require long periods of time to examine and often require multiple machines in computer forensic labs. Usually one forensic computer is used for acquiring and duplicating data, a second is used to conduct character string searches, and a third is used to analyze the search results from another hard drive. Investigators commented that the process takes too long. Forensic software needs to be designed to search large volumes of data more efficiently.

Table 2: Please indicate the urgency of the need in your department or agency to develop tools to address extremely large disk drives.

Not at all urgent	Moderately urgent	Extremely urgent	Don't know
7%	42%	45%	5%

Note: Values are rounded to the nearest integer.

3. THE INVESTIGATIVE PROCESS: LOG ANALYSIS

Background

After preliminary case data has been retrieved, investigators are required to devote significant analytical time to deconstructing log and other files to advance the case. System administrators usually maintain audit trails for normal operational tasks, but this may not be comprehensive enough for the successful investigation of a cyber attack. Setting the audit level to capture the maximum amount of activity significantly increases the chance of capturing, discovering, and tracking unauthorized activity. The drawbacks of increasing the audit level are increased storage space requirements and analysis time by company employees. In the current environment, most network owners do not maintain high audit levels. This means that when investigators receive log files, they usually contain limited data. Cyber attackers may erase or modify log files to mask malicious activity. The log analysis process, often done manually, is time-consuming and requires investigators with expertise in investigating cyber attacks. Relationships and attack characteristics are often difficult to discern in character- and text-based log files.

3.1 LOG COMPILATION

Findings and Analysis

As listed in Table 3, logs were given to investigators most often as text files. An events database[20] was encountered less often in the last three years.

Table 3: Please indicate how often you received necessary log files in each of the following informational formats during your investigations.†

	Very often	Often	Occasionally	Seldom	Never	Can't remember
Raw text file	19%	28%	21%	8%	15%	9%
Formatted text file	8%	31%	26%	12%	16%	7%
Events database	5%	15%	21%	20%	31%	7%
Other	1%	1%	0%	0%	93%	6%

Note: Values are rounded to the nearest integer.
† In the last three years.

The following considerations are salient to developers of law-enforcement-specific technological solutions for log-data compilation:

Recognize and Import Preliminary Investigation Data

As detailed in "The Investigative Process: Preliminary Investigation and Data Collection" (Section 2, page 14), investigators require detailed information from a compromised network to track a cyber attacker. Information must be collected and organized from

[20] Events database: A formatted file for storing system events in records for later query.

multiple data points in various formats. For example, investigators require information on operating system configurations in addition to log files. Simple text files may be created listing the names of individuals who use certain IP addresses. In some cases, investigators receive blocks of formatted information such as an events database. Technological solutions to compile data from the preliminary investigation must be flexible enough to import and organize input from the variety of data sources and file formats collected over the course of an investigation.

Recognize and Import Logs Across a Network

Most networks today are not based on a single platform or operating system, but are instead a collection of heterogeneous computing[21] platforms. For example, a network may comprise a Solaris-based system for the web and database server; UNIX, Linux, and Windows 98 for programming and development uses; and Macintosh for the computing needs of the office staff. Cyber attackers use weaknesses in each system to their advantage. Security holes in the firewall may allow an attacker to compromise a system's web server and from there gain access to the office computers. In this situation, logs from the firewall, web server, and the office machines are examined in order to track the intruder. As illustrated in this example, computers are configured with different operating systems, logging options, and file formats. Investigators encounter problems aggregating and analyzing these files. The lack of cross-platform investigative solutions for log analysis strongly points to the need for the development of new tools, approaches, and technologies. These solutions should search the network for logs, collect them regardless of platform, and prepare them for export to a different operating system or analysis environment. Law enforcement investigators endorsed automating this process in a compiler function. Study participants articulated the need for software wizards or other import utilities to recognize common operating systems' log files. Technological solutions with this functionality would significantly cut down the time investigators currently spend merging, aggregating, and transforming disparate data files.

Reconstruct Altered or Damaged Logs

A skilled computer intruder will take steps to avoid detection, such as altering log or system files to mislead a system administrator or defeat an intrusion detection system (IDS).[22] A less-sophisticated attacker will simply erase log files. In cases where log files are lost or corrupted, several of the respondents indicated a need for solutions to reconstruct logs or search for fragmentary digital evidence. Law enforcement investigators volunteered a possible solution for this problem: a nationally accessible repository of log file structures for multiple operating systems. This would provide an investigative resource for reconstructing digital fragments into a readable format for importation into log analysis solutions.

[21] Heterogeneous computing: The use of dissimilar computing hardware, software, and operating systems.

[22] Intrusion Detection System (IDS): A program or hardware device that monitors activity, security logs, or audit data on a computer or network in an attempt to detect an intrusion.

Similar issues are discussed in "National Information Sharing" (Section 6, page 40). This functionality may be integrated into software designed to compile log files for subsequent analysis.

Place Log Data into an Organized Timeline

Following the path of an intruder through a compromised system is critical to assessing the extent of the attack. As logs are generated, individual machines stamp network event times based on either their internal clocks or a network-based clock. These time stamps differ when networked computers' clocks are not synchronized. Needs assessment participants pointed out the obstacle that unsynchronized time stamps present during the analysis process. In order to build a case for prosecution, investigators want to be able to go to compromised systems and capture the clock settings. This allows the clock times to be compared and log files to be placed on a common timeline. By placing information that has been converted to a common time frame on a timeline, the activities and intent of the intruder may be more clearly identifiable.

Cyber-attack investigators indicated a need for a technological solution to automatically capture the individual time and date settings from compromised network computers. This task is often done manually. Study participants revealed that translating log files from multiple time zones to a common time frame is also often done manually. The automation of this process would significantly decrease the amount of time law enforcement investigators expend reconstructing attackers' methodologies.

Organize Output to a Common and Portable Format

Once log data has been aggregated and transformed into a common data set, investigators often need to export it to other software programs and transmit it to other law enforcement agencies. Participants indicated that no common format or standard is available for storing and sharing log files or analysis products. Without common data-sharing protocols, it is difficult for investigators to share information, even within their own organizations. Additionally, very large data sets of log files are cumbersome to manage and transmit. Data sets optimized for analysis and portability would provide law enforcement with a common platform for sharing and presenting log data. Common data sets could be compressed into a manageable size and imported and exported into common software available to law enforcement nationally.

The issues surrounding the creation of common data sets are part of a larger technological impediment facing law enforcement. The information revolution was largely made possible by the widespread adoption of interoperability standards for software and hardware. These protocols make it possible for heterogeneous computing platforms and software applications to communicate. The law enforcement community has not adopted standards for sharing cyber-attack information. This lack of interoperability hampers cyber-attack investigations when the sharing of information is critical. The law enforcement community involved in this survey supports widely adopted standard file formats, middleware[23], and protocols to address the specific needs of law enforcement.[24]

[23] Middleware: Software that connects two otherwise separate applications.

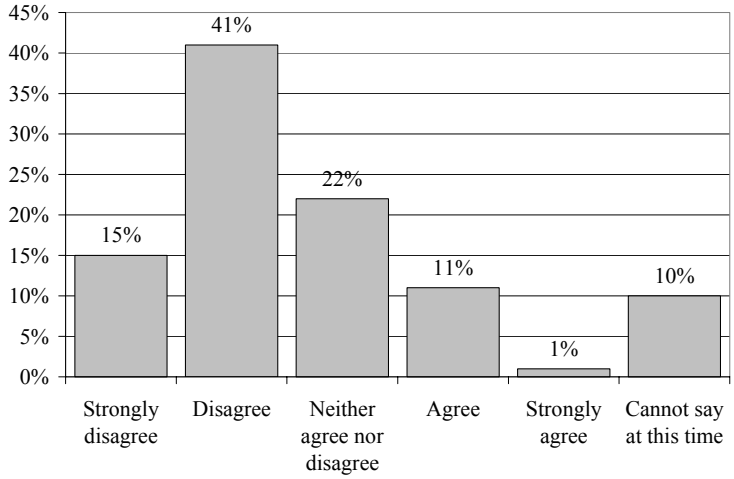
The national discussion, development, and adoption of common data-sharing and communications protocols would greatly improve the ability to share information across jurisdictions.

3.2 LOG ANALYSIS AND REPORTING

Findings and Analysis

Log file analysis is a critical, complex, and time-consuming task requiring investigation by individuals with significant training and experience. Survey respondents spent 23% of their time, in a typical investigation, interpreting and analyzing log files. Figure 5

Figure 5: In general, I am completely satisfied with the software tools I have available for the task of interpreting and analyzing log files.

