TOOLS IN SUPPORT OF ELECTRONIC CRIME INVESTIGATIONS

By

Leonardo S. Kermes

A Thesis submitted to the
Department of Computer Science
in partial fulfillment of the
requirements for the degree of
Master of Science

Degree Awarded:
Spring Semester, 2007
The members of the Committee approve the thesis of Leonardo S. Kermes defended on December 7, 2006.

Sudhir Aggarwal  
Professor Directing Thesis

Michael Burmester  
Committee Member

Breno deMedeiros  
Committee Member

Zhenhai Duan  
Committee Member

The Office of Graduate Studies has verified and approved the above named committee members.
To the Men & Women of:

The Florida Department of Law Enforcement,

National Institute of Justice,

Florida State University
FCI Research Group, & SAIT Lab,

&

My Friends and Family
Funding for the implementation of law enforcement tools discussed in the following manuscript was provided by grants received from United States Department of Justice - National Institute of Justice. I acknowledge the help of all the people who have worked and contributed on the above-mentioned grants, and all members of the Florida Cybersecurity Institute of Florida State University, Florida Department of Law Enforcement, and the National White Collar Crime Center.

National Institute of Justice Grant 2004-RD-CX-K154, PI Dr. Sudhir Aggarwal
National Institute of Justice Grant 2005-MU-MU-K077, PI Dr. Sudhir Aggarwal
# TABLE OF CONTENTS

List of Figures ........................................................................................................ vi
Abstract .................................................................................................................. vii
1. INTRODUCTION .......................................................................................... 1
   1.1 Law Enforcement .................................................................................... 1
   1.2 Electronic Crime Threshold ................................................................. 1
   1.3 Overriding Goal .................................................................................... 1
2. LEGAL & PRIVACY CONCERNS ................................................................. 2
   2.1 Exploratory vs. Actionable ................................................................. 2
   2.2 Confines of the Law ............................................................................ 2
   2.3 Privacy Concerns ................................................................................ 3
3. CONTRIBUTIONS .......................................................................................... 4
   3.1 PAPA ..................................................................................................... 4
   3.2 UnMask ................................................................................................. 4
4. PREDATOR PREY ALERT SYSTEM (PAPA) ............................................. 6
   4.1 Introduction to PAPA ........................................................................... 6
      4.1.1 General Idea of PAPA ................................................................. 6
      4.1.2 Cyberstalking ................................................................................ 7
      4.1.3 Cyberstaling Laws ....................................................................... 7
         4.1.3.1 Federal Laws ......................................................................... 7
         4.1.3.2 State Laws ............................................................................ 7
   4.2 The Cyberstalker’s Problem ................................................................. 7
   4.3 Requirements ...................................................................................... 9
   4.4 Constraints ......................................................................................... 9
   4.5 PAPA Architecture ............................................................................. 11
      4.5.1 Tier 1 ........................................................................................... 11
      4.5.2 Tier 2 ........................................................................................... 11
   4.6 Implementations ................................................................................ 12
      4.6.1 Victim’s Computer ....................................................................... 12
      4.6.2 Session Recorder ......................................................................... 13
      4.6.3 Officer’s Computer ...................................................................... 14
      4.6.4 VPN Concentrator .................................................................... 15
   4.7 Integration Method ............................................................................. 15
   4.8 National White Collar Crime Center (NW3C) Test and Results ........ 15
5. UNDERCONVER MULTIPURPOS ANTI_SPOOFING KIT (UnMask) ....... 17
   5.1 Introduction to UnMask ....................................................................... 17
   5.2 Requirements Solicitation ................................................................... 17
      5.2.1 ESMTP Email Address Veridication ........................................... 17
      5.2.2 Mail Exchange Records .............................................................. 18
      5.2.3 Traceroute .................................................................................... 18
      5.2.4 Whois .......................................................................................... 19
      5.2.5 DIG .............................................................................................. 19
   5.3 Acceptance of Tools ........................................................................... 20
5.4 Method of Automation

5.5 Status of UnMask

6. CONCLUSION

APPENDICES

REFERENCES

BIOGRAPHICAL SKETCH
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>The Cyberstalker’s Problem</td>
<td>8</td>
</tr>
<tr>
<td>2</td>
<td>The Prisoner’s Dilemma</td>
<td>8</td>
</tr>
<tr>
<td>3</td>
<td>The man-in-the-middle attack</td>
<td>8</td>
</tr>
<tr>
<td>4</td>
<td>The PAPA System</td>
<td>10</td>
</tr>
<tr>
<td>5</td>
<td>Victim Playing World of Warcraft</td>
<td>13</td>
</tr>
<tr>
<td>6</td>
<td>Offices desktop</td>
<td>14</td>
</tr>
<tr>
<td>7</td>
<td>Section of an email address</td>
<td>18</td>
</tr>
<tr>
<td>8</td>
<td>ESMTP Verification of an email address</td>
<td>18</td>
</tr>
<tr>
<td>9</td>
<td>UnMask System</td>
<td>22</td>
</tr>
</tbody>
</table>
This thesis presents the efforts made to develop new tools for non-technical law enforcement agents to effectively respond to electronic crimes. We present two tools: one that combats online predators and cyber-stalking; and a second tool that automates a law enforcement officer’s workflow in investigating email based frauds, threats, or attacks. In this manuscript we discuss the general limitation of law enforcement in investigating cyber crimes, the laws and privacy concerns that constrain the scope of our tools, the crimes we are combating, and the implementation of the Predator and Prey Alert System (PAPA) and Undercover Multipurpose Anti-Spoofing Kit (UnMask) as developed for the National Institute of Justice.
CHAPTER 1

1 INTRODUCTION

1.1 Law Enforcement

In general, Law Enforcement Officers are highly intelligent individuals vested by the state with a monopoly in the use of certain powers: the powers to arrest, search, seize, interrogate; and if necessary, to use lethal force. They are experts in the law of criminal procedure, which regulate officers' discretion, so that they do not exercise their powers arbitrarily or unjustly. Few are experts in: Programming, Operating Systems, Relational Databases, Data Communications, Computer and Network Administration, Computer and Network Security, Cryptography - elements of Computer Science.

1.2 Electronic Crime Threshold

When called upon to conduct inquire into electronic crimes there is a limited number of Officers familiar with Computer Science [22]. This scarcity of expertise causes a threshold of severity a electronic crime must meet in order to warrant the use of such limited resource, examples being Internet Crimes Against Children (ICAC), cases of high monetary loss, or cases that express an imminent threat against one’s life [9] [10] [20]. These thresholds provide the criminal the opportunity to structure his attack in such a manner as to avoid tripping the thresholds. A series of minor electronic crimes disturbed across multiple police jurisdiction could in fact be together a larger crime that warrants investigation. For example, use of spoofed email or website to defraud a million people across the nation out of a small amount of funds - what is commonly known as a salami attacks may not be investigated [4]. Harassment of an adult through electronic means by an anonymous assailant might not result in action by local police until the assailant escalates his actions to the physical world thus losing his anonymity [5] [14].

1.3 Overriding Goal

Thus the overriding goals of developing tools in support of electronic crime investigations is to provide robust tools with simple interfaces such that users of the tool requires only the computer literacy of a novice, and yet the tool performs highly technical tasks such as analyzing the source of an email through obfuscations and showing not a technical report of machine addresses, but rather physical locations of the machines which sent it. Our tools must bridge the gap the use of computer by the perpetrator creates between an officer and the criminal act so that an average officer is no longer technologically unable to effectively respond to the threat.
CHAPTER 2

2 LEGAL & PRIVACY CONCERNS

2.1 Exploratory vs. Actionable

Legal concerns vary on the intended use of the tools’ product, the information the tool creates or gathers. There are two separate general needs expressed by members of the Florida Department of Law Enforcement. The first is tools for gathering evidence for use in a criminal court. This evidence is normally collected under the issuance of a subpoena or warrant, and great care is taken in the collection, recording, and verification. Strict rules govern standardized evidence handling procedures and chain-of-custody records. Since most methods used in the collection, recording, and verification of digital evidence are new with no precedence in criminal courts, they are subject to a Daubert challenge that requires the judge to determine the validity and applicability of the technique or methodology used in the collection, recording, and verification [6]. The Daubert ruling provides for a judge to make an assessment based on the reliability and general acceptance in the scientific community of the techniques or methodologies used [6]. Thomas Duerr et al [8] states that digital evidence is more susceptible to post-collection alteration than analog, and we can agree digital evidence poses a challenge due to the ease in which it can be copied, deleted, and modified. Berg [2] has argued that digital evidence (specifically images) requires special care to document the collection, analyst procedures and the chain-of-custody to ensure admissibility. Clearly such careful endeavors come at a cost in effort, time, money and manpower. This is why there is a greater interest in tools that explore the crime and produce information which is used to determine if an incident requires the efforts needed to gather actionable evidence for the courts, issuance of a subpoena or warrant. Theses exploratory tools are also used to raise the crime above the ether of the digital or electronic world to the surface of the physical world. For example, in cyber abuse one goal of a tool may be to help law enforcement coax the perpetrator into physical meeting, or provide the officer the perpetrator’s location in which he is accessing the Internet. The two tools we present here, Predator and Prey Alert System (PAPA) [1] and Undercover Multipurpose Anti-Spoofing Kit (UnMask), are both exploratory tools. A paper by Burmester et al [4] does present an alternate implementation for cyberstalking, which would produce actionable evidence. Also, UnMask would produce actionable evidence if the emails it investigates were gathered by subpoena directly from the ISP servers hosting the email account. For more information on evidence handling please see the appendix A: “Evidence Handling in Proactive Cyberstalking Investigations: The PAPA Approach.”

2.2 Confines of the Law

Whether the tool is for gathering exploratory information or actionable evidence, the tool itself must operate within the confines of the law. Title III of the Omnibus Crime Control and Safe Streets Act of 1968 (Pub.L. 90-351, June 19, 1968, 82 Stat. 197, 42 U.S.C. § 3711) hold for wiretaps "that the Fourth Amendment did apply to searches and seizures of conversations and protected all conversations of an individual as to which he
had a reasonable expectation of privacy" this also applies to data communication. This requires that a warrant must be issued to intercept communications that are in transit. Thus such interceptions must be avoided in exploratory tools. The pen-trap provisions in 18 U.S.C. §§ 3121-3124 and amended by the USA Patriot Act allows capture of routing, addressing, or signaling information transmitted by an instrument or facility from which a wire or electronic communication is transmitted, provided, however, that such information shall not include the contents of any communication. Law enforcement agencies must get approval for surveillance by court order from a judge. However, they need only certify to the judge that the information likely to be obtained is relevant to an ongoing criminal investigation, at which point the judge 'shall' issue the order. This is the lowest requirement for receiving a court order under any of Electronic Communications Privacy Act’s three titles. This reflects the fact that pen registers do not gather the contents of any information, only the addressing or routing information. The Pen Register Act did not include an exclusionary rule. Evidence gained in violation of the Act can still be used against a defendant in court. Since there is no penalty for the violation of the act and permission form the court is simply asking for it routing and addressing information can be intercepted in an exploratory tool.

2.3 Privacy Concerns

Although we are only required to preserve the expectation of privacy afforded to us by the 4th Amendment and the laws expressed above, every courtesy must be afforded to the victim as to not violate her expectation of privacy. Disclosing the function and purpose of the tools to the victim prior to their use and allowing the victim to deny the use of the tool will allow us to operate with in her expectation of privacy. We must operate under her knowledgeable consent.
CHAPTER 3

3 CONTRIBUTIONS

To develop software for law enforcement within the constraints described in the previous section, this thesis introduces the methods we used to create systems and applications for law enforcement. Our efforts were shared by partnerships with the Florida Department of Law Enforcement (FDLE), and the National White Collar Crime Center (NW3C) to ensure the practicality of our solutions. Our research builds on previous and ongoing work on developing tools, software and techniques addressing electronic crimes. Our approach differs from previous work in this area.

