LLNL-TR-419023

Probabilistic Characterization of Adversary Behavior in Cyber Security

FY2009 SMS Project
Final Report

Carol Meyers, Sarah Powers, and Daniel Faissol

September 2009

This work performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344.

Probabilistic Characterization of Adversary Behavior in Cyber Security

FY2009 SMS Project
Final Report

Probabilistic Characterization of Adversary Behavior in Cyber Security

FY2009 SMS Project
Final Report

Abstract

The objective of this SMS effort is to provide a probabilistic characterization of adversary behavior in cyber security. This includes both quantitative (data analysis) and qualitative (literature review) components. A set of real LLNL email data was obtained for this study, consisting of several years' worth of unfiltered traffic sent to a selection of addresses at ciac.org. The email data was subjected to three interrelated analyses: a textual study of the header data and subject matter, an examination of threats present in message attachments, and a characterization of the maliciousness of embedded URLs.

1. Introduction

1.1 Motivation

Cyber defense is a vast and growing problem in national security. According to the FBI, the annual loss due to cyber crime was estimated at $67.2 billion for US organizations in 2005 (GAO, 2007). Numerous efforts have sought to quantify the impacts of cyber crime (Rantala, 2008; ICCC, 2008), but much less work has focused on characterizing the cyber adversaries themselves. Given that cyber security is such a vast problem, it is essential in constructing a defensive architecture to know who the cyber adversaries are and what kinds of threats they are likely to attempt.

The goal of this SMS effort is to provide a probabilistic characterization of adversary behavior in cyber security. This involves not only quantifying the number of cyber threats, but *who* is perpetuating these attacks, and *what methods* they use to bypass security protocols. Our study is both quantitative and qualitative in nature, exploiting the literature in cyber security to address attack vectors beyond those for which we were able to obtain real data.

1.2 Overview of the SMS Project Thrust Areas

The SMS project encompassed three main quantitative thrust areas and one qualitative study. The quantitative areas centered on analysis and characterization of the email dataset described in Sections 1.3 and 3. These thrust areas, and the sections within which they are covered, are as follows:

- Characterization of Textual Email Data (Section 3)
- Characterization of Viruses Present in Attachments (Section 4)
- Characterization of Malicious URL Content (Section 5)

The first of these addresses the textual and descriptive content of the emails themselves: the volume of emails over time, the distribution of countries of origin, subject matter, and methods of spoofing the header data. The second thrust area examines the content of the email attachments, obtaining distributions of the kind and frequency of attacks sent via such attachments. The third quantitative thrust area characterizes the content of web addresses embedded as URLs within the emails, focusing on malicious content and including a comparison of online website malware detection tools.

In addition, a qualitative study surveyed the published literature to create taxonomies of cyber adversaries and attack methods. The goal of this work was to provide a foundation for the quantitative work, in terms of the adversary types we might hope to identify, and also to extend the effort to attack vectors beyond those for which we were able to obtain quantitative data. Findings from this study are fully described in a companion report, entitled "Taxonomies of Cyber Adversaries and Attacks: a Survey of Incidents and Approaches" (Meyers, Powers, & Faissol, 2009). An overview of the content in this report is presented in Section 2.

1.3 Data Sources

The primary data source obtained for this effort was a set of unfiltered email data, from a selection of addresses at ciac.org, the former Computer Incident Advisory Capability at LLNL. For a nearly 20-year period (1989-2008), the CIAC team was responsible for incident response, reporting, and tracking of cyber threats at LLNL and several other sites within the US Department of Energy (Schultz, 1990). As part of their public presence, the CIAC team maintained an externally accessible website with information on reporting cyber incidents, security bulletins, and how to recognize and report internet hoaxes (hoaxbusters.ciac.org).