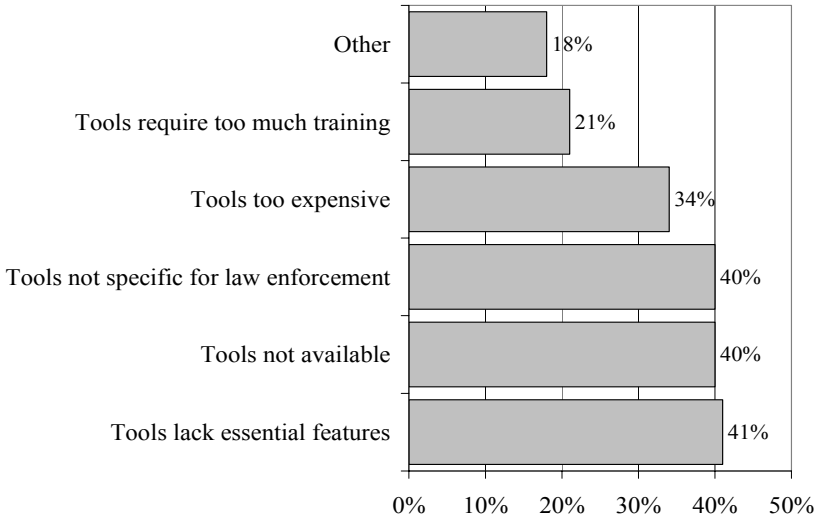


Note: Values are rounded to the nearest integer. Number of respondents = 73; excludes respondents who did not use software tools for interpreting and analyzing log files (78).

illustrates that over half (56%) of the respondents were dissatisfied with the tools they have available for interpreting and analyzing log files. In Figure 6 (page 26), those respondents dissatisfied with log analysis tools cited such issues as tools lack essential features (41%), tools not specific to law enforcement (40%), and no tools available (40%). Table 4 (page 26) shows more than half of the respondents indicated that they have difficulty filtering out irrelevant information (59%), managing large sets of log files (56%), and synthesizing information from multiple log files (53%).

[24] A national initiative toward this end, referenced by study participants, is called InfoTech. This research, part of the Advanced Generation of Interoperability for Law Enforcement (AGILE) Program at the National Institute of Justice, is working toward enabling technologies for information sharing between law enforcement agencies. Additional information can be found at <www.agileprogram.org/research/infotech.html>.

Figure 6: If you are less than completely satisfied with the software tools you have available for the task of interpreting and analyzing log files, please indicate the reason(s) for your dissatisfaction:



Note: Values are rounded to the nearest integer. Number of respondents = 65; excludes respondents who did not use software tools for the task of interpreting and analyzing log (78), who strongly agreed that they were completely satisfied with the software tools for the task of interpreting and analyzing log files (1), or who cannot say at this time (7).

Table 4: How frequently have you encountered each of the following obstacles in carrying out the task of analyzing log files?[†]

	Very often	Often	Occasionally	Seldom	Never	Can't remember
Difficulty filtering out irrelevant information	30%	29%	21%	7%	9%	5%
Difficulty managing large sets of log files	29%	27%	23%	5%	12%	3%
Difficulty synthesizing information from multiple log files	25%	28%	18%	7%	15%	7%

Note: Values are rounded to the nearest integer. Number of respondents = 150.
[†] In the last three years.

The following exchange is representative of cyber-attack investigators' perception of this problem:

> **Interviewer:** "Please describe the three most significant technological obstacles you face in the successful investigation of cyber attacks."
>
> **Interviewee:** "Analysis of logs, analysis of logs, analysis of logs."

Log analysis is usually done manually and accounts for, on average, one quarter of the time spent on an investigation. Many investigators recommended developing technological solutions that can automate log analysis to shorten the time it takes to track a cyber attacker. The following considerations are salient to developers of law-enforcement-specific technological solutions for log analysis and reporting:

Automate Log File Analysis

A network intrusion case may produce very large data sets of log files. These need to be extensively examined to determine the intrusion methodology and damage caused by a cyber attacker. As one investigator suggests:

> The task of log analysis needs to be automated or more manpower put into the task. In all successful cases that I am aware of, the case progressed very rapidly. The logs need to be reviewed rapidly in order to identify suspect IP addresses and then get the proper legal process served, before the [Internet service providers (ISPs)[25]] destroy their records. Cases that involve extensive amounts of log analysis usually languish unless they have adequate personnel to trudge through the logs quickly or have a very dedicated agent to do the review quickly in long hours over a very few days.

Many features were emphasized by law enforcement as important for log analysis solutions. Easy-to-use search functions are a basic need of cyber-attack investigators. Analytical tools that discover anomalies in large log files would be of great value to investigators who are currently undertaking this analysis manually. If common data sets are developed using relational database structures, powerful analytical applications may be utilized to reveal hidden relationships.

Develop Graphical Reporting

Many of the data files gathered and produced during the investigative process are in the form of detailed lists of events, network connectivity descriptions, and other character- or number-based technical information. Study participants expressed a clear desire for detailed technical data to be represented in a graphical format. Investigators emphasized that graphical representation of complex data sets would allow them to determine hidden relationships not apparent in character- or number-based outputs.

Graphical reporting was also mentioned by prosecutors as a valuable tool for the presentation of complex cyber-attack data in the courtroom. Participants commented that many individuals with technical expertise are unable to relate the issues to non-experts in such a manner that it can be presented in a courtroom. Prosecutors interviewed for this study suggest that the presentation tools for the courtroom should include a graphical reporting function.

[25] Internet service provider (ISP): A company that offers access to the Internet as a service to individuals or organizations via dial-up telephone lines or direct network connections. ISPs may be local, regional, or national.

4. THE INVESTIGATIVE PROCESS: IP TRACING AND REAL-TIME INTERCEPTION

Background

Many obstacles face cyber-attack investigators when they are tracing an IP address. The broadly adopted Internet protocol currently in use allows attackers to spoof[26] source IP information, resulting in investigative dead ends. Even if an IP address has not been spoofed, the attack may have been launched from a public access machine, limiting investigative options. Many Internet service providers do not keep current information on users. Furthermore, ISPs may not keep or maintain access logs which would show the use of an IP address by an account holder.

If IP address registration yields a valid physical address outside an agency's jurisdiction, the investigator faces additional impediments. To further the case, the investigating agency must often contact a law enforcement agency where the physical address is located. Identifying and involving other law enforcement agencies can be time-consuming. Cyber attackers perpetrating crimes from international locations present significant additional challenges. Existing legal processes and protocols are reported to entail unacceptable delays in communications, language barriers, and non-equivalent international criminal justice systems. The borderless nature of cyber attacks is a considerable obstacle hampering an investigator's ability to further a case.

If an investigation begins while an attack is ongoing, law enforcement may apply to use legally authorized surveillance. The use of publicly available counter-surveillance technologies by cyber attackers impedes investigators during this task. For example, the use of encryption and other data-hiding techniques by suspects can limit the usefulness of intercepted data.

4.1 IP TRACING

Findings and Analysis

As itemized in Table 5 (page 29), cyber attackers use spoofing, anonymizers[27], and publicly accessible machines to evade detection. Spoofed packets[28] used in DDoS[29] and other attack methodologies present significant obstacles to investigators. In these attacks, data is maliciously sent or directed to overwhelm available computing resources in a victim's network. This makes the system unavailable to legitimate users. By replacing the real IP source address with a false one, the origin and identity of the attacker can be disguised. Since the attacker only sends packets to expend the resources of this victim and does not receive any data in return, it is extremely difficult to locate an individual

[26] Spoof: In networking the term is used to describe a variety of ways in which the origin of data can be obscured.

[27] Anonymizer: A computer that acts as a proxy for Internet requests, hiding the identifying information of the requester from the service provider.

[28] Packet: A unit of data that is routed between an origin and a destination on the Internet.

[29] Distributed Denial of Service attack (DDoS): Action(s) by distributed computers that prevent any part of another computer system from functioning in accordance with its intended purpose.

who is perpetrating an attack using IP spoofing. Once the domain of sophisticated cyber attackers, IP spoofing is now a common attack methodology employed by amateur and expert attackers alike.

Table 5: How often during your investigations have you encountered each of the following obstacles in carrying out the task of finding the entity linked to an IP address?[†]

	Very often	Often	Occasionally	Seldom	Never	Can't remember
Attacker uses IP spoofing	6%	14%	35%	21%	23%	2%
Attacker uses anonymizers	5%	13%	35%	18%	25%	3%
Attacker uses public or shared machine	15%	40%	26%	5%	12%	1%

Note: Values are rounded to the nearest integer. Number of respondents = 150.
[†] In the last three years.

There are technological solutions to limit the effectiveness of spoofing attacks; however, current strategies often are underutilized. For example, spoofed IP addresses are easy to detect and stop near their source, since routers can be programmed to discard any outbound packets whose source IP address does not belong to the router's client networks. Such outbound or "egress" filtering is a relatively simple but not widely implemented validation procedure. Likewise, inbound or "ingress" filtering of any IP packets with un-trusted source addresses, before they have a chance to enter the network, can also be effective. Untrusted source addresses include those addresses reserved for private networks or not yet issued by the international authorities that assign Internet numbers. Countermeasures for DDoS can also include cooperation from upstream ISPs that send packets to their client networks. ISP routers can be programmed to limit the rate at which packets typically associated with attacks are sent downstream to client networks. By limiting these particular packets, the effects of a malicious flood can be minimized without seriously disrupting normal operations. These preventive measures are well within the capabilities of most ISPs, but often not employed. With a single technological solution unlikely, law enforcement will continue to face significant challenges in this task area.