3.1 PAPA

The original Predator Prey Alert System was proposed by Dr. Sudhir Aggarwal, et al, in a proposal {not funded} to the National Institute of Justice entitled “Anti-CyberStalking: Tools and Training for Law Enforcement” submitted in May 2003. In the proposal Dr. Aggarwal et al clearly presented an argument for the need for a new tool for addressing the crime of cyberstalking and online predators and proposed the first conceptual system. This was the basis for a white paper proposing PAPA submitted to the NIJ in January 2004 and subsequently funded.

In the spring of 2004, Melissa Kryder described in her thesis a possible software solution for addressing the problem of cyberstalking borrowing the name from the NIJ Proposal and presented another conceptual system [23]. In 2004/2005, I began with Sudhir Aggarwal and others to refine and implement the concept of PAPA under funding from the NIJ. By employing both hardware and software to produce a solution for cyberstalking we designed and implemented a functioning secure solution. I was primarily responsible for the implementation of the system.

My work on architecture and some aspect of implementation appears in [1] and also as appendix B. I also contributed to ensuring that PAPA would meet necessary legal requirements for evidence. Work on this aspect appears as [24] and appendix A.

3.2 UnMasK

Some tools exist to analyze the source of emails from websites through implicitly using Unix console commands. We contribute to this area by using a parser to automatically extract relevant information from the email and explore gathering further information via the use of Unix tools, where previous methods required a manual extraction of such information. Also, we collate the information into a relational database such that reports can be generated and business logic applied through the use of the structured query language (SQL). My contribution has been in the design of the UnMask architecture, and in developing a precise set of requirements and also the initial specifications.
The remainder of this thesis is organized as follows. Chapter 4 discusses the Predator Prey Alert System. Chapter 5 gives an initial design of the Undercover Multipurpose Anti_Spoofing. Finally, chapter 6 describes our conclusions.
CHAPTER 4

PREDATOR PREY ALERT SYSTEM (PAPA)

4.1 Introduction to PAPA

4.1.1 General Idea of PAPA

The Predator and Prey Alert System (PAPA) is an exploratory tool for combating online predators and cyber-stalking. The system is designed to help victims of cyber-stalking by permitting law enforcement agents to “shadow” the victim remotely and provide online assurance and guidance while securely capture and log data related to the cyber-stalking activity in order to support investigations, while not compromising the security of the victims computer system. Mechanisms of the PAPA system are designed to be user friendly, so that it requires little training to deploy and use in an investigation of cyber-stalking by non-technical users. The PAPA system was designed with the following goals in mind:

• permit law enforcement to remotely “shadow” a victim and provide assurance and advice when needed;
• capture and log appropriate data related to cyberstalking activities so that a law enforcement analyst can subsequently investigate and determine the identity of the predator; and
• capture evidence of probative value so that the predator can be successfully prosecuted.

Thus the design of the system supports observation, capture, indexing, and preservation of the integrity of collected evidence, and provides a WorkBench suite of tools for the forensic analysis of captured data [1].
4.1.2 Cyberstalking

Stalking is a pattern of behavior over time through which a stalker seeks to gain access to, or control over, a victim. Cyber-stalking, the convergence of stalking and cyberspace - stalking behavior found in: email, chat rooms, multiplayer games, or other electronic means. The term Cyberstalking is interchangeable with online harassment and online abuse. Internet anonymity empowers the predator as he follow the victims online activity gaining the knowledge he needs to escalate cyber-stalking, to make threats, intimidate or physical stalk. Stalking is not a single incident; rather it is a continuous process, which can be terrifying for the victims. Terrorizing at the risk of psychological trauma, and or physical harm [5] [11] [15] [22].

4.1.3 Cyberstalking Laws

Many stalking laws require that the perpetrator makes a credible threat of violence against the victim or the victim's immediate family and others require the alleged stalker's conduct constitute an implied threat [5] [11] [12] [14] [15] [17].

4.1.3.1 Federal laws. 18 U.S.C. §875 – Requires an actual threat to kidnap or injure and is punishable by up to five years in prison and/or a fine up to $250,000. The Communication Decency Act of 1996 addresses annoying, abusive, and threatening behavior punishable as a misdemeanor, while the 1996 Interstate Stalking Act addresses physical stalking which crosses state lines, which the Violence Against Women Act, passed in 2000, made Cyberstalking a part of the federal interstate stalking statute.

4.1.3.2 State laws. In Florida, HB 479 was signed into law on October 2003. Alabama, Arizona, Connecticut, Hawaii, Illinois, New Hampshire and New York have prohibitions against harassing electronic, computer or e-mail communications in their harassment legislation. Alaska, Oklahoma, Wyoming, and California, have incorporated electronically communicated statements as conduct constituting stalking in their anti-stalking laws.

4.1.4 The Cyberstalker’s Problem

Now that we established that cyberstalking is a crime we attempt to create a model for cyberstalking. Burmester et al [4] put fourth the Cyberstalker’s problem, a three party communication game related to the classical Prisoner’s dilemma and the man-in-the-middle attack [17] [18]. This game illustrates the stalker establishing a (virtual) harassment channel with the victim. This channel is what law enforcement wishes to capture the contents of with out the stalker having any knowledge of the presence of law enforcement. This means that the law enforcement must have a point of presence covertly to intercept the harassment channel (the man-in-the-middle) and law enforcement must be able to communicate with the victim with out the stalker knowing (prisoner’s dilemma). See Figures 1, 2, 3.
Figure 1. The Cyberstalker’s problem: the Monitor has to capture harassment data exchanged by Alice and Bob [4].

Figure 2. The Prisoner’s dilemma: Wendy has to prevent Alice from sending Bob a covert message [4].

Figure 3. The man-in-the-middle attack: Eve intercepts the communication between Alice and Bob: she may simply eavesdrop or corrupt communicated data. Alice and Bob want to establish a private and/or authenticated communication channel.
By considering the Papa Problem it is clear that we need to keep private and secret the communications of the PAPA system form the Cyberstalker. The other requirements are as follows.

### 4.3 Requirements

The Predator and Prey Alert System (PAPA) must:
- Permit law enforcement agents to “shadow” the victim remotely,
- Permit law enforcement agents to remotely control the victims computer,
- Provide a means of secure and private communication between the victim and a law enforcement agent,
- Securely capture and log data related to the cyber-stalking activity, if capturing network packets in their entirety, a Title III court ordered warrant is necessary.
- The victim must have the control to start and stop the monitoring,
- Not compromising the security of the victims computer system,
- Mechanisms of the PAPA system must be user friendly
- PAPA must be simple to deploy

### 4.4 Constraints

The PAPA system:
- Must not monopolize the network bandwidth of the victim
- Most not monopolize the system resources of the victim
- Victims internet service provider (ISP) dynamically assigns addresses
- Victims may or may not be a network address translated private address positioned behind a network firewall
- Officer may or may not be a network address translated private address positioned behind a network firewall
Figure 4. The PAPA System
4.5 PAPA Architecture


The design of the PAPA system employed a distributed dual tier approach. Not counting the stalker, the system consists of software on four machines: the victim’s machine, a session recorder, a virtual private network (VPN) concentrator, and an officer’s machine. The session recorder is local to victim’s the computer, sharing Internet access serially. The VPN concentrator can be located anywhere, provided it has a global ip address. The officer’s computer may exist anywhere with connectivity. The only known ip address is that of the VPN Concentrator. The VPN creates the secret private communications for the PAPA system.

4.5.1 Tier 1

The first tier provides the network infrastructure of the system via the VPN and network services offered by the session recorder. A private subnet is split between Dynamic Host Configuration Server on the session recorder and the VPN concentrator. Starting with Session recorder, it receives a random public address from the victim’s ISP and connects to the VPN concentrator and receives a known address in the private subnet of the VPN. When requested, the session recorder assigns a known address in the private subnet of the VPN to the victim’s computer. When the officer computer comes online it too connects to the VPN Concentrator and receives a known address in the private subnet. Now that all the systems have known address the VPN Concentrator and the session recorder routes the private traffic across the VPN subnet. The session recorder also acts as a network address translating firewall providing the victims computer with Internet access. This provides the network infrastructure to support the second tier application. Besides assigning known address to the machines in the PAPA system, the VPN provides a secure and private channel for communication, and access control by authentication to the VPN.

4.5.2 Tier 2

The second tier contains the application software for: law enforcement agents to “shadow” the victim remotely, to remotely control the victims computer, chat applications between the officer and victim, and other applications to gather information.
4.6 Implementations

The PAPA System was for the most part an exercise in network/systems administration and application software integration.

4.6.1 Victim’s Computer

On the first implementation attempt the victim’s computer (running Windows XP) received an installation of TightVNC server configured as a system service, which had to be started and stopped manually. Next, it had a custom application (see figure 5) that ran in the users space (as opposed to a system service) that performed several task. One, it started and stopped the TightVNC Server. Two, it was the chat application between the victim and the officer. Three it was the remote controls for starting and stopping session recorder gathering of evidence. The software written performed within expectations with only a few bugs due to undocumented differences between Berkley Socket Programming and Windows Socket programming. TightVNC performed with no apparent errors.

Virtual Network Computing (VNC) is a desktop sharing system using the RFB (Remote Frame Buffer) [16] protocol to remotely control another computer. It transmits the keyboard presses and mouse clicks from one computer to another relaying the screen updates back in the other direction, over a network. VNC being platform-independent allowed us to not only to use it to fulfill the requirement of the officer to be able to shadow and control the victim’s computer, but also allowed the Linux based session recorder to record the RFB. This is essential in order for this device to operate as an exploratory tool since VNC is recoding information that has already been delivered to the victim. Thus there is no expectation of privacy for the attacker. This enables us to avoid the need of a Title III warrant.

On the Second implementation attempt the custom application on the victim’s computer received new features. Visual indicators to the availability of the officer were added to the application. Exception handling, and logging to the windows event viewer of user interaction with the application as well as errors was added to the application.
Figure 5. Victim is Playing World of War craft while being shadowed by the officer.

4.6.2 Session Recorder

Tier 1 of the session recorded consisted of a ‘vanilla’ Linux Kernel from kernel.org with the options set specifically for the hardware of the machine along with option enabled for packet forwarding, routing, point to point tunneling, and virtual private networking. For the VPN client, L2TP/IPsec was attempted using the Linux Racoon Client, connecting to a Linux L2TP/IPsec Racoon server, but it proved to be an unstable solution and was replaced with the Linux PPTP client connecting to a Windows 2003 Server as the VPN concentrator. PPTP is less secure, but more stable. Next, Iptables was configured to apply a firewall when an address was assigned to the public interface. The address assigned by the ISP as mapped to the private addressee reserved for the victim’s computer and network address translation was applied. At the end of the firewall script the VPN tunnel was launched along with rules for forwarding packets to the VPN tunnel. Then network services for the support of the victim’s computer were configured in the common ways. The services included DHCP (dhcpcd), DNS (named), and NTP (ntpd) to insure indications of time were synchronized to a trusted source. At this point tier 1 is effectively built for the session recorder.
Tier 2 begins with the compilation of a program called vncrec, which simply records the RFB to a file. A more exciting application to integrate with its dependencies was Transcode, which had the potential of converting the linear rendering VNC recording to a common video play back format such as MPEG2 or Divx. Unfortunately the Transcode application failed to work consistent and was dropped from subsequent builds. The next application was the obvious candidate for the recoding of the network traffic directed at the victim’s computer was SnortIDS, which is a very popular packet sniffer, but sadly it did not play well with others having conflicting dependencies versions with other applications. An attempt was made to deploy virtualization technologies, the Xen Hypervisor, but sadly it was too taxing on the hardware used. The utility TCPDUMP replaced Snort IDS in the second Build. To simply the deployment to of the session recorder to the field, Apache web server was installed so the software needed on the victim’s computer could be simply downloaded off the session recorder in a simple web interface. Finally there was a demon application written to run and stop all the other applications mention above remotely by the application on the victim’s machine.