The addresses associated with our email dataset were principally affiliated with the hoaxbusters portion of the website. There were three primary addresses: webmaster@ciac.org, hoaxmaster@ciac.org, and hoaxbusters@ciac.org. The time period of the data collection was from January 2004-December 2008 and June-July 2009, with a monthly traffic of between 2000-7000 messages per month. This traffic represents both legitimate queries and spam, malware, and phishing emails; of these, legitimate queries represented a very small minority of the data. Thus, our analysis is primarily focused on characterizing this unsolicited, and possibly malicious, email traffic. A detailed description of the email data, including relevant statistics and characteristics, can be found in Section 3.

2. Types of Cyber Adversaries and Attack Methods

2.1 Cyber Adversaries

The study of cyber adversaries was initiated in the early 1980's, when personal computers began to come into the mainstream. The term 'hacker' first entered the lexicon to describe a person skilled at programming, and was modified to describe an individual engaging in malicious activity, following the arrests of computer criminals such as the infamous '414 gang' (Murphy et al., 1983; Raymond, 2003).

By the mid-1980's, there was a growing interest on the part of law enforcement in cyber crime, culminating in the first legislation aimed at prosecuting cyber adversaries (Eltringham, 2007). Profiling studies of cyber adversaries suddenly became important as a method of identifying these criminals and determining their modes of operation (Smith & Rupp, 2002). The work of Landreth (1985), Hollinger (1988), Chantler (1996), and Rogers (1999, 2001, 2006) was particularly influential in furthering the understanding of cyber adversaries and their motivations.

A taxonomy of cyber adversaries is given in Table 2.1, which is reproduced from the companion report to this one (Meyers, Powers, & Faissol, 2009). This table represents an amalgam of research in the area, most heavily influenced by the work of Rogers (2000, 2006).

Adversary Class	Skills	Maliciousness	Motivation	Method
script kiddies, newbies, novices	very low	low	boredom, thrill seeking	download and run already-written hacking scripts known as 'toolkits'.
hacktivists, political activists	low	moderate	promotion of a political cause	engage in denial of service attacks or defacement of rival cause sites
cyber punks, crashers, thugs	low	moderate	prestige, personal gain, thrill seeking	write own scripts, engage in malicious acts, brag about exploits
insiders, user malcontents	moderate	high	disgruntlement, personal gain, revenge	uses insider privileges to attack current or former employers
coders, writers	high	moderate	power, prestige, revenge, respect	write scripts and automated tools used by newbies, serve as mentor
white hat hackers, old guard, sneakers	high	very low	intellectual gain, ethics, respect	non-malicious hacking to help others and test new programming
black hat hackers, professionals, elite	very high	very high	personal gain, greed, revenge	sophisticated attacks by criminals/thieves; may be 'guns for hire' or in organized crime
cyber terrorists	very high	very high	ideology, politics, espionage	state-sponsored, well-funded cyber attacks against enemy nations

Table 2.1: A Taxonomy of Cyber Adversaries (Meyers, Powers, & Faissol, 2009)

In this taxonomy, there are eight different classes of adversaries, which are arranged in increasing order of skills and sophistication. In general, the maliciousness level of the adversary groups is proportional to their sophistication, with the major exception of white hat hackers, who are not intentionally malicious.

The least sophisticated of the cyber adversary groups are the script kiddies, who have limited programming skills and rely on pre-written scripts known as 'toolkits' in their exploits. Their overall maliciousness tends to be low, primarily due to their limited skills; however, with the increasing sophistication of toolkits, their ability to pull off large-scale attacks is on the rise, as in the case of the denial-of-service attacks perpetuated by 'Mafia Boy' in Canada (Rogers, 2006). The next most skilled group is that of hacktivists, who are motivated by political causes rather than personal gain. Their attacks tend to be focused against specific rival organizations, such as the email bombs used by the Internet Black Tigers in Sri Lanka, to gain publicity for the Tamil Tigers (Denning, 2001).