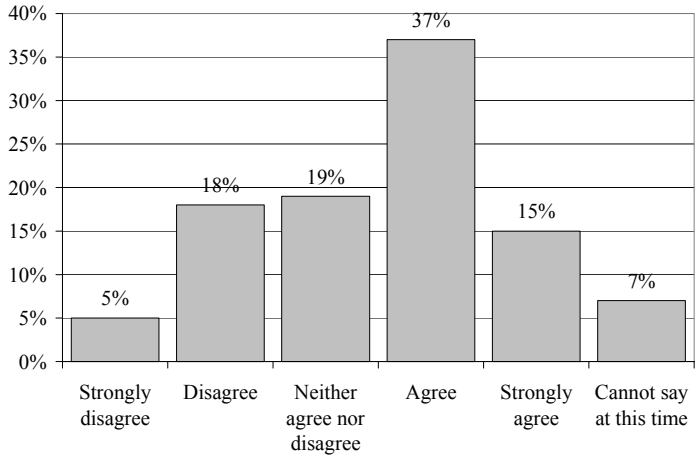
From the very beginning of this study, the non-technical issues surrounding IP tracing generated lengthy discussion. The level of concern expressed by the study population is a result of its experience with cases that could not be advanced into the final investigative stages due to the lack of network owner record keeping and the anonymous nature of many Internet access sites. Several individuals commented that they wanted ISPs to take more responsibility for the maintenance of records that could aid investigations. The

current environment allows ISPs to provide the public access to the Internet with only limited user verification or accountability. To counter this problem, law enforcement encourages ISPs to keep more complete historical logs, such as RADIUS logs.[30]

As evidenced in the log analysis discussion (page 27), investigators indicated that one of the most effective approaches to catching cyber attackers is tracking their actions and locations as quickly as possible. Law enforcement officials who provided data to this survey commented that they find the legal mechanism and information sharing between organizations, both nationally and internationally, inadequate for rapid IP tracing. Investigators indicated that in some cases weeks, months, and even years can pass before cooperation requests can be accommodated, particularly in international cases. The very limited time frames involved with capturing data to trace cyber attacks effectively hinder the successful investigation and prosecution of some incidents. Workshop participants, after discussions of the survey feedback on this data-sharing topic, pointed out the need for a database of who works on cyber-crime investigations in each jurisdiction, what experience they have, what training they have completed, and how to reach them. This issue is explored further in "Law Enforcement Cyber-Attack Contact Database" (Section 6.4, page 44).

When a cyber-attack investigation yields an IP address, investigators have a number of tools at their disposal to determine the registration and network information associated with it. In Figure 7, over half of the survey population was generally satisfied with the

Figure 7: In general, I am completely satisfied with the online tools I have available for the task of finding the entity linked to an IP address.

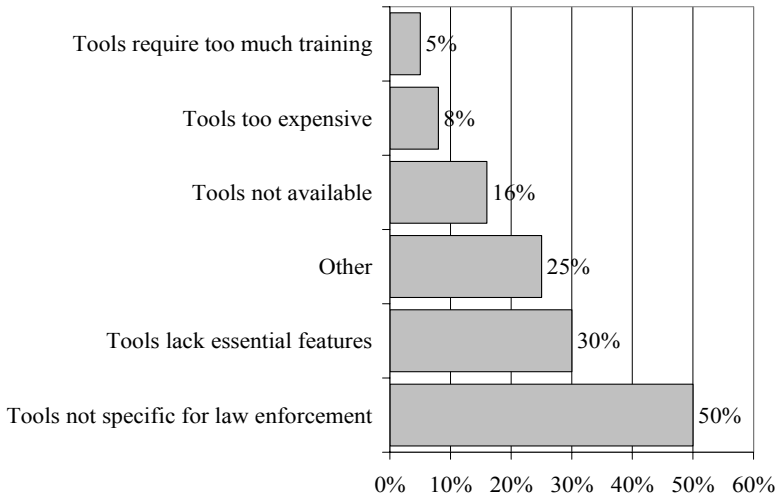


Note: Values are rounded to the nearest integer. Number of respondents = 150.

[30] RADIUS log: RADIUS is the acronym for Remote Authentication Dial-In User Service, an authentication and accounting system used by many ISPs. When users dial into an ISP they must enter their usernames and passwords. This information is passed to a

tools they had available for performing this task. In Figure 8, dissatisfaction is expressed with existing IP tracing tools because they are not specific to law enforcement (50%) and lack essential features (30%).

Figure 8: If you are less than completely satisfied with the online tools you have available for the task of finding the entity linked to an IP address, please indicate the reason(s) for your dissatisfaction.



Note: Values are rounded to the nearest integer. Number of respondents = 118; excludes respondents who strongly agreed that they were completely satisfied with the software tools for the task of finding the entity linked to an IP address (22) or who cannot say at this time (11).

4.2 REAL-TIME INTERCEPTION OF DIGITAL DATA

Findings and Analysis

The real-time interception of digital data is considered a form of wiretapping and falls under legal restrictions to prevent the unlawful interception of a person's communications. Survey respondents noted that in the last three years they had, on average, each worked on more than four cases that involved real-time interception (investigators worked an average of 15 computer-intrusion cases in the last three years). Although the survey data indicates that almost one third of an investigator's computer-intrusion caseload involved real-time interception, further analysis shows that this statistic is misleading. Cross tabulation of the survey data reveals a link between the experience and responsibility level of investigators and the use of real-time interception technology.

Investigators with less than a year of experience have each worked, on average, less than one case involving real-time interception. Investigators with one to four years of

RADIUS server, which checks that the information is correct, and then authorizes access to the ISP system. This information is recorded in the RADIUS log.

experience worked just under three cases involving real-time interception. Investigators with more than five years experience worked an average of seven cases involving real-time interception in the last three years. Supervisors worked an average of eight-and-one-half cases in the last three years; conversely, non-supervisors worked, on average, less than one case over the same time frame.

In instances where digital interception is warranted and is legally authorized, local and state law enforcement agencies have found themselves either without the necessary technology or in possession of technology that exceeds their level of training. Many participants commented that they had not conducted interceptions because they did not have the technology to do so. Several study participants indicated that the use of this technology would increase their effectiveness in combating cyber attacks. A study participant emphasized:

> [… the need for] state and local law enforcement to have software that enables them to conduct real-time court authorized intercepts. … There have been many investigations that we have identified the suspects but have been unable to develop evidence to prosecute. This technology would give us another investigative avenue that we don't have now.

Cyber-attack investigators believe that technological challenges lie in filtering large amounts of information in a data stream and then reconstructing messages from their corresponding packets. Among respondents who noted specific impediments to carrying out the task of real-time digital interception, 31% indicated that the inability to selectively monitor traffic was a common obstacle (Table 6, page 33). For example, law enforcement personnel are required to detail the nature of information they are looking for in their application to conduct legally authorized interception. If approved, the only information investigators are allowed to collect, analyze, and use in subsequent prosecution is that detailed in the application. Specific data-collection restrictions present both technological and legal impediments to using digital interception. Both the volume of data and rate of collection make it extremely difficult to extract specific information from a digital data stream. As indicated by workshop attendees, filtering the data is the issue. Digital interception is analogous to trying to take just a sip of water from a fire hose at full pressure. This dilemma points to a technological solution: an instrument for minimizing, isolating and analyzing data captured in the course of legally authorized data interception. National assessment and R&D efforts toward this end would increase law enforcement's capability to conduct legally authorized real-time digital interception of data.

Law enforcement investigators reported that instant messaging (IM)[31] and IRC[32] are commonly used by cyber attackers to plan coordinated attacks and discuss their exploits. The ability to perform legally authorized monitoring, and subsequently link these communications to a particular IP address or entity, would provide law enforcement with

[31] Instant messaging: A number of software applications that allow users to see if other individuals are actively using the same application over the Internet and to exchange messages with them. Instant messaging differs from ordinary e-mail in the immediacy of the message exchange and also makes a continued exchange simpler than sending e-mail back and forth.

[32] IRC: An acronym for Internet relay chat, a popular chat system.

additional investigative approaches during or following an attack. This issue was discussed by several investigators as being the only method by which certain attackers can be traced. However, gaining any information about an IM or IRC user is perceived as very difficult since there is no central repository of information. IM and IRC sessions often end with no trace of their existence.

Investigators commented that perpetrators often don't encrypt their IRC messages currently. This is supported by the survey results indicating that encrypted IRC is encountered only occasionally. New software is driving a trend toward encrypted IM and IRC systems that will challenge law enforcement capabilities in the future. New approaches that give law enforcement investigators additional capabilities in these task areas are perceived as a need by the participants of this study.

Table 6: How often during your investigations have you encountered each of the following obstacles in carrying out the task of real-time digital interception?[†]

	Very often	Often	Occasionally	Seldom	Never	Can't remember
Inability to selectively monitor traffic[*]	9%	22%	21%	15%	22%	10%
Inability to monitor peer-to-peer communications or instant messaging	10%	19%	20%	15%	27%	8%
Subject encrypts e-mail	2%	8%	34%	14%	34%	8%
Subject encrypts IRC or chat traffic	2%	5%	12%	27%	46%	8%

Note: Values are rounded to the nearest integer. Number of respondents = 59; excludes respondents whose cases involved no real-time digital interception and respondents with missing data (1). Item marked with an * number of respondents = 58.
[†] In the last three years.

5. EMERGING TECHNOLOGIES REQUIRING RESEARCH AND DEVELOPMENT

Background

This study queried participants about emerging technological issues that hamper their ability to successfully conduct investigations. Several of these technologies were topics within the survey mechanism. Additional information was gathered during the site visits and in the workshop sessions.