Figure 6. Officer’s desktop: Here you can see he office shadowing the victim while he plays World of Warcraft.

4.6.3 Officer’s Computer

The officer’s computer contains a copy of the TightVNC Client and custom application (see figure 6) for controlling the VNC Client and chatting with the victim. The VPN client was the native Windows application. Similar to the victim’s custom application, on the second implementation the application received new features. Visual
indicators to the availability of the officer were added to the application. Exception handling, and logging to the windows event viewer of user interaction with the application as well as errors was added to the application.

4.6.4 VPN Concentrator

The VPN concentrator was originally a Linux server, but due to reliability and Windows / Linux interoperability issues, it was replaced with a Windows 2003 server. This was the simplest part of the entire system to setup. Simply run the Windows Remote Access and Routing Wizard selecting it to host the VPN service, then create two user accounts, and designate for them an address to be assigned to the accounts when they connected to the VPN.

4.7 Integration Method

To simplify integration VMware Virtualization software was employed to take advantage of the “Snap Shot” feature which allowed us to revert the virtual system back to previously indicated state. Thus, the first step was to create a virtual machine of the Linux distribution that was to be used on the session recorder. Next was to test each individual piece of software alone against the Linux distribution. If available a pre-compiled binary (a binary compiled by the developer of the software) was used to install the particular application. As dependencies (an application or library that the primary application we are attempting to install is dependent on) arose we would also try to satisfy them using a pre-compiled binary. The same action was taken for dependencies of dependencies, which we encounter up to six layers of dependencies. It turned out that most dependencies needed to be recompiled with the options for the binaries to use a common threading library (pthreads) and to be set as shareable in the compilation. This required header files and sources to be added to the environment and properly pathed with the necessary tools installed to generate the necessary configuration and makefiles needed to compile. Once each application was successfully install, and functioning without any apparent errors with the Linux distribution, we began combining them on to a single virtual machine and worked on resolving version conflicts where two applications would required a common shared object, but each needed a different version of that object. This was the approach we used for the tier 2 applications. Tier 1 required modification to the Linux Kernel, so we did not use VMware in its development since the snap shot would just be hardware. The final phase was joining tier 1 to tier 2 and resolving issues that arose then. In hindsight it would have been simpler to build the system entirely form scratch, never using any open source software as it took nearly two years to work through conflicts and defects in the open source software.

4.8 National White Collar Crime Center (NW3C) Test and Results

The PAPA system did function with a few hiccups (unexplained pauses in execution), and lockups (system stopped responding) in the controlled environment of the laboratory. The system did function well enough to demonstrate proof of concept. During independent test outside the laboratory by NW3C, the PAPA Systems never fully
functioned and NW3C was unable to utilize remote access for an officer, the passive recording did function, but playback was limited. All network traffic was dumped to a binary file. This traffic included the VNC information packetized and intermixed with all the other traffic. NW3C found reconstructing the packetized information with a network diagnostic tool such as Ethereal beyond their skill set. Video Playback of the Victims desktop also failed to impress NW3C. Since Transcode was not able to render more than a few minutes of video before crashing, playback was limited to being played through the vncrec application linearly from the start of the video to the end without the ability to jump to a time of interest, or being able to fast-forward or rewind. NW3C did suggest an application that would convert the Macromedia Flash file format (SWF), but that also only plays linearly and the portably advantage is mute since a raw RFB file could be played back linearly in a Windows VNC client as well. If they found one that converted RFB to Flash Video (FLV) File Format then they would have had a superior solution. We believe NW3C were expecting a production quality application rather than a prototype. Open source / freeware hardly ever reaches such high reliability. Even Cisco System has know issues with their VPN solution failing. Perhaps our goals were too ambitions, but to my satisfaction the prototypes did show that our ideas were possible and there were moments where the systems did work flawless. Counting all the dependencies there were over a hundred pieces of software from different groups making up the PAPA system. Just getting the system to run briefly was a massive effort. The true flaw of the PAPA system’s development perhaps was that NW3C was kept separate from the development of the system so that they could validate the system in a proper independent study. However, for proper software engineering the end users must be involved with the development. This mistake was not repeated in the next tool developed - UnMask.
CHAPTER 5

5 Undercover Multipurpose Anti-Spoofing Kit (UnMask)

5.1 Introduction to UnMask

The Undercover Multipurpose Anti-Spoofing Kit’s goal is to automate a law enforcement officer’s workflow in investigating email based frauds, threats, or attacks; such as: phishing, emailed threats against life or property, extortion, or a malicious payload in the email. The measure to judge the success of UnMask is to simply demonstrate that it can generate a report faster than the hours it takes an officer to perform the task manual. Since we are simply automating a process already performed by the police the legality of the tool is not in question, but we will address them quickly. The email message is already delivered to the destination prior to it being submitted to law enforcement thus it is not a pen trap or wiretap. The information gathered to analyze the email is in the public domain and has no expectation of privacy.

5.2 Requirements Solicitation

We moved from the broad strokes of a grant proposal to a design by observation. We performed a series of interviews with Florida Department of Law Enforcement, as well as examined training material, and best practices prescribed to law enforcement for examining an email, and then worked to improve the efficiency of their efforts. What law enforcement does is examine each part of an email and extracting our email and server address, and then use websites the offer utilities such as traceroute, and the ability to retrieve DNS records and WHOIS DNS Registers information to produce a report describing where the message most likely came from. This report would be used to assess future actions: issuing a subpoena to a local ISP for customer record, forwarding the case to another jurisdiction such as the Secret Server, or abandoning the case due to a lack of evidence.

5.2.1 ESMTP Email Address Verification

We proposed the developing a parser to extract email address from field common to all email regardless of the client used to send the message - email addresses found in originator fields, destination address fields, resent fields as defined in RFC2822 [25], and addressee marked us as such in the hypertext markup language (HTML) found in the message body of an email. Using the ESMTP Commands for debugging Addresses defined in RFC 2821[26] we can verify if a username found in the local part of an email address is a member of the domain found in the domain part of an address [25].
Figure 7: Sections of an email address

<table>
<thead>
<tr>
<th>Local-part</th>
<th>Delimiter</th>
<th>Domain-part</th>
</tr>
</thead>
<tbody>
<tr>
<td>John.Dow</td>
<td>@</td>
<td>Some.site.com</td>
</tr>
</tbody>
</table>

Server: 220 mail.cs.fsu.edu ESMTP Postfix
Client: HELO cs.fsu.edu.com
Server: 250 Hello cs.fsu.edu.com
Client: VRFY kermes
Server: 250 kermes@cs.fsu.edu
Client: QUIT
Server: 221 BYE

Figure 8: ESMTP Verification of an email address; VRFY request asks the mail server to verify that an email address is on that server. If the server accepts the request, it may provide information about the address in a server-defined format. Code 250 means the address is valid, code 251 typically means that mail to the address is forwarded, and 252 means that the server doesn't know whether the address is valid. [26]

5.2.2 Mail Exchange Records

This verification may indicate that the sender address may be false, but this not conclusive since the user account may have been temporary. The next logical step would be to retrieve the MX recorded (Mail exchange record, RFC 974 / RFC 2821) [27] [26] for each email address domain by interrogating a series of Domain Name Servers (DNS). We proposed using DNS severs distributed around the globe (Tallahassee, FL; London, UK; Cologne, DE; Tokyo, JP, Los Angeles, CA), so that we can propagation of a change in a DNS record which on average takes a day or more to transverse the entire internet. Using path information in the MX record we can compare the path shown in the email header. A miss-match in the two paths shows the injection point of a spoofed email.

For examination of Uniform Resource Identifier (URI) [28] of hosts found in MX records, the domain-part of the email addresses, trace fields of the message header, and HTML links in the message body we use a set of UNIX commands: traceroute, whois and dig. [29]

5.2.3 Traceroute

Traceroute is a network tool that determines a possible route for packets across an IP network. Traceroute produce a list of hosts that packets traverse en route to the destination by incrementing the "time-to-live" value of successive batch of packets sent do the destination. The first packets have a time-to-live (TTL) value of one such that they make a single hop. The next packets have a TTL value of 2, and so on. [30] When a
packet passes through a host, the host decrements the TTL value by one, and forwards the packet to the next host. When a packet with a TTL of one reaches a host, the host discards the packet and sends an ICMP time exceeded (type 11) packet to the sender. The traceroute utility uses this ICMP return packet to identify the hop.

Unix and Linux use UDP datagrams with a destination port number starting at 33434, and Windows uses ICMP echo request (type 8) instead to perform a traceroute. There are also implementations that use TCP packets, such as tcptraceroute or Layer Four Trace. The websites used by the officers of the Florida Department of Law enforcement never disclose which method of traceroute they use. The method could have an effect on the results the utility produces because some routers may block arbitrary ports such as 333434 or ICMP messages, or a router may behave differently to different packet type by performing quality of service (QoS). By examining the different implementations of the utility we found that the Berkeley Software Distribution (BSD) implementation allowed us to specify all the different methods of performing a traceroute, and we propose using all methods.

5.2.4 WHOIS

The UNIX command line tool ‘whois’ is used to retrieve administrator contact information for host on the Internet. We have an advantage using the command line tool over the use of a website by being able to specify the ‘whois’ database we wish to retrieve the information from. We can script the command line tool to ask all the available generic databases: Internet Assigned Numbers Authority (IANA), Network Abuse Clearinghouse, Network Solutions Registry for Internet Numbers, and Verisign. As well as databases for specific groups of domain: American Registry for Internet Numbers (ARIN), Asian/Pacific Network Information Center (APNIC), R’eseaux IP Europ’eeens (RIPE), Russia Network Information Center (RIPN), Latin American and Caribbean IP address Regional Registry (LACNIC), US Department of Defense database ‘.mil’, and US non-military federal government database ‘.gov’. Finally we can gather information for specific countries such as from ac.whois-servers.net - Ascension Island to zw.whois-servers.net – Zimbabwe.

5.2.5 DIG

Dig (domain information groper) is a UNIX command line tool, which is used to gather information from the DNS servers. Where the website currently used by law enforcement only queried a subset of the possible resource record types we have the ability to ask for them all. The complete list of DNS resource record types is documented in RFC1034 and RFC1035 [31] [32]. The websites omit obviously useful resource record for law enforcement such as LOC defined in RFC1876 [33], which gives the latitude and longitude of the host. An example would be to get the LOC of the ip assigned to computer used to write this manuscript. The query returns latitude: 30.462900, longitude: -84.244598; which is Tallahassee, Florida. Using the Google Maps API we can simple generate a graphical estimation of the IP address’s location in the real world.
We propose retrieving the following records:

- **A** - address record maps a hostname to a 32-bit IPv4 address
- **NS** - name server record maps a domain name to a list of DNS servers authoritative for that domain. Delegations depend on NS records.
- **CNAME** - canonical name record makes one domain name an alias of another. The aliased domain gets all the subdomains and DNS records of the original.
- **WKS** - Well Known Service Record
- **RP** - Responsible Person, used to indicate the person responsible for the domain.
- **GPOS** - geographical location in longitude, latitude, and altitude
- **AAAA** - IPv6 address record maps a hostname to a 128-bit IPv6 address
- **LOC** - a means for expressing location information in the Domain Name System. It contains WGS84 Latitude, Longitude and Altitude information together with host/subnet physical size and location accuracy.
- **NAPTR** - Naming Authority Pointer, a newer type of DNS record that support regular expression based rewriting.
- **DNAME** - provides aliases for a whole domain, not just a single domain name as with CNAME
- **HINFO** Host Information in terms CPU and operating system
- **MINFO** Mailbox or mailing list information, specifies the mailbox that should receive messages or errors for this object
- **MX** - mail exchange record maps a domain name to a list of mail exchange servers for that domain.
- **TXT** - allows an administrator to insert arbitrary text into a DNS record.