Most commonly covered in the press is the class of cyber punks, who have similar motivations but greater programming skills than individuals in the novice category. These hackers seek attention and occasionally mature to become security consultants, as in the case of Kevin Mitnick, who served five years in jail following numerous intrusions into restricted computer systems (Mitnick, 2002). Even more dangerous is the group of insiders, who represent the greatest risk to companies and who are most commonly motivated by revenge against their current or former employer (Kowalski et al., 2008). These criminals often seek to sabotage systems, such as the logic bombs planted by the disgruntled employee Michael Lauffenberger, and as such have the potential to cause a lot of damage (Shaw et al., 1998).

Groups at a high level of skill include coders and white hat hackers, who both have the power to construct sophisticated scripts. Coders write the toolkits that are used by the script kiddies, and they are motivated by power and prestige as well as a sense of mentoring younger hackers (Rogers, 2006). The white hat hackers are the only non-malicious group, who consider themselves ‘purists’ and engage in the intellectual challenge of testing security systems; we include them for the sake of completeness. These individuals are commonly hired as security analysts to test cyber defenses, and the National Security Agency even offers a certification course in such ‘ethical hacking’ activities (Taylor et al., 2006).

The most dangerous and sophisticated groups of cyber adversaries are the black hat hackers and cyber terrorists. Black hat hackers are professional criminals, who are motivated by money and greed and use their hacking to support themselves financially. Such adversaries may commonly be employed by organized crime; unfortunately, while they represent some the most malicious hackers, they are also the group about which the least is known (Rogers, 2006). Cyber terrorists engage in state-sponsored warfare on information technology, performing attacks that try to disrupt and destroy the cyber assets of an enemy nation. Examples of such activities include massive distributed denial-of-service attacks against Estonia (in 2007, following the removal of a Russian World War II monument), and the Republic of Georgia (in 2008, preceding the conflict between Russia and Georgia) (Landler & Markoff, 2007; Markoff, 2008).

2.2 Cyber Attacks

The formal study of cyber security began in the mid-1970’s, when computers first became installed in government and universities. Most of this early work in cyber security did not explicitly consider the different types of attacks that might be performed. The first studies of cyber attacks in particular arose in the early to mid-1980’s, around the same time that cyber adversaries first entered the public eye. Stoll (1986) describes the methods a German hacker used to break into computer systems at Lawrence Berkeley National Lab, and how researchers there used the hacker’s activities to track him. The Computer Emergency Response Team (CERT) was founded by DARPA in 1988, after a high-profile crippling of the internet via the Morris worm (Scherlis, 1988; Kehoe, 1992).

One of the major issues in studying cyber attacks is that the notion of an “attack” itself is very broad: it can encompass attack vectors, operating systems, hardware and software targets, access schemes, attacker objectives, specific implementation and design vulnerabilities, and the attack payload (Howard, 1997; Hansman & Hunt, 2005). Efforts to classify the space of cyber attacks have typically focused on case studies of cyber security incidents, using some subset of attack characteristics. Notable work in this area includes the studies of Landwehr et al. (1994), Howard (1997), Howard and Longstaff (1998), Hansman and Hunt (2005), and Kjaerland (2006).

A taxonomy of cyber attacks is given in Table 2.2, which is reproduced from our companion report (Meyers, Powers, & Faissol, 2009). It is most heavily influenced by Hansman and Hunt's (2005) work.

Attack Class	Subtypes	Description
viruses	file infectors, system/boot record infectors, macros	self-replicating program that replicates through infected files; attached to an existing program
worms	mass mailing via botnets, network aware	self-replicating program that replicates through networks or email; no user interaction required
trojans	remote access, data destruction	program made to appear benign that serves a malicious purpose
buffer overflows	stack-based overflows, heap-based overflows	process that gains control or crashes another process via buffer overflowing
denial of service	host (resource hogs, crashers), network (TCP, UDP, ICMP flooding), distributed	attack that prevents legitimate users from accessing a host or network
network attacks	spoofing, web/email phishing, session hijacking, wireless WEP cracking, web application attacks	attack based on manipulating network protocols, against users or networks
physical attacks	basic, energy weapon (HERF gun, EMP/T bomb, LERF), Van Eck	attacks based on damaging the physical components of a network or computer
password attacks/ user compromise	guessing (brute force, dictionary attacks), exploiting implementation	attacks aimed at acquiring a password or login credential
information gathering	packet sniffing, host mapping, security scanning, port scanning, OS fingerprinting	attacks in which no damage is carried out, but information is gathered by attacker