5.1 ENCRYPTION

Findings and Analysis

The ability to defeat encryption was the leading priority of the emerging technology issues addressed during this study. Information that is encrypted with robust algorithms is often impossible to decrypt regardless of available computing power or investigative resources. Based on the unacceptable time constraints involved with attempting brute-force attacks on encrypted data, the outright defeat of encryption may not be a reasonable short-term research goal[33], although it may be a primary focus of mid- to long-term research efforts.

Members of the law enforcement community expressed the need for solutions to address this obstacle. Additional focus and research must be directed toward increasing law enforcement's ability to circumvent the obstacle of encrypted data. This could include methods for detecting the presence of encrypted data and providing law enforcement with the solutions to perform legally authorized keystroke monitoring[34] to determine pass-phrases. There are some technologies which have been developed for these purposes, but they may only be available to certain law enforcement agencies[35], and some may still be developmental in nature.[36] Earlier efforts to develop methods for circumventing encryption included installing, with a search warrant, a keystroke logging device to capture passwords on-the-fly.[37] Other approaches may include collecting data that resembles encryption keys from a cyber-attacker's computer(s) and applying them to the encrypted data. Development and dissemination of solutions toward this end would provide additional investigative approaches for cyber-attack investigators.

---

[33] Encryption methods use keys, or strings of characters, to encrypt and decrypt data. A modern computer (such as a 2 GHz Pentium IV) can attempt approximately $2^{\wedge 26}$ (~67 million) keys per second. A brute-force attack on a 128-bit key will require $2^{\wedge 102}$ seconds to try all possible combinations. Even assuming that only 50% of the possible combinations would need to be attempted, $8.2*10^{\wedge 22}$ years would be required to crack a 128-bit encrypted key. $8.2*10^{\wedge 22}$ years is longer than the time scientists believe the universe has been in existence (National Research Council, Board on Physics and Astronomy, *Cosmology: A Research Briefing* available at <www.nap.edu/readingroom/books/cosmology/>). The example given here is based on 128-bit encryption. Even stronger encryption (e.g. 1,024-bit encryption) is commonly available. The hypothetical case examined here gives a concrete look at the problems that encryption poses for cyber-attack investigators.

[34] Keystroke monitor: A program that records the keystrokes typed at a computer, typically saving the keystrokes to disk for later retrieval.

[35] One example is the FBI's DCS-1000, formerly known as Carnivore (see www.fbi.gov/hq/lab/carnivore/carnivore2.htm).

[36] The "Magic Lantern" program, a remotely installed key logging device, has been described as "under development."

[37] For example, the United States v. Scarfo, 180 F. Supp. 2d 572 (2001) case.

5.2 WIRELESS TECHNOLOGIES

Findings and Analysis

Technological advances in wireless communication are presenting new challenges to the law enforcement community. Members of the needs assessment working group ranked wireless devices as their second-highest emerging technology concern behind encryption. Cyber-attack investigators indicated that wireless devices, such as cell phones, are evolving into multi-functional appliances. Integration with the Internet makes them increasingly susceptible to cyber attacks. Increased functionality has also allowed wireless technology to be used to launch cyber attacks.

Wireless sniffing, or covertly capturing data, is seen as an emerging threat that is increasingly employed by cyber attackers. Investigators noted that it is not uncommon to find the publication of unprotected wireless network addresses on hacking web sites. As the popularity and functionality of mobile devices continue to increase, study participants believe this technology will play a larger role in cyber attacks.

Determining the location of the intruder is a significant challenge during a cyber-attack investigation. Wireless technologies allow an attacker to drive into a parking lot, compromise a local network, and simply drive away when the attack is completed. Investigators are not equipped to deal with this threat. Many study participants called for the development and use of technologies for wireless networks that can recognize unauthorized access for both network owners and law enforcement investigators. Further, technological solutions to trace an attacker's physical location are required for wireless cyber-attack investigations.

Advances in technology have made satellite Internet access widely available. The advantages that satellite-based Internet connections provide attackers were raised by investigators during this study. Law enforcement investigators voiced a concern that high-speed satellite connections will allow attackers the freedom to operate nearly anywhere on the planet. The challenges and impediments concerning satellite-based Internet connections were similar to those discussed with wireless networks. These cases present significant challenges if the technology to locate the geographic location of an attacker is not readily available to law enforcement investigators. Study participants indicated that they perceive a need for mid- and long-term research into mobile communication solutions in order to increase their investigative capabilities.

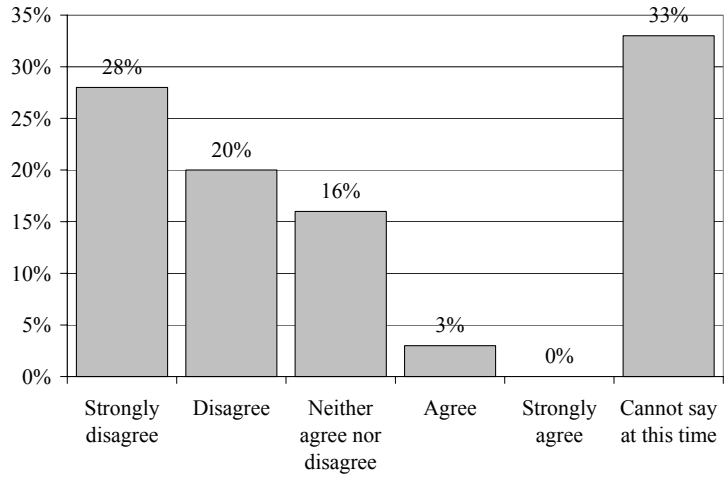
5.3 STEGANOGRAPHY

Findings and Analysis

The use of digital steganography[38] is viewed by cyber-attack investigators as an emerging technological obstacle. Commercial steganographic software programs and countless home-grown tools use any number of approaches and algorithms to hide messages or data. These methods make no perceptible change to the source file. If the presence of an embedded message can be detected, extracting the message without knowing the original algorithm can prove as difficult as decrypting an encoded message. Many steganography programs employ encryption as an added layer of security, increasing the likelihood that messages cannot be found or understood.

The survey population was asked about their experience with steganography and respondents stated that they have worked, on average, less than one case in the last three years involving this technology. Figure 9 reveals 48% of the respondents to the question expressing dissatisfaction with available tools for detecting and recovering data hidden using steganography. Of the respondents who were dissatisfied, 63% reported that tools were unavailable for the detection of data hidden by steganographic means (Figure 10, page 37). Currently, the law enforcement community believes it has not encountered steganography in many cyber-attack cases, but as one interviewee stated: "What you don't know about you can't address."

Figure 9: In general, I am completely satisfied with the tools I have available for the task of detecting and recovering data hidden using steganography.

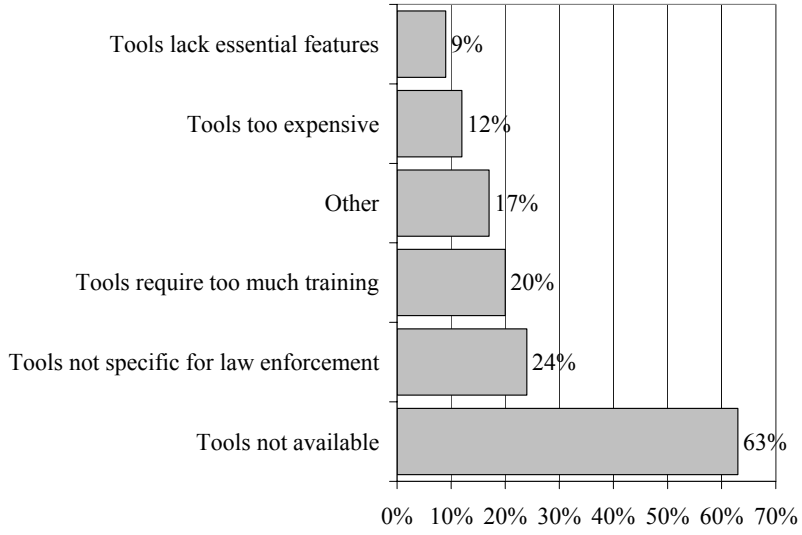


Note: Values are rounded to the nearest integer.

[38] Digital steganography: Hiding a secret message within a digital file in such a way that typical viewers cannot discern the presence or contents of the hidden message. For example, a message might be hidden within an image by changing the least significant bits that represent each pixels' color to message bits.

One investigator revealed that he looks for signatures of commercial steganography programs during forensic examinations. If found, additional steps would be taken to examine the drive for common data-hiding file types, such as image and music files. The members of the law enforcement community that participated in this survey expressed a need for a national data service containing digital steganographic programs and signatures that could be consulted during a forensic analysis of a seized computer. Furthermore, the study participants articulated a requirement for a technological solution that could, at a minimum, flag digital files that might contain steganographic messages. Several commercial programs are available to detect steganography utilizing different technological approaches; however, there is no national standard for benchmarking these tools. Law enforcement would welcome a comprehensive review of the technological solutions available for detecting the use of steganography in digital files to help gauge their effectiveness.

Figure 10: If you are less than completely satisfied with the tools you have available for the task of detecting and recovering data hidden using steganography, please indicate the reason(s) for your dissatisfaction.



Note: Values are rounded to the nearest integer. Number of respondents = 101; excludes respondents who cannot say at this time (50).