### 5.3 Acceptance of Tools

Once we have all the tools need to perform the email analysis in the UNIX Shell it is obvious to a person with a professional level of computer literacy to see that one could simply use Perl, a dynamic programming language with strengths in string processing, to parse the email and analyze each email and host address found in the message and produce a simple report. But in order for the novice to accept this method, it must be explaining and demonstrated to them. But acceptance of the tool can be made even simpler if we separate the report logic out of the Perl. To do this we use a relational database. This allows us to lower the technical skill set need to manipulate the results of the Unix tool’s information. Commercial off the shelf software such as Microsoft Access can be used as a database front-end allowing a simple secretary to create custom reports and perform queries. This is a clear advantage over the cost of a programmer writing a new Perl application.

### 5.4 Method of Automation

Our method for automation is a simple one. The law enforcement agent submits an email via a website front-end to a relational database that accepts email saved as a file. On receiving the file, the database invokes a parser to extract address within the message and populates tables with that information. This action invokes the database to request the
use of Unix tools to gather additional information from network services. The tools are evoked from within an application, which submits the result of the tool back to the database. Once all the information is gathered a report can be generated using stored SQL queries against the data sets to present the information in a meaningful way. The use of a database also allows us to mine data in interesting ways like how many crimes are associated to a particular website hosting company. Finally the use of a database allows us to apply logic to the data. The database could be queried to show the path an email should have traveled according to the mail exchange (MX) records and compare it to the path it took according to the received headers of the email message to show the rogue mail server which injected email with a spoofed source. Most importantly a presentation quality report is available nearly instantly for officers to act on.

Another advantage of the use of a database is that it allows us to scale the system. All modern database can be distributed across multiple servers, and using an Open Database Connection (ODBC) give use a standard interface for our font-ends and the capability for multiple front-end. Database security and best practices are well established, and he access controls, which allows us to have a secure application relatively easily. See figure 7.
5.5 Status of Unmask

At the time of this manuscript Unmask had just begun implementation. Several Requirement and Design documents have been written and the end-user is active in the development.
CHAPTER 6

6 CONCLUSIONS

Law Enforcement Officers are highly intelligent individuals and their time comes at a premium. The most important item to keep in mind when developing software for the support of electronic crime investigations is to actively peruse the law enforcement community’s contribution. Unlike traditional end users law enforcement has additional constraints on them in order to protect our liberties. Law Enforcement is hesitant to engage with new procedures and tools out of fears of violation their constraints. When developing for law enforcement you must not only satisfy their functional requirements, but also educate them on how satisfaction was achieved.

Simplicity is probably the most crucial problem to be resolved when creating a tool. The tool must be simpler to use than the task it performs. Only then will the tool be used and a burden lifted from the officer. In this thesis, we presented two tools for combating electronic crime.
APPENDIX A

Evidence Handling in Proactive Cyberstalking Investigations:
The PAPA Approach

S. Aggarwal, P. Henry, L. Kermes, & J. Mulholland
Department of Computer Science & Florida Cybersecurity Institute
253 Love Building, Florida State University, Tallahassee, FL 32306-4530
sudhir@cs.fsu.edu, henry@cs.fsu.edu, kermes@cs.fsu.edu, judiemul@cs.fsu.edu

Abstract
Stalking is the malicious, unsolicited intrusion on another’s personal space, and cyberstalking extends this to cyberspace via Internet technology. All fifty U.S. states criminalized stalking in the 1990’s, and many have passed cyberstalking statutes as well. The anonymity and reach of the Internet, and the difficulties in capturing, recording, and verifying digital evidence combine to create new challenges to law enforcement agencies trying to prevent and detect the crime and apprehend the criminals. In particular, the “expectation of privacy” afforded to all participants of live-wire communication makes it difficult to bind the actual perpetrator with his or her online persona. The PAPA system is a comprehensive toolkit that captures all relevant cyberstalking data with the potential for admissibility in a court of law. To this end, and as far as possible under existing federal, state, and international statutes, it captures data with the goal of producing evidence that is admissible, authoritative, reliable, complete, and believable.

1. Introduction

The Predator and Prey Alert (PAPA) toolkit is a proactive hardware and software system of forensic tools designed to provide high quality, verifiable evidence for the prosecution of cyberstalking cases. It can capture, record, and verify critical data directly from the computer of the victim. Law enforcement agents are able to analyze all network, IM, video, audio, email, and even encrypted traffic to discover evidence of cyberstalking, as well as to identify and help capture the perpetrator. Aggarwal et al provide an overview of the contemporary cyberstalking problem and the technical specifics of the PAPA system [1].

This paper explicates legal and technical issues of evidence handling in the fight against cyberstalking from a system design perspective. The specific goals of the PAPA system include:

* This work is supported by the National Institute of Justice under grant 2004-RD-CX-K154.
• **record** the user experience of the cyber-stalking victim in a high-resolution, lossless, verifiable format;
• **ensure** the integrity of the evidence recorded by the PAPA Session Recorder,
• **extend** the evidentiary “chain of custody” to the Computer Under Investigation (CUI),
• **capture** all available digital forensic evidence of the victim’s experience during the cyber-stalking;
• **prevent** undetected pre- and post-computing of evidentiary data;
• **provide** strong time-based verification of data reception, recording, and encryption through massive redundancy; and
• **enable** flexible playback and queries of the cyberstalking experience of the victim.

This proactive approach avoids the investigatory issues of anonymity, lack of records, and under-reporting inherent in computer crime cases. Proactive monitoring of cyberstalking activity also takes advantage of the repetitive nature of the crime as the predator repeatedly enters the victim’s cyberspace. Furthermore, the predator’s expectation of anonymity helps law enforcement undercover agents monitor activity because agents can assume the online identity of the victim to solicit further personal information or to set up a rendezvous with the stalker. From a technical perspective, PAPA hardware and software capable of recording all video, audio, and keystroke activity on the victim’s computer are currently feasible, but proactive investigatory techniques raise questions regarding compliance with federal wiretapping statutes. This article focuses on the ability to legally and contemporaneously record and identify the parties and activities involved in cyberstalking with the aim of gathering the admissible electronic evidence essential to a successful prosecution.

2. PAPA system overview

The Predator and Prey Alert (PAPA) system consists of a set of integrated tools designed to support law enforcement in assisting victims and capturing evidence of cyberstalking. In successful cases, it binds the online identity to the actual perpetrator for subsequent prosecution. The PAPA system was designed in pursuit of the following goals:

• **permit** law enforcement to passively “shadow” a victim remotely;
• **provide** real-time assistance and advice from a law enforcement agent;
• **capture** and log relevant data related to cyberstalking activities;
• **bind** the action to the identity of the predator; and
• **capture** evidence of probative value so that the predator can be successfully prosecuted.

In pursuit of these goals, the system supports the robust capture, indexing, and verification of data, “remote desktop” control of the CUI, online chat technology for continuous communication between the victim and the agent, a networked Dispatcher that monitors the state of the Session Recorder, and a Workbench tool for forensic analysis of the captured data.

3. Cyberstalking and the law
3.1 Federal law

The U.S. Department of Justice defines stalking as “harassing or threatening behavior that an individual engages in repeatedly” [2]. Cyberstalking in its simplest form is stalking via the medium of the Internet [3]. On the Federal level, there are currently four areas of the United States Code that are germane to cyberstalking: 18 U.S.C. §875(c), 18 U.S.C. §2261A, 18 U.S.C. §2425, and 47 U.S.C. §223. The three most important are:

(i) 18 U.S.C. §875(c) which makes the transmission of any communication in interstate and foreign commerce containing an actual threat to kidnap or injure another, punishable up to five years in prison and/or a fine of up to $250,000 [4],

(ii) The Communications Decency Act of 1996 (CDA), 47 U.S.C. §223, which makes a communication used with the intent to annoy, abuse, threaten or harass another in interstate or foreign commerce by means of a telecommunication device punishable as a misdemeanor [5], and

(iii) The 1996 Interstate Stalking Act, 18 U.S.C. 2261A, which makes it a felony for an individual to travel across state lines with the intent to injure or harass another and in the course of or as a result of that travel, placing that person in reasonable fear of death or serious bodily injury to that person or a member of that person's immediate family [6].

These Federal laws are primarily concerned with interstate and foreign commerce communications that are not within the jurisdiction of a single state. The Interstate Stalking Act in particular criminalizes the crossing of state lines with the intent of injuring or harassing another, but there are cases in which the courts consider any use of the Internet as interstate commerce. For example, in United States v. Kammersell, the defendant threatened a victim located in the same state. Although the defendant argued that the threat was not transmitted in interstate commerce the court held that since the electronic message traveled from Utah to Virginia and back to Utah, it met the requirement of being a communication in interstate commerce [7].

3.2 State law

On the state level, there has been a growing trend toward the criminalization of cyberstalking. After the widely publicized 1990 murder of actress Rebecca Schaeffer, California became the first state to criminalize stalking. By 1995 all fifty states and the District of Columbia had followed suit. Although the codification of the criminalized behavior varies from state to state, all of the statutes require some type of action, such as following, harassing, or engaging, in a “course of conduct” that is directed at a victim [8]. As Table 1 shows, other than this similarity, there is a great deal of variety among the states.

<table>
<thead>
<tr>
<th>Requirement</th>
<th>States</th>
</tr>
</thead>
<tbody>
<tr>
<td>Require prosecutors to prove the commission of certain acts and the presence of a specific intent to harass or instill fear in the</td>
<td>AL, AR, CA, FL, GA, ID, KS, KY, LA, MA, MD, MS, MO, NE, NM, NC, OK, PA, RI,</td>
</tr>
</tbody>
</table>
Most state statutes treat cyberstalking as an extension of physical stalking and thus require physical pursuit or contact. Others require prosecutors to prove the malicious intent of the perpetrator. Ironically, these laws ignore the phenomenon of erotomania that ultimately led to the death of Schaeffer. Stalking laws that require the prosecution to prove specific intent place a heavy burden on the state and thus potentially limit the number of successful prosecutions. More progressive laws look more closely at the damage done to the victim and specifically include electronic communication in their definitions of acts that may constitute the crime.

Florida codified its anti-cyberstalking law in 2003 with Florida Statute 784.048(1)(d). The Florida statute is broader than the California “credible threat” model because of its two-tier approach, which does not require a credible threat for a stalking conviction. This statute defines cyberstalking as a:

means to engage in a course of conduct to communicate, or to cause to be communicated, words, images, or language by or through the use of electronic mail or electronic communication, directed at a specific person, causing substantial emotional distress to that person and serving no legitimate purpose [9].

The Florida Statute makes such cyberstalking a first-degree misdemeanor [10], but “aggravated stalking,” defined as making a “credible threat with the intent to place that person in reasonable fear of death or bodily injury” is punishable as a third degree felony [11]. Also included as third degree felonies are continued stalking after any court-imposed prohibition of conduct toward the subject [12], and stalking of a minor under sixteen years of age. Persons who have been sentenced for sexual battery [13], lewd or lascivious offenses [14], or an order on no contact [15] are also subject to third degree felony charges for stalking.
3.3 International law

The Canadian Criminal Code [Section 264. Criminal Harassment (2)] prohibits repeated, unsolicited communications with another person, known or unknown, and this extends to cyberspace [16]. The Australian Criminal Code Stalking Amendment Act 1999 Chapter 33A-Unlawful Stalking, 359B(c)(ii) criminalizes “contacting a person in any way, including, for example, by telephone, mail, fax, e-mail or through the use of any technology” that would cause the stalked person apprehension or fear, reasonably arising in all the circumstances, of violence to, or against property of, the stalked person or another person; or causes detriment, reasonably arising in all the circumstances, to the stalked person or another person [17]. The United Kingdom Malicious Communications Act 1988 Section 1, as amended by Section 43 Criminal Justice and Police Act 2001, criminalizes the sending of “a letter, electronic communication or article of any description which … cause[s] distress or anxiety to the recipient or to any other person to whom he intends that it or its contents or nature should be communicated.” The U.K. Telecommunications Act of 1984, Section 43 also makes punishable anyone who, “by means of a public telecommunication system, a message … of an indecent, obscene or menacing character or … for the purpose of causing annoyance, inconvenience or needless anxiety to another,” or who “persistently makes use for that purpose of a public telecommunication system” [18].