Table 2.2: A Taxonomy of Cyber Attacks (Meyers, Powers, & Faissol, 2009)

This taxonomy includes nine different classes of cyber attacks, each of which contains several different subtypes. We note that many cyber incidents employ more than one of these attack methods, so they should not necessarily be viewed as mutually exclusive alternatives.

Two of the most prevalent kinds of cyber attacks are viruses and worms, which are both types of self-replicating programs. Viruses are spread via user execution of the virus code, and for this reason are often found attached to legitimate programs. The most destructive virus to date is the ILOVEYOU virus, a visual basic scripting exploit which caused 10 to 15 billion dollars of damage worldwide in the year 2000 (Jones, 2006). Conversely, worms do not require user interaction to propagate; they spread over networks and typically exploit vulnerabilities in operating systems. Worms commonly install a 'backdoor' on infected systems to allow remote control, as in the massive spam 'botnet' created by the Sobig worms in 2003 (Levy, 2003).

Trojan and buffer overflow attacks both masquerade as legitimate programs or processes, but which conceal a malicious purpose. A trojan is usually attached to a program that performs a real function, while secretly installing a ‘backdoor’ on systems to allow remote access. As opposed to viruses and worms, trojans are not self-replicating and rely on the distribution of their host program to spread, as in the 2008 distribution of the Mocmex virus via digital photo frames (Soper, 2008). Buffer overflows function by forcing a seemingly benign program to write more information into the buffer (temporary memory storage) than the space allocated to it, allowing alteration of local variables and the running of user-introduced code. These exploits are often used in conjunction with other attack methods, a technique used by the Code Red and SQL Slammer worms to force malicious code execution (Chen & Robert, 2004).

Attacks which function directly on networks include denial of service and network attacks. In a denial of service attack, routers or servers are made inaccessible to users by being overloaded with bogus requests for data. This “flooding” of requests is often distributed among many systems, and has been used to cripple high-profile government (Estonia, Georgia) and commercial (eBay, CNN) websites and accounts for days (Garber, 2000; Markoff, 2008). Network attacks function by manipulating network protocols to exploit others, and include IP spoofing, web and email phishing, session hijacking, and cross-site scripting attacks, all of which trick users into divulging private data or resources. These attack methods can also be used with other types of attacks, such as denial of service, and can be very costly: for example, an estimated $1.2 billion were lost in phishing attacks in the year 2003 (Emigh, 2005).

The last three types of attacks tend to be a bit more prosaic in nature, yet still have the potential to cause a good deal of damage. Physical attacks include the destruction of hardware using physical force, but can also involve the sophisticated manipulation of electromagnetic waves, as in HERF guns and EMP/T bombs, which fry a computer’s motherboard and other components (Schwartau, 1996). The US government’s TEMPEST component standards are designed to mitigate the risk of these kinds of attacks (Russell & Gangemi, 1991). Password attacks have the objective of gaining control of a particular system or user’s account, and can be based on social engineering or forms of dictionary search. In a recent study of MySpace passwords, fully 4% consisted of dictionary words, and another 12% were a word followed by a single number (Evers, 2006). Information gathering attacks involve scanning a network to determine what programs and operating system are used, and potential vulnerabilities; such actions are not inherently malicious, but are often used as a precursor to other attacks. Worms such as Sasser, Slammer, and Code Red used scanning as a method of determining hosts to compromise (Kikuchi et al., 2008).