5.4 MAGNETIC MICROSCOPY

Findings and Analysis

Investigators clearly indicated a requirement for new technologies, or the adaptation of existing technologies, to recover digital evidence. One example is magnetic microscopy.[39] Study participants indicated that this technology may be particularly pertinent in those cases where a hard drive from a victim's computer was either erased or damaged by the attacker. Although they recognize that the core technology already exists for other applications, investigators believe there are non-technology barriers preventing the development and adoption of magnetic microscopy for law enforcement use as a new approach to digital data recovery. One participant commented that this technology may be too expensive for most investigations since the skills and equipment needed to perform the examination would be limited to a centralized forensic facility. Study participants indicated their support for new research that may encourage market forces to commercialize magnetic microscopy technology, making it broadly available.

5.5 FORENSIC DATA ARCHIVING

Findings and Analysis

The law enforcement community must capture and archive increasingly large data sets from computers discovered in the course of their investigations. Prior to court proceedings, investigators or examiners may issue a unique number based on the particular order of the binary information contained on a storage device. This number, based on an MD5[40] algorithm or similar technology, may be referenced in the trial, and in subsequent years during appeals, to ensure the integrity of the source data. Law enforcement investigators surveyed during the course of this study indicated that problems with this system exist due to the natural degradation of magnetic data. The unique sum of the data will be changed if any of the magnetic bits[41] are flipped from a "1" to a "0" or vice versa due to the unstable properties of the storage media. Although the substance of the data will not be altered, the unique number provided by the MD5 algorithm will change. This may call into question the integrity of the data. Many commercial data storage companies deal primarily with the storage of short-term back-up data and accept a certain level, albeit very low, of data degradation. This technology does not meet the long-term preservation requirements outlined by the law enforcement participants of this study.

Once computer files have been used in court proceedings, law enforcement may be required to preserve the evidence for a number of years. Study participants expressed concerns regarding the cost of storing excessively large amounts of data. During an

[39] Magnetic microscopy: A scientific technique by which data may be recovered from magnetic disk drives even after an overwrite has occurred. The process includes examination of the disk surface with an electron microscope for the traces of remnant patterns left on the recording medium.

[40] MD5: An algorithm (a formula or set of steps for solving a particular problem) that is used to create unique digital signatures.

[41] Bit: Short for binary digit, the smallest unit of information on a computer.

interview, one investigator stated that his team was presently working on a case where they had seized and mirrored a RAID array[42] of four racks each with 20-plus gigabyte[43] hard drives, and had nearly used up the team's storage budget for the entire year with this one case. Survey participants expressed a requirement to store evidence for anywhere from seven to 20 years. One of the many reasons for the need for long-term storage is that, while each separate case develops in its own right, over time patterns and linkages between cases may develop, pointing to broader investigations. If the evidence from each individual case is not properly preserved, the links between cases may either not be discovered or not be prosecutable. With the size of storage media continually increasing, the task of maintaining significant bodies of digital data is becoming a significant issue for investigative agencies. The participants in this study expressed the need for the assessment, research, and development of solutions to securely store very large data sets they encounter.

[42] RAID array: An acronym for Redundant Array of Independent Disks, a category of disk drives that employ two or more drives in combination for fault tolerance and performance.

[43] Gigabyte: A measure of computer data. A byte usually denotes eight bits, which the computer treats as a single unit. Although mega is Greek for a million, a megabyte actually contains 1,048,576 bytes. One gigabyte is equal to 1,024 megabytes. Gigabyte is often abbreviated as G or GB.

6. NATIONAL INFORMATION SHARING

Background

Several topic areas throughout this study included a recurring theme: a need for multiple information-sharing services pertaining to cyber-attack issues. The data-collection needs varied. For example, investigators cited a need for databases of cyber-attack profiles, virus and worm attack signatures, and cyber-attack investigation and prosecution contacts. The common thread is the need to capture, store, and facilitate the retrieval of information for investigators and prosecutors. Study participants continually stressed that sharing of cyber-attack information at all levels is a critical issue. Knowledge transfer is not currently done systematically. Participants communicated that, while they understand that pieces of the needed data sets in many cases already exist, they feel that the mechanisms are not in place to efficiently access the information from all levels of the law enforcement community — i.e., they know the information is out there; they just cannot get at it. If they were able to rely on national data repositories of continually updated information, not only could they conduct their investigations more efficiently, but they could also save money and time by not replicating the efforts of others. This could also contribute significantly to breaking down barriers of communication that sometimes exist between FSL law enforcement, enabling the entire community to coordinate activities, learn from each other on an ongoing basis, and — most importantly — increase the ability to investigate, arrest, and prosecute individuals perpetrating cyber attacks.

6.1 CYBER-ATTACK-PROFILE DATABASE

Findings and Analysis

During the analysis of log files, certain patterns may become apparent that indicate a familiar attack methodology. Recognizing an attacker's profile, beyond the specific attack tools used in the attack, during a manual search requires expert log data interpretation skills and knowledge of continually changing attack exploits. Several law enforcement personnel stated a desire for a database of common cyber-attack profiles that could be cross-referenced with their log data for a particular intrusion. Participants reported that they have no readily available source for recognizing known attack profiles. Law enforcement investigators indicated that a law enforcement database to collect and recognize attack profiles, in concert with a solution for performing technical exploit matching discussed in "Attack Tool Signature Database" (Section 6.3, page 43), would be valuable. Investigative solutions that are designed for cyber-attack data sharing would provide the foundation for a national cyber-attack-profile database.

Widely adopted tools and technologies with data-sharing capability would greatly empower law enforcement. For example, the borderless nature of the electronic infrastructure means that cyber attackers are free to attack from one network to another just as they move from one computer to another in a compromised system. In the current environment, investigators have no way to run a search to see if similar cyber attacks have been undertaken across the country. A study participant commented that a nationwide information-sharing system created for computer-attack profiles would cut the time spent calling other law enforcement agencies trying to identify similar attacks.

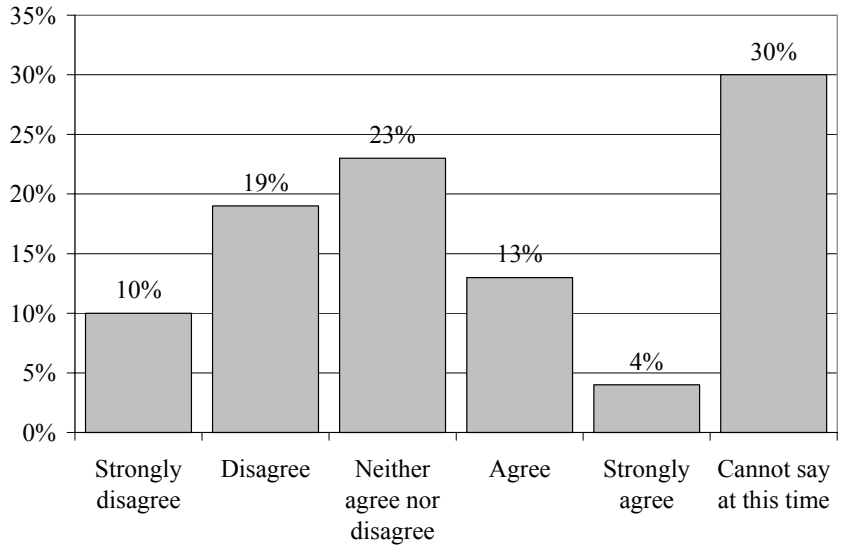
This capability is especially urgent for smaller agencies with limited resources. A national database, based on information generated by the analysis of previous attacks, would allow investigators to quickly assess if their case is a component of larger criminal activity.

6.2 VIRUS AND WORM SIGNATURE DATABASE

Findings and Analysis

Viruses and worms are common cyber-attack weapons that are prolific and destructive. In the last three years, survey respondents have worked on average four cases that involved these programs. Many participants could not say if they were satisfied with the tools they had available to analyze viruses and worms (Figure 11). Investigators who were not completely satisfied with virus and worm analysis software, as shown in Figure 12 (page 42), indicated that tools were not available in most cases (43%).

Figure 11: In general, I am completely satisfied with the tools I have available for the task of analyzing viruses and worms.

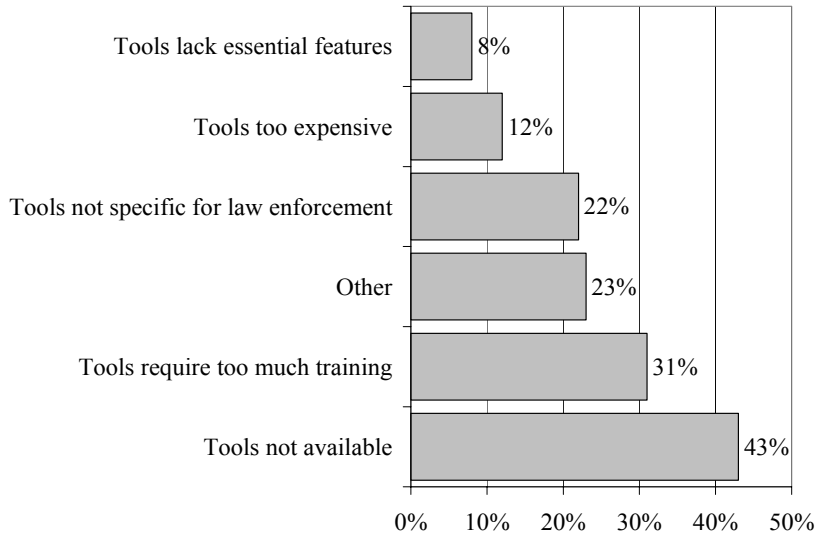


Note: Values are rounded to the nearest integer.