4. Characteristics of evidence

Evidence admissible in court consists of testimony, exhibits, or other admissible documentary material. It must be relevant, authentic (as in the “best evidence rule”), and introduced either by or through a witness (who is under oath and is subject to direct examination and cross-examination) or, exceptionally, without a witness (if it bears adequate certification). Parties in litigation carry a “burden of proof” and pieces of evidence that they introduce are admissible if they are relevant and material. The body of evidence presented by a party is then put to the test of being beyond a reasonable doubt, being clear and convincing, and attaining preponderance of the evidence. Any material that does not meet the above requirements is inadmissible as evidence, including “hearsay” - statements made not under oath or not subject to cross-examination. The five characteristics of good evidence can be identified as follows:

(i) **Admissibility**: conforming to the specific legal rules of the court;
(ii) **Authenticity**: positively tied to the incident of crime;
(iii) **Reliability**: nothing about how the evidence was collected and subsequently handled casts doubt about its authenticity and veracity;
(iv) **Completeness**: illuminating the whole story and not just a particular perspective; and
(v) **Believability**: being readily believable and understandable by a court [19].

These five categories provide a useful framework to classify and evaluate the evidence produced by PAPA. Electronic evidence is a fairly recent development in Anglo-Saxon jurisprudence, but it has become increasingly important as computers and information technology are used in crime. Computer evidence is similar to any evidence in that it needs to conform to the requirements above. In the next section, we will consider the
specific characteristics of electronic evidence and how the PAPA system facilitates admissibility, authenticity, reliability, completeness, and believability of evidence collected in cyberstalking cases.

5. Admissibility

In order for the PAPA system to be effective in the prosecution of cybercrime cases, the evidence it collects must be admissible in a court of law. The two primary impediments to admissibility in stalking cases are the privacy protection of the perpetrator and the deniability of online identity that downgrades the perpetrator’s online statements to hearsay.

5.1 Privacy

Privacy of online communication is specifically protected under two primary sources of American law: the Fourth Amendment to the U.S. Constitution; and the statutory privacy laws codified at 18 U.S.C. §§ 2510-22, 18 U.S.C. §§ 2701-12, and 18 U.S.C. §§ 3121-27. Although constitutional and statutory issues overlap in some cases, most situations present either a constitutional issue under the Fourth Amendment or a statutory issue under these three statutes.

The Fourth Amendment limits the ability of government agents to search for evidence without a specific warrant issued by a judge:

The right of the people to be secure in their persons, houses, papers, and effects, against unreasonable searches and seizures, shall not be violated, and no Warrants shall issue, but upon probable cause, supported by Oath or affirmation, and particularly describing the place to be searched, and the persons or things to be seized [20].

Although the Fourth Amendment seems to require warrants for searches, two exceptions have emerged from Supreme Court rulings. First, as held in Illinois v. Andreas, 463 U.S. 765, 771 (1983), if the investigation does not violate a person’s “reasonable expectation of privacy,” then it does not constitute a Fourth Amendment “search” and no warrant is required. Second, as ruled in Illinois v. Rodriguez, 497 U.S. 177, 185 (1990), even if a warrantless search violates a person’s reasonable expectation of privacy, it may be “reasonable” and constitutional if it falls within an established exception to the warrant requirement.

Just as courts distinguish the physical pages of a letter from its addressed envelope, they now distinguish between “information,” the actual delivered content of an electronic communication, and the addressed TCP/IP packets that carry it. Thus, even if the parties to Internet communication have a subjective expectation that third parties will keep the “information” confidential, they cannot reasonably expect to retain control over the mere information actually delivered to third parties. United States v. Miller, 425 U.S. 435, 443 (1976), held that customers could not expect privacy in regard to bank account information. Smith v. Maryland, 442 U.S. 735, 743-44 (1979) established that there was no reasonable expectation of privacy in phone numbers dialed into the telephone system. Couch v. United States, 409 U.S. 322, 335 (1973) held that a client’s financial
information supplied to an accountant similarly retains no reasonable expectation to privacy.

Because the courts have deemed computer data “information,” individuals who send data over communication networks lose Fourth Amendment protection of the data after it reaches its intended recipient. United States v. Meriwether, 917 F.2d 955, 959 (1990) ruled that an electronic message sent via a pager is “information,” and United States v. Charbonneau, 979 F. Supp., at 1184, ruled that

E-mail transmissions are not unlike other forms of modern communication. We can draw parallels from these other mediums. For example, if a sender of first-class mail seals an envelope and addresses it to another person, the sender can reasonably expect the contents to remain private and free from the eyes of police absent a search warrant founded upon probable cause. However, once the letter is received and opened, the destiny of the letter then lies in the control of the recipient of the letter, not the sender, absent some legal privilege.

Thus an email message, like a letter, cannot be afforded a reasonable expectation of privacy once that message is received. Moreover, a sender of e-mail runs the risk that he is sending the message to an undercover agent [21].

In White v. White, 344 N.J. Super, 211, 781 A.2d 85 (2001) a woman in a custody case used email found on the family computer as evidence, while her husband argued that admission of the documents violated the New Jersey state wiretap act, and common-law right of privacy. The court held that, under the state wiretap act (which is identical to the language of a similar federal statute) a violation occurred only upon authorized access to an electronic communication stored “incidental to the electronic transmission thereof” [22].

There are two federal statutes that set the standard for real-time electronic surveillance in federal criminal investigations. First is “Title III,” the “wiretap” statute, 18 U.S.C. §§ 2510-2522, appended as Title III of the Omnibus Crime Control and Safe Streets Act of 1968. Second is the “Pen/Trap” statute, or the “Pen Registers and Trap and Trace Devices Chapter of Title 18,” 18 U.S.C. §§ 3121-3127. Title III permits government investigators to obtain the contents of wire and electronic communications while in transmission only with a court order, and the Pen/Trap statute limits surveillance to the real-time collection of addressing and other non-content information relating to those communications. Together, they permit the capture of sensitive and relevant information of ongoing communication deemed criminal by the judge issuing the court order. Originally drafted to tap into telephone conversations, they are now also used to capture suspect TCP/IP packets and email. Internet Pen/Trap devices capture addressing information, while Internet Title III intercept devices capture and retain content. For email, Pen/Trap devices capture the addressing and routing information from the “header,” while Title III devices capture the content within the “body.” Interception of TCP/IP packet payloads and email contents, including the subject line, requires careful compliance with the strict dictates of Title III, and it is relatively difficult to obtain court orders under Title III.

Under the Pen/Trap statute, a government attorney can apply to a court for an order authorizing the installation of a pen register and/or trap and trace device only if “the information likely to be obtained is relevant to an ongoing criminal investigation” [23]. The Pen/Trap statute also requires that government agents use “technology reasonably available to it” in order to avoid recording or decoding the contents of a wire or electronic
communication [24]. The USA PATRIOT Act of 2001 confirmed that the Pen/Trap statute applies to a wide range of communication technologies [25]. Previously, the definition of “wire communication” explicitly included “any electronic storage of such communication,” but the PATRIOT Act specifically deleted this phrase and amended § 2703 of the Electronic Communications Privacy Act, 18 U.S.C. §§ 2701-12 (ECPA). The ECPA, however, still restricts access to stored account records and contents from commercial service providers, such as Internet service providers (ISPs), telephone companies, cell phone service providers, and satellite services.

“Electronic communication” is a broad, catch-all category that includes most electric or electronic signals that do not fit the definition of “wire communications.” Almost all Internet communications, including email and chat, qualify as electronic communications. 18 U.S.C. § 2510 (12) defines electronic communication as:

- any transfer of signs, signals, writing, images, sounds, data, or intelligence of any nature, transmitted in whole or in part by a wire, radio, electromagnetic, photo-electronic or photo-optical system that affects interstate or foreign commerce, but does not include:
  - any wire or oral communication;
  - any communication made through a tone-only paging device;
  - any communication from a tracking device; or
  - electronic funds transfer information stored by a financial institution in a communications system used for the electronic storage and transfer of funds [26].

It is important to note that communication obtained from electronic storage is not subject to Title III court order requirements because acquisition of the contents of both stored electronic or wire communications is governed by § 2703(a) of ECPA, not by Title III. Most courts consider interception of communication to refer only to their real-time acquisition at the time of transmission between the parties. Steve Jackson Games, Inc. v. United States Secret Service, 36 F.3d 457, 460-63 (1994) held that investigators had the right to access stored e-mail communications [27]. Because communications are intercepted only if acquired contemporaneously with transmission, a keylogger, which is extremely intrusive technology, will intercept communications if it is configured such that keystrokes are recorded when the computer is online [28].

In cybercrime cases, seven exceptions to Title III’s prohibition of the intentional interception, use, or disclosure of wire and electronic communications [29] have been established:

- interception pursuant to a § 2518 court order;
- the ‘consent’ exception, § 2511(2)(c)-(d);
- the ‘provider’ exception, § 2511(2)(a)(i);
- the ‘computer trespasser’ exception, § 2511(2)(i);
- the ‘extension telephone’ exception, § 2510(5)(a);
- the ‘inadvertently obtained criminal evidence’ exception, § 2511(3)(b)(iv); and
- the ‘accessible to the public’ exception, § 2511(2)(g)(i).

Pursuant to cyberstalking cases, the “consent” and “accessible to the public” exceptions are most pertinent. 18 U.S.C. § 2511(2)(c)-(d) states:

(c) It shall not be unlawful under this chapter for a person acting under color of law to intercept a wire, oral, or electronic communication, where such person is a
party to the communication or one of the parties to the communication has
given prior consent to such interception.

(d) It shall not be unlawful under this chapter for a person not acting under color of
law to intercept a wire, oral, or electronic communication where such person is
a party to the communication or where one of the parties to the communication
has given prior consent to such interception unless such communication is
intercepted for the purpose of committing any criminal or tortious act in
violation of the Constitution or laws of the United States or of any State.

This language authorizes the interception of communications if one of the parties to the
communication consents to the interception. This implied consent doctrine permits
monitoring either if a computer network has been properly “bannered” as being
monitored, or if one of the parties to a communication consents. The PAPA system relies
on the explicit consent of the victim for its authorization.

There is a consensus that in Internet use, items are only accorded privacy protection if
they can be reasonably assumed to be private. Thus, since items posted to public bulletin
boards, chat rooms, and other forums are widely accessible to the public, they cannot be
considered private [30]. Winick and Cutrera argue that:

posting a message in the publicly accessible areas of a BBS can be viewed as
either putting the message into ‘plain view,’ or as voluntarily disclosing the
information to all other parties. One loses any expectation of privacy in an
otherwise private item by placing the item into plain view. As a result,
outsiders such as law enforcement officials may monitor BBS communications
if those communications are stored or transmitted in a manner that is accessible
to the public. Similarly, voluntary disclosure of information to another permits
the other party to relay that information to law enforcement personnel without
offending the Fourth Amendment [31].

PAPA takes advantage of this vulnerability of cyberstalkers, who use public cyberspace
and delivered communication to intimidate their prey. When such attacks occur in
publicly accessible regions of cyberspace, they are subject to capture and recording by
technology like PAPA. Although the PAPA system is capable of tapping the TCP/IP
traffic between the predator and the prey, it does not record the body of the packets in
human-readable form. Because PAPA looks at chat and email communication after it has
been delivered to the application layer of the CUI, it records only “information” that
carries no expectation of privacy in a format the agents can read.