3. Characterization of Textual Email Data

3.1 Volume and Time Frame

The CIAC email dataset consists of text files of emails in the form that they “came off the wire,” i.e., no formatting or changes have been made. Viruses, malware or other malicious vectors of attack may or may not be present in a given email; these emails were therefore handled with care on a computer dedicated for this purpose.

The data obtained were received between February 2004 and early December 2008, and also from June to July of 2009. The website maintained by the CIAC team was publicly accessible through January of 2008, at which point it was shut down as the Department of Energy transitioned their cyber response teams. The data on hand result from individuals contacting the webmaster or requesting further assistance; as expected, many non-legitimate emails were also received. Figure 3.1 shows the trend in email volume over time. A steady increase in the number of emails is seen, perhaps due to increased visibility of the website or simply a longer time in existence that allowed the webmaster's email to land on more "spam" lists. A sharp decline is seen starting in 2008, which corresponds to the CIAC website being shut down.

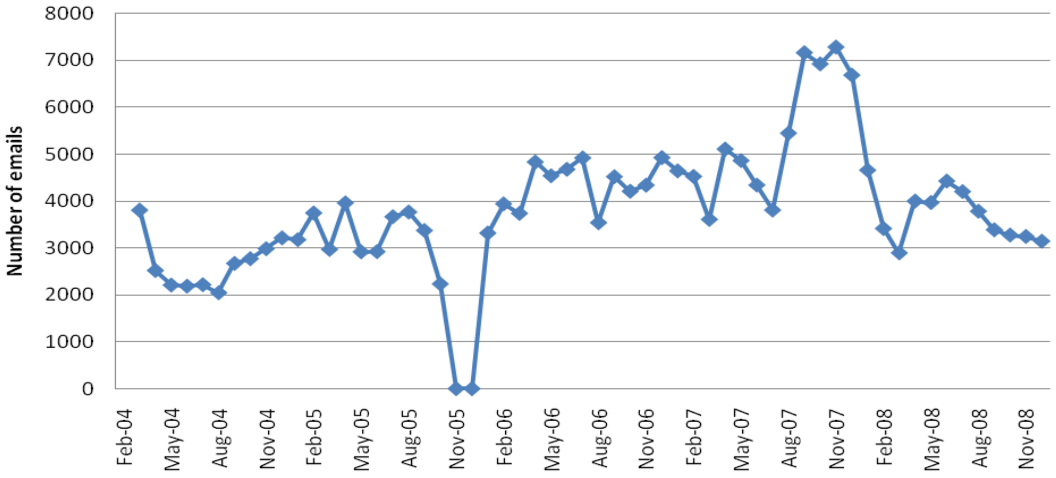

Figure 3.1: Quantity of Emails Received by Month, February 2004 – November 2008

It appears in general that the summer months attract more email traffic, while early and mid-spring time periods see a lull. Data were not provided for the months of October and November 2005, which accounts for the dip in the curve.

3.2 Countries of Origin

The countries from which the emails were sent were obtained by performing a reverse lookup on the IP addresses. (It should be noted that the IP addresses themselves may not be reliable, due to the potential for spoofing, and therefore this analysis is credible only to the extent that the IP addresses are valid.) The emails were either sent directly or relayed via several servers. The distribution of the number of places an email was sent before reaching the CIAC inbox is shown in Table 3.1.

Number of hops	Frequency	Percent of Total
0 (no data)	552	0.2485%
1	173348	78.05%
2	35931	16.18%
3	9749	4.39%
4+	2532	1.14%

Table 3.1: Distribution of the Email Routing

On average, an email traversed 1.28 servers before reaching its destination. Emails that traveled through multiple servers before reaching their destination were occasionally associated with multiple IP addresses, and thus potentially multiple "countries of origin." The term "origin" is used loosely, as it may or may not have been coming from the last "received" line indicated in the email.