The survey data should be viewed in the context of the wide commercial availability of products to protect, detect, and repair software against virus and worm attacks. The broad adoption and use of virus-detection software has raised public awareness of this type of attack and provides a means for users to rectify the issue themselves. Moreover, of those in the sample population, few respondents were involved in investigating these types of attacks. The reasons for the lack of broad virus attack investigative experience among respondents are many, not least of which is that unless the attack is big enough to draw significant media or public attention, it may not be reported as a crime. If it is not a big enough attack, the victim(s) will usually treat a virus attack as a nuisance, not a crime. At least as important is the difficulty of investigating a virus attack and tracing it to its

source. Successful investigations of virus attacks have required significant investigative resources that state and local investigators do not usually have at their disposal.

Figure 12: If you are less than completely satisfied with the tools you have available for the task of analyzing viruses and worms, please indicate the reason(s) for your dissatisfaction.



Note: Values are rounded to the nearest integer. Number of respondents = 99; excludes respondents who strongly agreed that they were completely satisfied with the software tools for the task of analyzing viruses and worms (6) or who cannot say at this time (46).

Thus, a criminal investigation into a virus attack is generally only undertaken at the federal level. Those who are responsible for investigating viruses have specific technological requirements. Many investigators had to employ manual detection techniques when analysis tools are not available. For example, a study participant conducted research on the Internet to determine characteristics of the viruses and worms and then analyzed case logs to look for exploit signatures. Technological solutions to automate the process of characterizing viruses and worms specifically for cyber-attack cases would be of value to investigators.

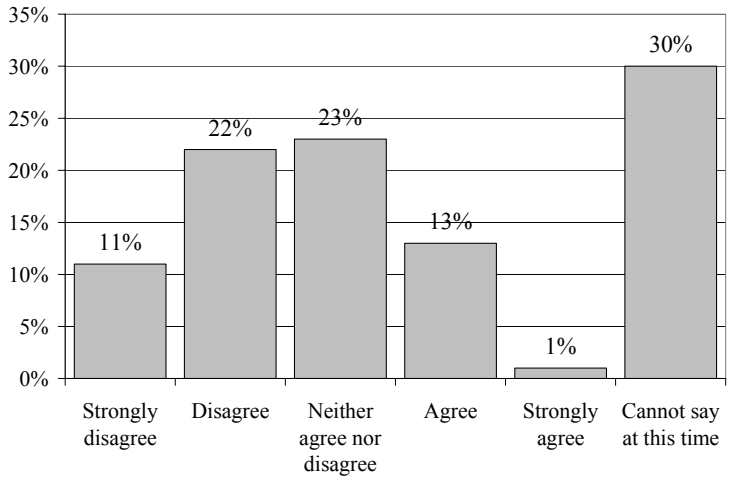
Determining the origin and author of a virus or worm is a significant challenge. Innovative technological solutions to facilitate this process would be welcomed by cyber-attack investigators. Investigators expressed an interest in applying emerging pattern recognition software for this purpose. There are commercial, subscription-based services that note and flag virus and worm signatures, but investigators feel that these services are beyond the budgetary limits of most agencies. A law-enforcement-specific resource to store and compare new virus code to existing examples could yield new insights into ongoing investigations. Additionally, this data source could serve as a national resource for research, information sharing, and analysis. As law enforcement populates the data pool with virus and worm examples from their investigations, research entities could provide analytical products to examine trends.

6.3 ATTACK TOOL SIGNATURE DATABASE

Findings and Analysis

One component of a cyber attacker's profile is the inventory of tools used to exploit computer networks. Root kits, Trojan horses, and other tools have been seen in almost half of the respondents' cyber-attack cases in the last three years. Figure 13 reveals that 33% of the survey's respondents are less than completely satisfied with the tools that are available to detect these sorts of attack tools. Respondents indicated that detection technology is generally not available (Figure 14, page 44).

Figure 13: In general, I am completely satisfied with the tools I have available for the task of detecting Trojans, root kits, and other intrusion tools.



Note: Values are rounded to the nearest integer.

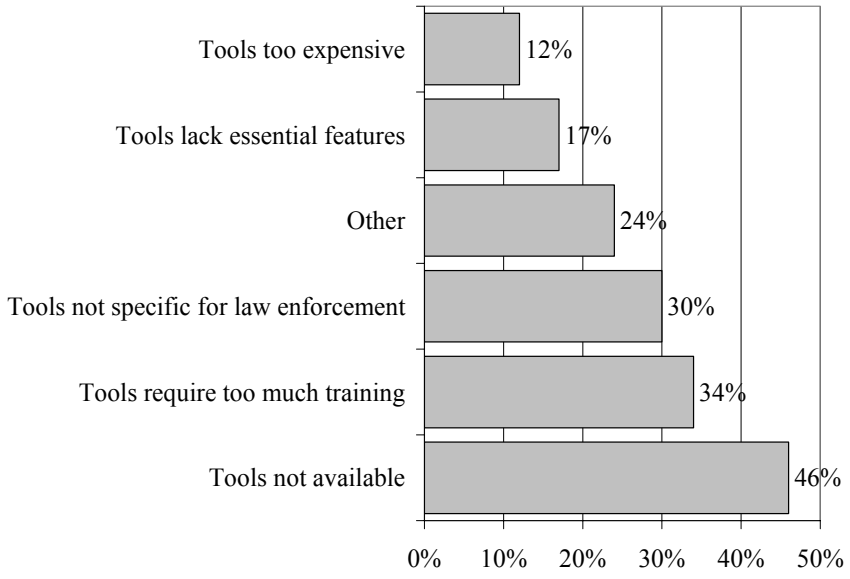
The survey data indicates that solutions for determining if known attack tools are present on a compromised system are less than satisfactory at this time. Site-visit interviews with cyber-attack investigators confirm this analysis. Determining if these programs are present on a particular computer may involve significant manual searching. Available automated search tools are often produced and distributed by questionable sources. For example, these tools are often found at hacker and warez[44] sites. Investigators indicated that a law-enforcement-specific database to search for Trojans, root kits, and other known attack tools would be a valuable data resource.[45] By linking to a trusted data source that is

[44] Warez: Slang for commercial software that has been pirated and made available to the public.

[45] Such a resource might resemble the National Institute of Standards and Technology, which creates and maintains the National Software Reference Library Reference Data Set (RDS), a repository of commercial, off-the-shelf software. The RDS consists of file profiles made up of file-identifying information and hash values from the individual files provided in a software package's installation media for forensic use. Additional information can be found at <www.nist.gov/srd/nistsd28.htm>.

continually updated, investigators would be afforded relevant and timely attack analysis capability.

Figure 14: If you are less than completely satisfied with the tools you have available for the task of detecting Trojans, root kits, and other intrusion tools, please indicate the reason(s) for your dissatisfaction.



Note: Values are rounded to the nearest integer. Number of respondents = 104; excludes respondents who strongly agreed that they were completely satisfied with the software tools for the task of detecting Trojans, root kits, and other intrusion tools (1) or who cannot say at this time (46).

6.4 LAW ENFORCEMENT CYBER-ATTACK CONTACT DATABASE

Findings and Analysis

A recurring frustration for investigators involved in cyber-attack investigations is the difficulty involved in identifying and communicating with other cyber-attack investigators. This was especially salient during the course of real-time intrusion investigations. Reaching out to other investigators is especially important to law enforcement in cyber-attack cases due to the relatively narrow window of opportunity available to collect relevant data to further the investigative process. One survey participant noted that the lack of cooperation results, all too often, in circumstances where cases are shelved and information not shared. There are organizational protocols in place to handle such interdepartmental and cross-jurisdictional cases. However, slow-moving bureaucracies and law enforcement personnel in other jurisdictions working their own caseloads often make outside requests low priorities.

Cyber-attack investigators indicated that they largely rely on their personal network of contacts and a series of phone calls and call backs to further an investigation in a timely manner. Participants pointed out that they need technological resources to facilitate and

coordinate strategic approaches to cyber-attack investigations. The borderless nature of cyber attacks necessitates the cooperation of all levels of American law enforcement. Workshop participants, after discussions of the survey feedback on this data-sharing topic, pointed out the need for a database of cyber-crime investigators in each jurisdiction, including their experience, training, and contact information. This need points to a national data repository of cyber-crime investigation contacts as an achievable solution.

6.5 LEGACY HARDWARE AND SOFTWARE DATABASE

Findings and Analysis

Investigators noted that no single source exists that maintains a repository or electronic warehouse of legacy software or the location of legacy hardware that can be used by investigators for forensic examinations. If investigators and forensic analysts are continually training to the cutting edge of technology, where the majority of network intrusions are likely to be carried out, they may lose the ability to perform thorough investigations on aging hardware and software. Study participants cited a need for the "maintenance of tools for historical purposes" and "[the] ability to work with legacy hardware and software" for investigative and prosecutorial applications. Another participant noted, "Encountering obsolete technologies unexpectedly requires lots of scavenging to obtain suitable equipment for the given situation." Commercial products are not likely to retain data-analysis capabilities for outdated technologies, since the market is driving manufacturers to continually update to the leading standards. This issue may have a technology solution in the form of a warehouse of legacy software and hardware. This type of repository would be a valuable resource for agencies responsible for investigating cyber attack and cyber crime.