5.2 Hearsay

Hearsay is an example of inadmissible evidence in court, and it has its counterpart in
computer records. Computer records containing text fall into two categories: computer-
generated records, and records that are merely computer-stored [32]. Computer-stored
records are electronic documents written by human beings, including email messages,
word processing files, and Internet chat room logs. In order to comply with the hearsay
rule, records must demonstrate the circumstances indicating that the human statements
contained in the record are reliable and trustworthy and the records must be authentic.
Computer-generated records include the output of computer programs, without human
“statements,” and are subject to different evidentiary rules than computer-stored records.
Hearsay rules allow for the testing of human assertions in court through cross-examination [33]. This is not possible when a computer makes an assertion since it cannot be called for cross-examination at trial. Because of this, Federal Rule 801(a) states that an assertion cannot contain hearsay if it was not made by a human and Federal Rule 801(b) states that “a declarant is a person who makes a statement” [34].

The PAPA system is built on open-source software and its functionality is designed to be readily apparent to expert analysts. This accessibility to the underlying code thus avoids the introduction of hearsay in the form of either human-manipulated or machine-manipulated output.

6. Authenticity

Computer records must be shown to be authentic for admissibility in court under Rule 901(a) by offering evidence “sufficient to support a finding that the [computer record or other evidence] in question is what its proponent claims” [35]. Challenges to the authenticity of computer records include questions of alteration, manipulation, or damage, the reliability of the computer program generating the records, or the identity of their author. For example, United States v. Tank [36] involved logs of Internet chat room conversations. The court found that a co-conspirator’s Internet chat room log printouts were relevant and authentic, even though they were not complete documents, and that undetectable “material alterations,” were possible prior to the government’s seizure of the computer. Prosecutors established a connection between the defendant and the screen name he used when he participated in the chat room.

PAPA uses authentication routines throughout its operations of recording sessions, installing software, monitoring the state of the Session Recorder, distributing cryptographic tickets, and data analysis. The system requires prior authentication from the victim before any case is initiated and at the beginning of each session. Additionally, during a session, the victim can flag, via a “panic button,” an incident perceived as an attack. This is primarily to aid in indexing relevant points in the potentially lengthy session files, but it also gives the victim the ability to identify incidents and authenticate them as suspicious activities.

PAPA aims to recreate as much of the actual desktop experience of the victim during the session as possible. To this end, it gathers as much circumstantial evidence surrounding cyberstalking attacks as possible, including video, text, TCP packet headers (including timestamps), and even audio. The aim of this redundancy is to tie, as strongly as possible, the actual perpetrator to the evidence of the attack. This can be extremely difficult in other cybercrimes, especially when they involve sophisticated attackers, but in the case of cyberstalking, the perpetrator is often known by the victim and the attacks are repetitive and pattern-forming. PAPA uses all these aspects of stalking to create a profile of IP addresses, screen names, email addresses, attack patterns, etc. to more fully bind the online attacks with the human perpetrator. PAPA’s Active Mode in particular allows an expert agent to elicit additional personal information about the stalker or to draw the stalker into a rendezvous at which identity is concretely established and apprehension is possible.

Data generated by PAPA is analyzed by forensics experts via the WorkBench software module. It is expected that any testimony in court will use such analysts as expert
witnesses. The circumstantial data available to the analysts enables them to closely tie the evidence with the crime in the court of law.

7. Reliability

The PAPA system always operates as a passive recorder of any desktop session the victim authorizes it to capture, but the PAPA Session Recorder can optionally be configured as an Ethernet bridge between the CUI and its broadband connection. In this configuration, it is capable of capturing all incoming and outgoing TCP/IP and other Internet traffic. Without a court order or exception, of course, investigators would be in violation of the Title III Wiretap Statute when capturing data content. In terms of capturing data headers, appropriate exceptions to the Pen/Trap statute would permit PAPA to collect such information that could subsequently be used via the WorkBench.

Because the PAPA system uses separation of duties, the field agent is incapable of decrypting the output. Only with a key held by a reliable third party are analysts able to analyze these packets. This maintains the expectation of privacy afforded to all live wire communications, but provides the possibility of decrypting the packets if a court order is obtained. Nevertheless, packet payload capture is redundant because the content of the packets is delivered to the application layer of the CUI, streamed to the Frame Buffer, captured by PAPA’s Victim Module software, and recorded by the Session Recorder, and this functionality is offered primarily for analysis of routing, origin, and timestamps contained in the packet header.

As a networked system, PAPA uses several methods to confirm and document the collaboration of its modules. The Dispatcher acts as a monitor of the Session Recorder’s state and it coordinates the participation of law enforcement agents. If the connection to the Dispatcher or Agents is broken, the PAPA system “fails on,” but clearly indicates this state of affairs as metadata within the output video file. The ability to constantly monitor and record the state of the system gives analysts important indicators of the reliability of the system and the evidence it collects.

8. Completeness

Current PAPA prototypes contain large hard drives dedicated to data storage. This capacity allows the system to record copious amounts of information, and the design has massive redundancy in data capture. For example, the video stream uses the highest resolution (raw mode) of the VNC server because bandwidth and storage capacity are not limitations. This allows for high-resolution “screen shots,” if desired, for physical evidence. In addition to the huge amounts of video data captured the system also records the textual content of online communications, the VNC control traffic and chat dialogue between the victim and agent, the TCP/IP headers (and, optionally, the encrypted payloads), as well as any audio stream that the victim flags as offensive. All this data, as well as the “heartbeat” and authorization signals for the remote Dispatcher, are stored with their corresponding timestamps to aid in the reconstruction of cyberstalking-related events.

In addition to the data captured by the Session Recorder, the PAPA system allows law enforcement agents to experience cyberstalking remotely, in real time. Through the
coordination of the Dispatcher, multiple agents, even operating in different jurisdictions, are able to shadow the activity. A further benefit of this capability is the mobility of the agents through wireless technology, which could aid in physical apprehension of the perpetrator.

9. Believability

Digital evidence is unusually fragile and easily deleted, modified, or corrupted. PAPA is designed to give prosecutors the ability to clearly and accurately recreate the experience of the victim as it occurs during the attack. The high resolution of the video capture and the extensive metadata indexing of the video files allows the analyst, as expert witness, to demonstrate to the court a clear video of the action during the recorded session. Additional functionality includes keyword searching, fast-forward and rewind, and high-resolution “still shots” of significant events during the session, and these features assist prosecutors in the presentation of their case.

Because of the highly technical nature of PAPA evidence, an expert witness will be needed to present it in court. The PAPA system has been designed and implemented to be a transparent as possible in all of its functionality. The use of open source code allows the expert witness to gain confidence in the operation and integrity of the system and its output, and instill that confidence in the courtroom. PAPA collects evidence in several redundant formats, and this redundancy also increases the believability of the output.

10. Conclusion

Cyberstalking has created considerable challenges for victims and law enforcement because of the anonymity and reach that the Internet provides, because of the fragility of electronic evidence, and because of privacy statutes that protect all forms of live-wire communication. The PAPA system takes advantage of the cyberstalker’s aggressive and repetitive behavior by recording the victim’s desktop experience of attacks and by giving investigators the ability to proactively intervene. This intervention can be total, with the agent assuming the online identity of the victim, or collaborative, with the agent and victim working together via a second communication channel.

The PAPA Session Recorder collects the video, text, TCP/IP packets, and, optionally, the audio from the CUI during a session. It receives authorization for its operation from the GUID of the CUI and the PAPA Dispatcher Module. In the absence of such remote authorization, it can optionally record anyway, but with encryption to preserve the privacy of the victim. By capturing and recording email, chat text, and other web content from the video buffer and text from a keylogging process, PAPA avoids most of the Title III wiretap restrictions that prevent other proactive surveillance techniques. By recording only session-flagged TCP/IP headers in the clear, it conforms to current Pen/Trap restrictions. By optionally encrypting TCP/IP payloads, it maintains the privacy of the electronic communication until such time as a court order is issued to decrypt and analyze it. Thus, PAPA produces high quality evidence of cyberstalking activity while preserving the online privacy of both the predator and the prey.

The PAPA system is designed to collect evidence that conforms to admissibility standards by avoiding surveillance methods that violate privacy statutes. PAPA assures
that evidence will be authentic and reliable if sworn law enforcement agents and analysts use standard forensic protocols for the deployment and operation of the system and the analysis of the results. The electronic evidence collected is as complete as possible by including as much relevant circumstantial metadata about the session as possible, and is believable even by a non-technical jury because it faithfully captures the desktop experience of the victim in a lossless format that allows for video-like playback, pausing, and searching.

Acknowledgements
The authors wish to thank Bob Breeden and the Florida Department of Law Enforcement’s Computer Crime Center for their support and help in this work. We would also like to thank Mark Gage, Wade Grant and the National White Collar Crime Center for their helpful input. The authors also thank Michael Burmester, Alec Yasinsac, Dave Kelley, Sahil Cooner, Newton Etienne, Sandeep Kalburgie, and Kim Vogt for helpful discussions.

Reference:


[9] Florida Statute 784.048(1)(d)

[10] The punishment for such first degree misdemeanors can be found at: http://www.flsenate.gov/Statutes/index.cfm?App_mode=Display_Statute&Search_String=&URL=Ch0775/Sec082.HTM, and
http://www.flsenate.gov/Statutes/index.cfm?App_mode=Display_Statute&Search_String=&URL=Ch0775/Sec083.HTM.

[11] The statutes can be found at:
http://www.flsenate.gov/Statutes/index.cfm?App_mode=Display_Statute&Search_String=&URL=Ch0775/Sec082.HTM,
http://www.flsenate.gov/Statutes/index.cfm?App_mode=Display_Statute&Search_String=&URL=Ch0775/Sec083.HTM, and
http://www.flsenate.gov/Statutes/index.cfm?App_mode=Display_Statute&Search_String=&URL=Ch0775/Sec084.HTM.

[12] Examples are at:
http://www.flsenate.gov/Statutes/index.cfm?App_mode=Display_Statute&Search_String=&URL=Ch0784/Sec046.HTM, and
http://www.flsenate.gov/Statutes/index.cfm?App_mode=Display_Statute&Search_String=&URL=Ch0741/Sec30.HTM.

[13] URL:
http://www.flsenate.gov/Statutes/index.cfm?App_mode=Display_Statute&Search_String=&URL=Ch0794/Sec011.HTM.

[14] URL:
http://www.flsenate.gov/Statutes/index.cfm?App_mode=Display_Statute&Search_String=&URL=Ch0800/Sec04.HTM.

[15] URL:
http://www.flsenate.gov/Statutes/index.cfm?App_mode=Display_Statute&Search_String=&URL=Ch0921/Sec244.HTM.

[16] URL:

[17] URL:

[18] URL:


APPENDIX B

Anti-Cyberstalking:
The Predator and Prey Alert (PAPA) System*

S. Aggarwal, M. Burmester, P. Henry, L. Kermes, and J. Mulholland
Department of Computer Science & Florida Cybersecurity Institute
253 Love Building, Florida State University, Tallahassee, FL 32306-4530
sudhir@cs.fsu.edu, burmester@cs.fsu.edu, henry@cs.fsu.edu, kermes@cs.fsu.edu,
judiemul@cs.fsu.edu

Abstract
Stalking is a crime typified by repeated harassment of another person and intrusion
upon his or her privacy. Cyberstalking extends stalking into the realm of cyberspace
wherein a predator stalks a victim or prey through internet technologies such as emails,
chat rooms, and instant messaging. This paper describes the Predator and Prey Alert
(PAPA) system. PAPA consists of a set of integrated software and hardware modules and
tools designed to support law enforcement in helping victims of cyberstalking, facilitate
the investigation of such crimes, and maintain evidence for the potential prosecution of
the cyberstalker.