Country Analysis 1: Final IPs in multi-routing emails or single hops

Figure 3.2 displays the percentage of emails originating from each country, in terms of the final IP address. Figure 3.3 gives a close-up picture of those countries corresponding to values greater than 0.5%. As can be seen from these graphs, over half of the email traffic originated from China and the United States.

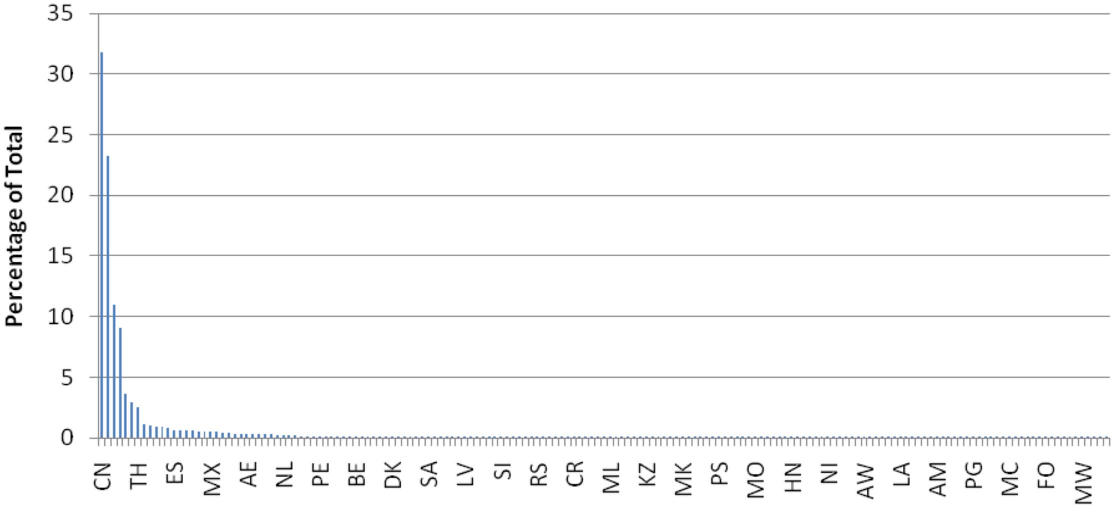

Figure 3.2: Countries of "Origin" for Final IPs in Multi-Routing Emails or Single Hop Emails

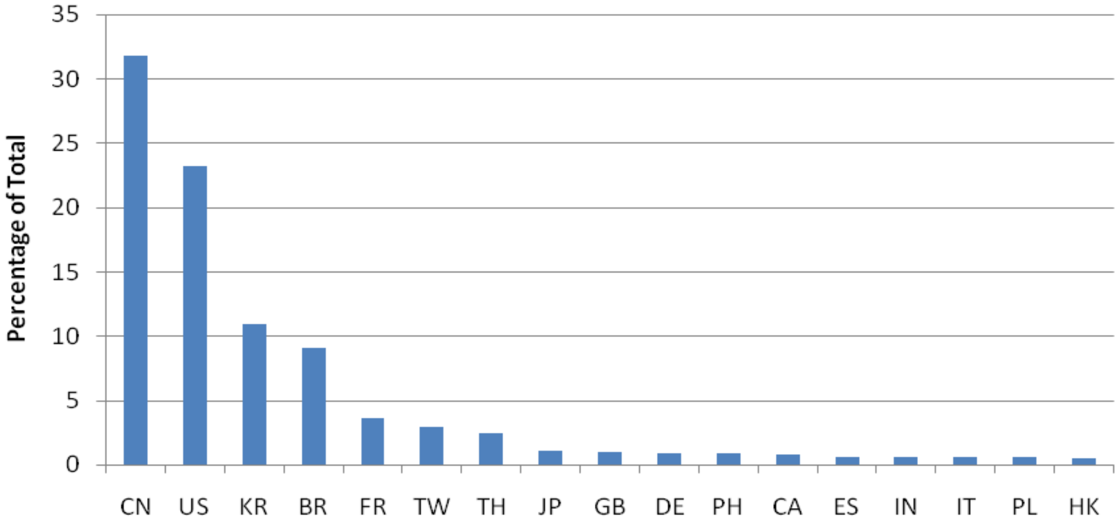

Figure 3.3: Countries of "Origin" for Final IPs in Multi-Routing or Single Hop Emails with Values > 0.5%