7. LAW-ENFORCEMENT-SPECIFIC DEVELOPMENT ISSUES

Background

Investigators and prosecutors volunteered specific details of characteristics that they perceive as important for any technological solution developed for use by cyber-attack investigators. Specifically they commented on users' skill levels and help files. The following considerations are salient to developers of law-enforcement-specific technological solutions.

7.1 SKILL LEVELS

Findings and Analysis

Throughout the course of the survey, it became apparent that different levels of tool complexity are required to accommodate different investigators' skill levels. Many of the less experienced investigators were very interested in a point-and-click user interface, while the more experienced investigators were comfortable with a command-line-based interface. Several experienced investigators expressed concern that new technologies would slow down their investigations if the tool was developed for too broad a range of technical abilities. Therefore, solutions must be robust and have built-in capability to match the skill level and complexity of the user with appropriate functionality. On the whole, the development of solutions specifically for the law enforcement community should recognize that users will have varying skill levels.

7.2 HELP FILES

Findings and Analysis

Tools and technologies used in the course of a cyber-attack investigation will have many users. For example, a first responder may use a forensic analysis tool to isolate information that will later be handed over to a forensic investigator. Upon completion of the forensic examination, the prosecutor and defense attorney may both use the tool to view the investigators' results. Each of these parties will have a different level of expertise and different expectations regarding the functionality of the tool. Several investigators noted that having clear and unambiguous help files that are specifically written for multiple law enforcement users would have a significant positive impact on the success and usefulness of the tool.

8. TRAINING

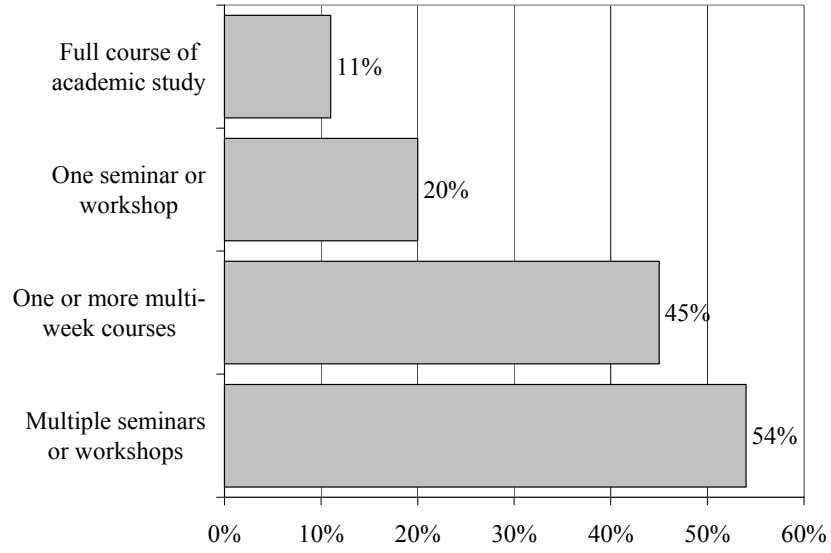
Background

Training is an ongoing process for all law enforcement personnel. The skill set required to successfully investigate and prosecute cyber attacks demands technical competence and ability. There are a number of excellent private and public training programs available to investigators, including programs provided by the National White Collar Crime Center and the National Infrastructure Protection Center. However, without national standards, a certification process, or a commonly recognized training curriculum for cyber-attack investigators, law enforcement faces significant impediments to increasing its ability to investigate cyber attacks at all jurisdictional levels.

Findings and Analysis

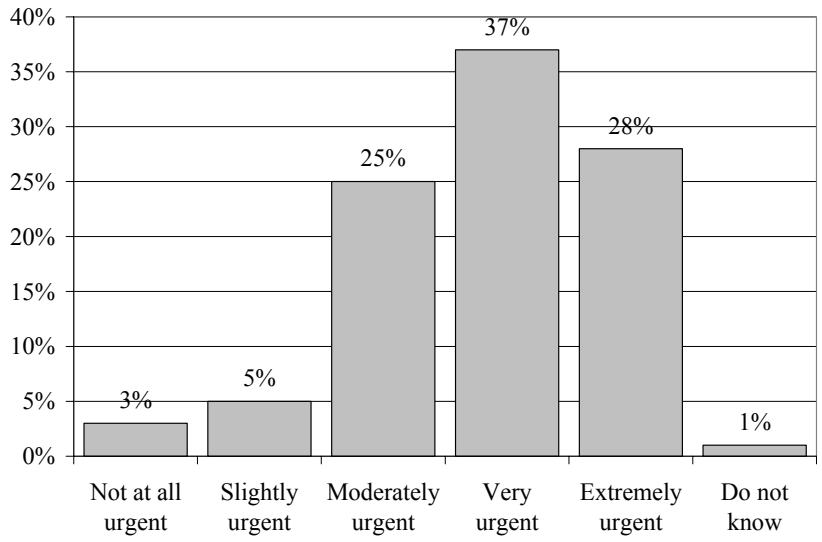
The majority of the survey participants indicated that they had been afforded training in the investigation of computer intrusions or attacks. Only 7% of the survey population reported that they have had no formal training. Significantly, however, only 11% have completed a full course of academic study in a related computer field (Figure 15). When asked about the need for more formal cyber-attack training in their departments, survey respondents were unequivocal. A full 90% of those surveyed stated that the need for additional training was urgent (Figure 16, page 48).

Figure 15: How much formal training in the investigation of computer intrusions or attacks have you had?



Note: Respondents were asked to check all that apply. Values are rounded to the nearest integer. Graph excludes respondents who indicated they had no formal training (10).

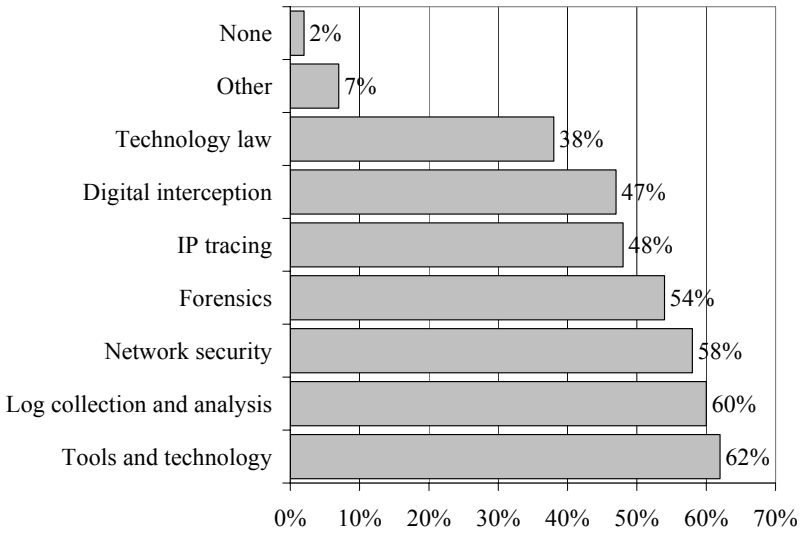
Figure 16: How urgent is the need for more formal training in your department or agency?



Note: Values are rounded to the nearest integer.

Several topic areas were presented in the survey as areas for additional training. As depicted in Figure 17, all were indicated as significant areas where further training was needed.

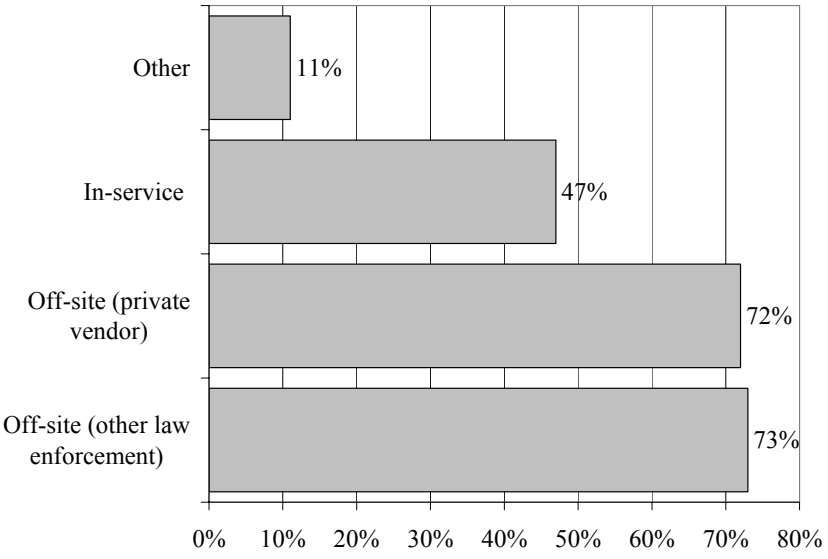
Figure 17: In which of the following areas would you like more formal training?



Note: Values are rounded to the nearest integer.

Investigators received almost equal amounts of training from other law enforcement agencies as they did from the private sector (Figure 18). Many cyber-attack investigators have received in-service training as part of their skill set development.

Figure 18: Where have you received your formal training?