1. Introduction

Stalking is a traditional crime typified by repeated harassment of another person or
intrusion upon his or her privacy with the intent of intimidation or harm [1]. On the
physical level it traditionally involved the continued pursuit or observation of another [2],
as predators typically stalked ex-husbands or wives and ex-boyfriends or girlfriends,
teachers, doctors, or celebrities. Predators usually know their prey and are generally
considered normal, although some are sociologically obsessed [3, 4], and sometimes
target complete strangers as well. They also use, and often try various tactics to acquire,
private information about their prey. At its worst, stalking can lead to threats or violence.
Some jurisdictions have criminalized all intentional stalking [5], while most others deem
it illegal only when accompanied by physical threats or danger to the prey.
The Internet has provided new opportunities for predators to stalk their prey [6, 7], and
such activities are commonly included under the rubric “cyberstalking.” The anonymity
and ubiquity of the Internet seemingly allow the stalker relative safety and freedom [8],
within the new forums of websites, chat rooms, internet relay chat, and email. The
increasing prevalence of online stalking activity presents some new challenges for law
enforcement because:

* This work is supported by the National Institute of Justice under grant 2004-RD-CX-K154.
(i) the behavior patterns associated with predators and prey are not only complex, varied, and unpredictable, but they are difficult to recognize, investigate, assess, and prevent;

(ii) there is no single or standard predator / prey profile to assist investigators;

(iii) internet anonymity not only emboldens the dynamics of the relationship(s), but it makes it all but impossible to determine the true identity of the source of the communication [9], and

(iv) the technical ability of the cyberstalker often exceeds that of the victim and the law enforcement agent involved.

Cyberstalking often involves verbal or emotional intimidation or the pursuit of further personal information, although in and of itself does not present a direct physical threat to the victim. Cyberstalking often has psychological and emotional effects on the victim, and can escalate into physical stalking and dangerous liaisons with the predator.

As the public and lawmakers have come to perceive the dangers of cyberstalking [10], laws have been enacted around the U.S. California became the first state to establish an anti-cyberstalking law in the early 1990s. As a further case in point, Florida criminalized cyberstalking with Florida Statute 784.048(1)(d) in 2003 [11].

This paper describes the Predator and Prey Alert (PAPA) system that consists of a set of integrated tools designed to support law enforcement in helping victims of cyberstalking, facilitate the investigation of such crimes, and maintain evidence for subsequent prosecution of the perpetrator. An alpha version of the system, to be used for testing and evaluation, has been built, and a second, more complete beta version is under development to incorporate feedback from testers of the alpha version. In the next section, we give an overview of the architecture of the PAPA system, covering both hardware and software. In section 3, we describe the modes of use and the user interface of the PAPA system. In section 4, we discuss the security issues that are germane to the use and effectiveness of PAPA. In section 5 we give an overview of related work in combating cyberstalking.

2. PAPA architecture

The PAPA system was designed with the following goals in mind:

(i) permit law enforcement to remotely “shadow” a victim and provide assurance and advice when needed;

(ii) capture and log appropriate data related to cyberstalking activities so that a law enforcement analyst can subsequently investigate and determine the identity of the predator; and

(iii) capture evidence of probative value so that the predator can be successfully prosecuted.

Thus the design of the system supports observation, capture, indexing, and preservation of the integrity of collected evidence, and provides a WorkBench suite of tools for the forensic analysis of captured data.

The system environment of PAPA is illustrated in Figure 1. It is assumed that the victim has
a primary communications channel with the predator via the internet. A *Session* is the period of time from when a victim logs in until the time the victim logs out from the Computer Under Investigation (*CUI*), when the PAPA system records the actual desktop experience of the victim.

A hardware *Session Recorder* is connected to the CUI either through a USB or an Ethernet connection. This approach has several advantages over collecting data directly on the CUI or sending video through slower channels. First, it is good practice to minimize, as far as possible, any modifications to the CUI. Second, software solutions are potentially more insecure and susceptible to manipulation, and evidence collected by them may be less reliable. Third, it is much easier and less disruptive to transfer the large-capacity, dedicated disk from the Session Recorder to the analysis laboratory while preserving the necessary chain of custody. Fourth, private and sensitive data on the CUI, irrelevant to the investigation, is not compromised. Another solution might be to transmit captured video and other data to a remote agent via the Internet, but this may not be practical in low bandwidth environments such as dialup. In general, it is important that the management of captured data be separate and independent of the other functions of the CUI.

An independent second channel is used to communicate with the law enforcement agent. This channel could be via phone line, another high-speed internet connection, cellular, wireless, satellite, etc. Authorization and case coordination is achieved through communication with a remote *Dispatcher*, which informs the agent when a victim requests assistance, authenticates the agent’s connection to the CUI via the Session Recorder, and continually monitors the state of the Session Recorder through a “heartbeat” protocol.

### 2.1. Information flow in the PAPA system

We assume that the victim is connected to the Internet through a high bandwidth connection or possibly a dial-up connection and receives cyber threats in the following way:

(i) the victim is using instant messaging or other real-time means to communicate with an individual (most likely known to the victim) and the cyber threats occur during these exchanges;
(ii) the stalker harasses the victim in a chat room or other discussion group; or
(iii) the victim checks email and the threat is contained in one or more email messages.

For completeness, PAPA captures all the video information from the framebuffer on the victim’s machine during the session. It also captures other meta-information related to this session as we will discuss later.

In order to capture information, the Session Recorder is connected to the CUI via a USB2 connection, or via Ethernet to the LAN, where available. This high-speed connection permits us to capture the video information in “raw” mode yielding a high-resolution video image of activity on the CUI during the session. The session recorder also captures keystrokes made by the victim during the session and metadata for subsequent analysis. Both video and keystroke data are time stamped. Optionally, metadata captured by filtering communication to the victim’s machine such as IP header information and predetermined auxiliary textual and temporal information in the packet data such as target words, screen names, email addresses, avatars, etc. are also stored and used to index the video files.

The Session Recorder has a second logical channel to the agent and Dispatcher. This second channel is secured using a Virtual Private Network (VPN) tunnel. The second channel is logically viewed as an independent channel, for example, using a wireless connection or a second high-speed connection. The purpose of this second channel is to hide this traffic from the predator, who communicates on the primary Internet connection channel with the victim and may have the technical ability to detect unusual traffic over the primary channel. If necessary, however, the second channel could be implemented over the primary connection, with the caveat that bandwidth may suffer for the victim over the first channel and with the increased possibility of the predator detecting the communication.

The recorder is connected to the Dispatcher to keep track of its status, and the Dispatcher also establishes a connection between the agent and the victim. The channel between the agent and the victim operates transparently through the Session Recorder and can be viewed as over the independent second channel.

The channel between the agent and victim is primarily implemented through a customized version of Virtual Network Computing (VNC) [12] open source software. VNC supports a wide variety of remote viewing and control modes between the remote desktop of the CUI and the agent. It requires a client running on the agent’s machine (the Agent Module), and a modified VNC server (the Victim Module) running on the CUI. The VNC based connection permits law enforcement to view exactly what is happening on the victim’s machine, as well as take control of the CUI when necessary. PAPA further implements an additional “chat channel” between the agent and the victim to allow the agent to interact with the victim independent of the communication between the victim and the predator. This chat channel is also implemented through the independent second channel and the Session Recorder.

2.2. PAPA components

2.2.1. Session Recorder: The Session Recorder lies at the heart of the PAPA system. In our current prototype, the Session Recorder connects to the victim’s machine via a
special USB2 cable that allows transfer rates of up to 400 Mb/s. The prototype Session 
Recorder is built from off-the-shelf components and consists of a dedicated Secure Box 
with two hard disk drives and a micro-ITX motherboard. This board has dual Ethernet 
connections for versatility and 512MB of system memory, a dedicated 40GB hard disk 
for the OS and configuration storage, with all recorded data stored on a dedicated, 
removable, and lockable 300GB IDE hard disk. This modular design reduces the load on 
the CUI, and allows the evidentiary disk to be taken to the lab for analysis and replaced 
with a new disk when necessary.

The PAPA Session Recorder uses a customized version of Security Enhanced Linux. 
No unnecessary access is allowed to the processes, services, or ports of the recorder, 
based on best practices, separation of duties, and strict application of IP firewall rules. 
The session recorder box has tamper-proof case security, with vandal-proof case screws 
and a lockable hard drive caddy. The session files stored on the Session Recorder are 
cryptographically hashed to establish the integrity of the data. Thus the session recorder: 
(i) records all video information displayed on the victim’s machine, 
including communication between the agent and victim, between the victim and 
predator, and between the Session Recorder and the Dispatcher; 
(ii) adds relevant metadata such as time stamps and potentially relevant 
TCP/IP packet header and data information; 
(iii) indexes suspected attacks within the potentially large video files created; 
(iv) preserves the integrity of the evidence with robust storage, verification 
and access control mechanisms; and 
(v) coordinates network communications between the victim’s computer and 
law enforcement systems by means of an independent secure second 
communications channel.

2.2.2. Victim module: The victim software module includes a VNC server and access 
control and authentication software that implements other functions required by the 
PAPA model. Components of this module detect when the victim logs on, and only then 
initiate recording for that session. To insure privacy, once a recording session is 
activated, the PAPA software provides a continuous indication of its operation to the user 
via blinking icons and message screens. It is assumed that an agent is not always 
available to monitor or assist the victim. Thus, the victim module indicates to the victim 
whether a law enforcement offer is available, and provides a mechanism to alert law 
enforcement that monitoring is required. Furthermore, the victim is also informed when 
law enforcement has assumed control of the session, and the chat channel is used by the 
victim to interact in real-time with the law enforcement agent.

We are currently assuming that the victim is running a modern Windows Operating 
System such as 2000 or XP. In normal operation, all software is installed via CD or 
pushed from the session recorder after the USB connection is established. Optionally, the 
law enforcement agency could supply the victim with a dedicated computer to ensure 
absence of viruses, spyware, etc.

2.2.3. Agent module: The agent software module allows an investigator to monitor the 
CUI remotely, communicate with the victim via chat technology, and optionally assume 
control of the victim’s computer. The authentication protocols in the PAPA system
ensure that only authorized agents can monitor a session. Using features of VNC, PAPA can be deployed across multiple agents and jurisdictions and operates optimally through the independent second channel that allows the agent and victim to interact with reduced risk that a predator might detect that the victim is being monitored. The agent module can also incorporate a keystroke logger to provide additional evidence of the agent’s activities.

2.2.4. Dispatcher: The Dispatcher is a software module, hosted on a remote server, that authenticates connections between PAPA components, monitors the state of the Session Recorder, informs agents of user requests for monitoring, and uses a database to maintain law enforcement case file data on victims. The dispatcher can coordinate multiple agents, even from different jurisdictions, as active or passive observers during sessions.

3. Using the PAPA system

The PAPA system is primarily designed for use in law enforcement applications, but it could also be used to record the activity of computers in other scenarios.

3.1. Modes of use

For convenience, we consider the following modes of using the PAPA system, using some illustrative examples.

- **Passive**: a computer (say, of a parolee) is continuously monitored by a probation officer who maintains exclusive administrative access to the parolee’s computer and can monitor and track all activities of the parolee’s use of the CUI.
- **Active**: an agent masquerades as prey in order to attract the attention of online predators.
- **Consensual**: the parent, legal guardian, or victim contacts law enforcement and gives permission to use the PAPA system to remotely shadow ongoing interactions between the stalker and the victim.
- **Non-consensual**: a customized network version of the PAPA system is covertly deployed on the CUI, and law enforcement monitors or controls it. For criminal investigations, we assume that a court order for a Title III wiretap has been secured by law enforcement prior to installing the system on the prey’s machine.

Note that both active and passive modes can be used in consensual and non-consensual modes. Our work primarily focuses on the consensual mode and this is reflected in the basic PAPA configuration, but the design is flexible to allow for unforeseen applications.