Country Analysis 2: Second to last IPs in multi-routing emails

Figure 3.4 shows the percentage of emails originating from each country in terms of the second to last IP. Here, a high percentage of the IPs has no mapping. In general, this corresponds to IPs that map to the "localhost" (for example: 127.0.0.1) or "private internets" (for example: 172.20.115.201). These usually correspond to machines that operate on an internal network without direct access to the WWW. To access the WWW, users on these networks must go through a router or a proxy system.

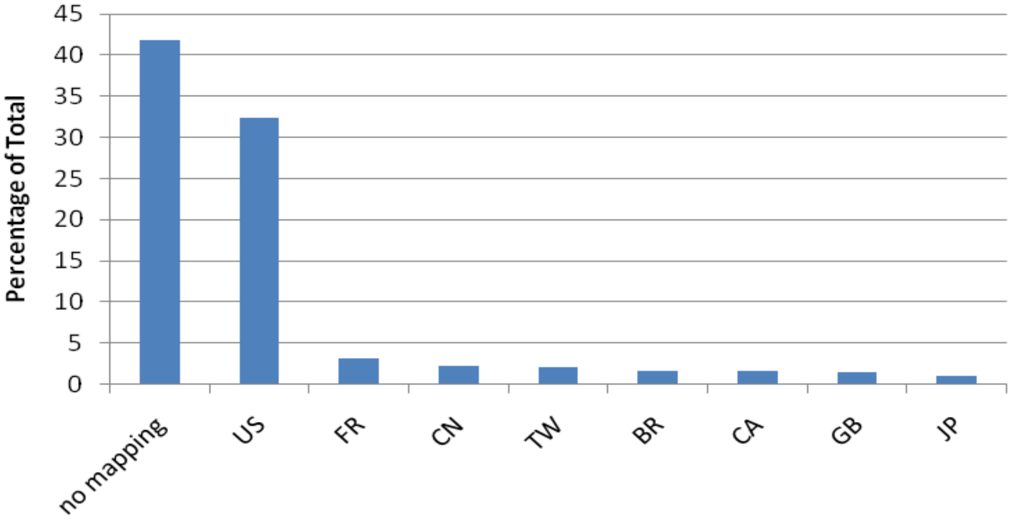

Figure 3.4: Countries of "Origin" for Case 2 IPs with Frequencies Greater than 0.5%

Country Analysis 3: Original IPs in single or multi-hop emails

Figure 3.5 indicates the countries of origin of the originating IP addresses, as specified in the email headers. As before, China and the US comprise the bulk of the data, with South Korea and Brazil following. Some origins cannot be determined, but this represents a relatively small proportion of the total emails.