Note: Values are rounded to the nearest integer. Number of respondents = 141; excludes respondents who indicated they had no formal training (10).

The information captured for this study indicates that maintaining up-to-date training in the rapidly changing world of computer security and intrusions is a difficult task. One study participant acknowledged, "I know the need for training is critical. The latest trends are measured in hours and days; we're working in Internet time on these investigations. It seems by the time we get the training, the bad guys are still one step ahead of us. We're always playing catch-up." Once a minimum level of training is obtained by investigators, they are faced with the need to continually keep abreast of cyber-attack tools and techniques. This presents law enforcement agencies with a significant challenge and necessitates the delivery of timely and relevant training. This is frequently provided by private industry.

Currently, law enforcement must go off-site to obtain the majority of the training they require. However, because investigators are afforded only small blocks of time to take in-service training, off-site training is becoming impractical. The study participants spoke favorably about potential new technological solutions for training. For example, network video gaming technologies may warrant additional study for their application to law enforcement training problems. Investigators also indicated that advanced distributed learning (ADL) solutions could provide remote training at their agencies leveraging limited resources. Cyber-attack investigators generally have access to high-speed Internet connections and the technical ability to undertake ADL training on a regular basis.

Supervisors commented that ADL delivery of training materials allows a group of individuals across the nation to receive the standardized training without associated travel costs. Further assessment, research, and development efforts are warranted in this field.

There are significant opportunities for distance learning programs to address cyber-attack investigators' specific needs. One example is the law enforcement request for the timely delivery of training updates on tools and technologies. Law enforcement contends that training should be engineered into software products from the ground up, not added as an afterthought. For example, when service packages are used to upgrade software functionality, cyber-attack investigators require complimentary training updates. In-service training utilizing ADL technologies could accompany major service packages. This may help upgrade investigators' skills and allow them to fully exploit increased software functionality.

Survey respondents indicated that log collection and analysis, network security, computer forensics, IP tracing, digital interception, and technology law were areas in which they require urgent training. Qualitative study data indicates that training needs extend beyond those involved in actual investigations. Several respondents noted that their supervisors had not received sufficient technical training to understand cyber-attack investigations in depth. Similarly, the respondents stated that there is a shortage of prosecutors trained and skilled in cyber-intrusion cases. In a related area, many respondents noted the need for training to enable investigators to testify knowledgeably about not only the evidence that they gather in the course of their investigation, but also the science behind their findings.

While new technologies may provide innovative ways of introducing and maintaining cyber-attack investigative skill sets, this study reveals that the underlying problems are not entirely technology-based. Study participants indicated that demand far outweighs supply for the best commercial training programs. Existing training programs are often prohibitively expensive for state and local law enforcement agencies. With notable exceptions[46], cyber-attack investigators point out that few colleges and universities offer technology or engineering degrees that are specific to cyber-attack investigators' needs. Only a small number of investigative units nationally have a computer scientist or other technically trained individuals on staff. A national strategy should include a common curriculum and standards for training courses for cyber-attack investigators.

To address these issues, at least two national assessments should be conducted. The first should be a collaborative effort to develop a baseline curriculum for training and certification of law enforcement cyber-attack investigators. The second should be the collection and assessment of available training programs.[47] These assessments would allow a gap analysis to be performed to identify what needs are not addressed through

[46] The National Security Agency's National INFOSEC Education & Training Program and the National Science Foundation's Federal Cyber Service: Scholarship for Service are two examples of the excellent programs that are available. There are also encouraging initiatives making their way through Congress, such as H.R. 3394, the Cyber Security Research and Development Act, and S.1901, the Cybersecurity Research and Education Act of 2002, that are focused on increasing the nation's ability to conduct research and educate students.

[47] One example of a national effort to identify training and education opportunities for forensic digital evidence, information assurance, and security is the National Cybercrime Training Partnership Training Database, available at <www.nctp.org>.

existing programs and permit industry and academic researchers to develop pilot programs to fill these gaps. An additional study could be conducted to assess and provide guidance for the creation of focused degree programs in this discipline at colleges and universities. Without these studies, research and development of additional training resources may not adequately address the needs of cyber-attack investigators.

9. CONCLUSION

Cyber attackers are employing advanced technologies that present significant challenges to law enforcement investigators. The scientific community cannot adequately develop the tools and technologies necessary to address the obstacles faced by law enforcement without a clearly defined statement of law enforcement's needs and the development of an R&D strategy to address those needs. The *Law Enforcement Tools and Technologies for Investigating Cyber Attacks: A National Needs Assessment* provides a detailed look at the problems as law enforcement perceives them.

The entities that develop technological solutions to the obstacles outlined in this study have a singular opportunity. Since the existing technology does not meet cyber-attack investigators' needs, the solutions that are developed may become widely adopted by the law enforcement community. Similar to the rise in popularity of the Windows operating system, well-designed software for law enforcement that integrates users' needs could revolutionize the speed and effectiveness of cyber-attack investigations.

The next step in developing a national R&D agenda is a comprehensive national evaluation of existing tools and technologies. By comparing law enforcement requirements with existing solutions, the gaps in existing technology can be determined. The R&D community can then prioritize its research to fill these identified gaps. By working together, researchers in academia, industry, and government can give our public servants the tools they need to address one of the critical public security and national security issues of the 21st century.

10. LIST OF FIGURES

11. LIST OF TABLES

12. ACKNOWLEDGMENTS

The Institute for Security Technology Studies extends its sincere appreciation to the many individuals and organizations from government, industry, and academia that participated in the *Law Enforcement Tools and Technologies for Investigating Cyber Attacks: A National Needs Assessment.*

The following agencies and organizations made their expert staff available to ISTS for consultation and review throughout the duration of this project.

California Office of Criminal Justice Planning
Connecticut Department of Public Safety
Connecticut United States Attorney's Office
Department of Defense Computer Forensics Laboratory
Department of Defense Joint Task Force for Computer Network Operations
Department of Justice Criminal Division, Computer Crime & Intellectual Property Section
Department of Justice National Institute of Justice
Federal Bureau of Investigation
Florida Department of Law Enforcement
Hartford, Vermont, Police Department
High Tech Crime Investigators Association
International Association of Chiefs of Police
Massachusetts Office of the Attorney General Criminal Bureau
Mitre Corporation
National Aeronautics and Space Administration
National Association of Attorneys General
National Infrastructure Protection Center
National Law Enforcement and Corrections Technology Center West
National White Collar Crime Center
New Hampshire United States Attorney's Office
New Jersey State Police
New York Electronic Crimes Task Force
North Bay High Technology Evidence Analysis Team
Ohio Bureau of Criminal Identification and Investigation
Pacific Institute for Computer Security, San Diego Supercomputer Center
Pennsylvania Governor's Office
Pennsylvania State Police
Sacramento Valley High Tech Task Force
San Diego Computer and Technology Crime High Tech Response Team
South Carolina Law Enforcement Division
Southern California High Tech Task Force
Stroz Associates
United States Secret Service

13. CONTACT INFORMATION

Please address comments or questions to:

Law Enforcement Tools and Technologies for Investigating
Cyber Attacks: A National Needs Assessment
The Institute for Security Technology Studies
45 Lyme Road
Hanover, NH 03755
Telephone: (603) 646-0700
Fax: (603) 646-0660

Project e-mail: <needsassessment2002@ists.dartmouth.edu>.

The ISTS web site is available at <www.ists.dartmouth.edu>.

The ISTS Law Enforcement Programs web site is available at
<www.ists.dartmouth.edu/lep>.

Director:

Michael A. Vatis

Research Staff for the Report:

George Bakos
Vincent Berk
Daniel Bilar
Jay Bregman
Dan Burroughs
Kathleen Cassedy
Henry "Chip" Cobb
Julie Cullen
Garry Davis
Edward A. Feustel, Ph.D.
Matt Funk
Paul Gagnon
Trey Gannon
Eric Goetz
Nicole Hall-Hewett
David Koconis, Ph.D.
Stacy Kollias
Andrew Macpherson
Dennis McGrath
Mark Noel
Kevin O'Shea
Richard Scribner, Ph.D.
William Stearns

The following outside organizations and individuals directly contributed to the creation of this study:

Theresa D'Orsi
John Elder
RAND Corporation
Emily Reber, Ph.D.
Ventana East Corporation

14. PUBLICATION NOTICE

Supported under Award number 2000-DT-CX-K001 (S-1) from the Office of Justice Programs, National Institute of Justice, United States Department of Justice.

Points of view in this document are those of the author(s) and do not necessarily represent the official position of the United States Department of Justice.

The authors of this report have made every effort to provide original definitions, use definitions provided in past ISTS publications, and acknowledge the sources of publicly available common knowledge definitions integrated into this document. Footnote definitions were compiled from multiple sources (including <www.cnet.com>, <www.foldoc.org>, <www.lycos.com>, <www.sans.org>, and <www.techtarget.com>) in addition to ISTS scientists. Due to the complexity and public availability of many of the technical definitions used in this study, ISTS acknowledges the possibility that one or more definitions may resemble definitions offered in other non-ISTS sources. We invite the authors or readers of these non-ISTS sources to notify ISTS in writing if similar language is found in this document. Upon notification we will takes steps to verify the claim. If appropriate, ISTS will insert language crediting the appropriate non-ISTS source or to change our own definitional language.