3.2. Victim and agent interfaces

A sample operational interface from the victim’s perspective is displayed in Figure 2. The
PAPA victim module is designed for full disclosure of its monitoring operations. When the victim logs on and session recording starts, splash screens indicate this. The PAPA installation process mounts a link (1) icon on the victim’s desktop. On startup, the PAPA Status box (5) and system tray icon (6) appear to give continuous indication of PAPA’s operation. The first IM window (2) and the email window (4) represent two of the possible modes of communication between the predator and prey. The second IM window (3) is for direct communication with the agent.

Figure 3 illustrates the agent module interface. When the agent starts monitoring a victim, the PAPA icon (1), the victim’s desktop, and all running applications on the CUI, as seen in Figure 2, appear within a window on the agent’s machine. The victim’s entire desktop (2) appears here within a browser, although a native, non-browser mode is also available. The PAPA Status Box (3) remains visible as a reminder of monitoring activity. The side channel IM window (4) corresponds to one of the IM windows on the victim’s desktop. A system tray icon (5) is provided for additional convenience.

3.3. Law enforcement use
The PAPA system can be used by law enforcement agents for:

**3.3.1. Incident response and live analysis:** The PAPA system enables law enforcement agents to maintain constant remote vigilance, as needed, over victims who are involved in cyberstalking cases. In active mode, PAPA will establish an ongoing chat-based communication stream between the victim and investigator. At any time during a session, the investigating agent can seamlessly and unobtrusively take over the desktop. Even when the agent is not present, the Session Recorder is maintaining a complete record of the victim’s display and keystrokes, in addition to the metadata. The secure recording is used to not only record evidence but also to protect the agent because all actions of the agent are also recorded, including when the agent took control of the victim’s machine.

**3.3.2. Event reconstruction:** Using the PAPA WorkBench suite of analysis tools, investigators can reconstruct the victim’s experience during a cyberstalking attack. PAPA can playback the evidence in a non-technical simple format (time-stamped video), which can be viewed anywhere – from the laboratory to the courtroom – in order to show unambiguously the intrusion, intimidation, and abuse of the victim by the alleged stalker. The PAPA system records the actual video stream sent to the victim’s monitor and ensures the integrity of the evidence through robust hashing of the session files. The video/audio evidence stored on PAPA can be played back in both linear and fast-forward modes. Similarly, PAPA allows direct skipping to an event that has been earmarked by the victim and it allows the retrieval of an incident via metadata gathered during the recoding of a session. Because of the size of the video files created and the time needed to view them in real time, PAPA allows the victim, keylogger, or agent to flag incidents of interest as they happen, so that analysts can save time and effort. The metadata that PAPA captures for quick analysis is consistent with privacy restrictions by capturing information only after the content has been actually delivered to the victim's machine (at the application level), and when applicable by intelligent sniffing of TCP/IP packets for predefined data. Capture of the packet headers (and their timestamps) provides valuable circumstantial evidence about the source, path, and chronology of stalking incidents that can help prosecutors reconstruct patterns that bind online activity with its human sources.

PAPA combines both proactive and reactive approaches to cyberstalking. Proactively, it collects evidence as the crime is unfolding. This is facilitated because stalking is a repetitive, “power” crime in which the predator usually makes little or no attempt to conceal his or her identity. PAPA allows both the cooperation and the direct participation of the victim when it might increase the likelihood that the predator will be apprehended. On the other hand, using the active mode, an agent is able to block the victim and assume his or her online identity to avoid potential sympathy for the stalker, as in the “Stockholm syndrome.”

**3.4. Non-consensual mode**

PAPA can also operate non-consensually to allow agents, (e.g. probation, parole), to remotely monitor the online activity of the CUI. In this approach, we assume that the CUI has a high-speed connection to transfer video traffic to a remote Session Recorder. Note
that the action of the Session Recorder is the same as in consensual mode, except for possible reduction in the resolution of video that can be sent to the recorder. The modified victim module in this mode is termed the Covert Module. Note that the Covert Module would operate clandestinely without informing the victim of its presence, and keystroke logging and text capture become more important to the evidence gathering procedure. As much as possible, PAPA hides the operation of the covert module by using techniques such as “Trojan Horse” style installation, hiding evidence of the module’s operation from the Windows Task Manager and Registry, writing hidden, encrypted log files, and obfuscating network traffic.

It is not reasonable to expect that all covert module activity will go unnoticed by a technologically astute victim, even when the hardware Session Recorder next to the CUI is not used. A suspicious parolee under investigation, for example, might detect the increased network traffic used to stream video to the remote session recorder. Techniques to reduce network traffic include lowering the resolution, stopping or local buffering of the video stream, and relying primarily on textual data from the keystroke logger.

Truly covert operation presents other technical challenges for investigators in the form of firewalls, NAT translators, etc. For example, firewalls may necessitate the use of more sophisticated techniques such as reverse tunneling, a process that can increase the computational load on the CUI.

4. Security considerations

Because the PAPA system combats Internet crime and is distributed physically over the Internet, it is useful to view its components with respect to their logical security perimeters. The CUI and the Session Recorder derive their security partially from the physical custody of the victim, but the CUI must be free from existing malware. Because of this, the consensual operation of PAPA is dependent on the cooperation of the victim with the investigators. The Dispatcher and Analysis Console are physically within the custody of law enforcement and need to adhere to all relevant internet security policies. The Dispatcher is also potentially connected to multiple Recorders and agents in the field via the Internet, so it must also be protected by firewall technology while being always accessible to the Session Recorders and agents in the field.

4.1. Integrity

To ensure the highest quality of collected evidence, and to minimize the possibility of corruption or tampering, it is desirable to include a mechanism to verify the integrity of all data recorded. The PAPA Session Recorder receives its input via USB or Ethernet connection with the CUI. The aggregate traffic (video, text, etc.) is recorded as one large time-delimited session file. The Session Recorder performs a cryptographic hash on these files as each session concludes, and the hash value is securely stored (either locally, remotely, or both) in a case database as an integrity reference to the state of the recently recorded session. Analysts who examine these files for evidence can use this hash value to check the integrity of the copy of the file under analysis.

4.2. Authorization
Because PAPA is intrusive software that allows the agent to completely monitor and control the victim’s computer, it compromises the confidentiality, integrity, and accessibility of the CUI. Thus, agents must show restraint in the use of PAPA, and victims must consciously permit the intrusion caused by a PAPA session. The Session Recorder records all actions of the agent, directly through framebuffer recording and indirectly by capturing all TCP/IP communication and VNC control traffic between the victim and the agent.

Because of confidentiality concerns surrounding this intrusive technology, PAPA will display clear indicators of its status to the victim and agent in consensual mode. PAPA session indicators include: a system tray icon with indicator lights displaying system status, a pop-up status text box, and a control panel that launches from a right-click on the tray icon. This control panel contains additional details of system configuration, status, and options for the victim. Another key feature of the consensual mode is the chat channel, which allows the victim to directly communicate with the agent. It also gives an indication of the participation and shadowing activity of the agent. Note that in non-consensual mode, court authorization may be required to install such intrusive software and hardware. Non-consensual software gives no indication of its operation. In fact, it makes every attempt to obfuscate its process and minimize communication traffic.

4.3. Authentication

PAPA requires authentication for all parties involved in every phase of its deployment, operation, and analysis. Each PAPA module requires duty-based authentication using certificates and a Kerberos authentication scheme. Certificates are generated and exchanged between PAPA components when a new case is initiated. In active mode, the Dispatcher is central to the authentication process. It constantly monitors the Session Recorder’s state by listening remotely for a “heartbeat,” and allows only agents officially assigned to the case to connect as VNC clients for shadowing and remote control.

The Session Recorder is central to all connectivity and PAPA sessions, and agents and administrators must authenticate themselves within the law enforcement trust domain. For example, the Session Recorder will not record sessions if the Windows GUID does not match that of the victim involved in the case. Thus it can be used on CUIs with multiple users and still maintain the privacy of other users.

4.4. Encryption

PAPA sessions are recorded locally and there is no compelling need for encryption of the video data over the local connection between the CUI and the Session Recorder. Capture of TCP/IP header data during a session, when permitted under current Pen/Trap limitations, can also be recorded in the clear. Title III and Pen/Trap statutes determine the type of data that investigators can record during live wire communication. If the Session Recorder is used to capture and record all TCP/IP packets, including payloads, these may be encrypted to prevent human readability. Because the agent and Dispatcher are typically accessible only via an insecure network (the Internet) all traffic between these
components and the Session Recorder must be encrypted, typically using a Virtual Private Network (VPN).

5. Related and previous work

PAPA draws its functionality from many sources, including the Remote Frame Buffer (RFB) protocol, VNC, tcpdump, chat, and VPNs. VNC and other similar tools have traditionally been used for remote access, training, and administration. The novelty of our work is the automation and control mechanisms that we introduce to apply this previous work to combat cyberstalking by capturing, indexing and preserving the integrity of data. PAPA is designed with a well-defined trust model to be a system that is simple to use and robust enough to be deployed and operated widely.

There are several other existing open-source and proprietary tools that could be useful in combating cyberstalking. These range from the notorious “Trojan Horse” style rootkits and spyware, to other more benign tools developed by hackers. PAPA makes no use of tools from such sources.

Several commercial applications have been developed to monitor, analyze, filter, and block undesirable, predatory, and sexually explicit computer traffic [13]. These tools are designed to be used by parents or guardians, but as far as we are aware, data they collect is not designed with evidentiary value in mind. Some of these tools have also been used in parole and in probation scenarios. In addition to these stand-alone monitoring tools, private services are also becoming available for victims of cybercrime. These include technical [14], and legal [15], and practical support services [16].

6. Conclusion and further research

Research is ongoing into the effects of predator and prey relationships, the responses of law enforcement, and the toll they take on victims and families. The anonymity and reach of the Internet has changed the dynamic by providing stalkers new opportunities to intimidate their victims. Electronic evidence is central to the detection, analysis, prevention, and prosecution of cyberstalking cases, and the authors are developing the PAPA system to address these new phenomena. An alpha version of the PAPA system has been built that incorporates most of the functionality discussed. This is being used for testing and evaluation purposes. Based on feedback, a beta version, incorporating user suggestions, full dispatcher functionality, security enhancements, and non-consensual capabilities is being developed.

PAPA gives law enforcement, parents, and businesses the ability to detect, record, verify, and play back evidence of cyberstalking or other computer activity. Although our prototype Session Recorder, for example, is notebook sized for development and testing and uses off-the-shelf components, it could easily be miniaturized and streamlined for a smaller footprint, or deployed via Ethernet in a LAN for clandestine operation. PAPA has been designed to be flexible in its deployment, and, although primarily targeted toward law enforcement applications with the aim of producing quality evidence, it could also be used to monitor desktop experiences and document the activity observed.
Acknowledgements

The authors wish to thank Bob Breeden and the Florida Department of Law Enforcement’s Computer Crime Center for their support and help in this project. We would also like to thank Mark Gage, Wade Grant, and the National White Collar Crime Center for their helpful input. The authors also thank Alec Yasinac, David Kelley, Sahil Cooner, Newton Etienne, Sandeep Kalburgie, and Kim Vogt for their help in the software design and development.

7. References


REFERENCES


[23] Melissa Kryder, "Predator and Prey Alert System" (Masters thesis supervised by Dr. Mike Burmester.)


[34] , “”, Network Working Group, The Internet Society,
BIOGRAPHICAL SKETCH

I have attended Indian River Community College, Palm Beach community College, the University of Central Florida, Pennsylvania State University, and Florida State University; and I have been awarded an Associate of Arts degree from Indian River Community College, and a Bachelors of Science degree from Florida State University. I am certified as a NSTISSI No. 4011 Information Systems Security Professional, and I have worked in the following research groups at Florida State University: the Security Assurance in Information Technology Laboratory, and the Florida Cyber Security Institute.