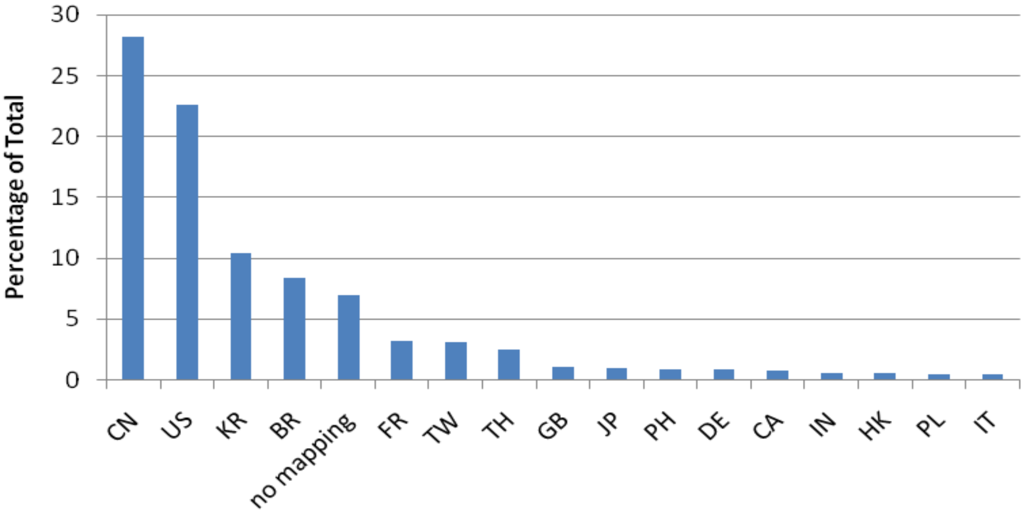

Figure 3.5: Countries of "origin" for case 3 with IPs that have values frequency greater than 0.5%

The emails themselves contain a plethora of information, and will be discussed in greater detail in the next section.

3.3 Methods of Spoofing Email Header Data

All data has limitations, which can be due to factors such as missing values or erroneous testing. The analysis in the current study is limited by the validity and veracity of the email data obtained. Because this study focuses on the cyber adversary, it is a given that in many cases these individuals will want to and attempt to fool the email recipients. In this section, we briefly review the different components of an email header and how these might be spoofed.

Email headers are often overlooked by most people or viewed as "garbage in my email." As a result, what can be a key source of information regarding spam is frequently disregarded, allowing spam to slyly slip through. Figure 3.6 shows an example email header.

```
From someone@somewhere.com Fri Mar 2 04:19:56 2005                       [1]
Return-Path someone@somewhere.com                                        [2]
Received: from autoturn.net.uk ([88.11.23.313])                          [3]
          by ciac.org (………) with ESMTP id 122CmhT21558                   [4]
          for webmaster@ciac.llnl.gov; Fri Mar 2 04:19:56 2005 -800 (PST) [5]
From: "Edgar" someone@somewhere.com                                      [6]
To: webmaster@ciac.org                                                   [7]
Subject: crazy                                                           [8]
Date: Fri Mar 2 04:19:56 2005 -800 (PST)                                 [9]
Message-ID: 01c75cc9@somewhere.com                                       [10]
```

Figure 3.6: Email header example

We now provide a brief explanation of the header file, along with possible methods for exploitation.

Line [1]: *From someone@somewhere.com Fri Mar 2 04:19:56 2005*

This line is to be distinguished from the traditional "From:" line, which includes a colon. It is not part of the actual email, but it is inserted by the mail transfer software when the email is received. One use for this is by UNIX mailers, who use the line to separate messages in a folder. This line is not always inserted, depending on the mail transfer software used. If it is, it will be the first line in the email header.

How it might be exploited: Although it is possible to forge this line, this is not always done, as well as it being possibly complex. Possibilities for exploitation include empty lines (i.e., the data is suppressed or unavailable) or the email looking like it is coming from the "postmaster" as a returned email.

Line [2]: *Return-Path someone@somewhere.com*

This is the address that error messages are supposed to be sent to (i.e., where an email is sent when it "bounces"). The email address here can be different from the one listed in the "*From*" line. For example, the address might be:

Return-Path: owner-somelistserv-l@LISTSERV.someschool.edu

This line is added at the receiving end by the server who makes the final delivery; as such, it can be more difficult to forge. Specifically, it is taken from the return path information given in the SMTP command MAILFROM, which is the "from" address of the "SMTP envelope."

How it might be exploited: The address could be forged, thus coercing people into thinking they are emailing a legitimate place while the information is actually being harvested for ill-use. Alternatively, the address may simply be forged to mask the identity of the sender.

Another related line is sometimes seen in addition to or instead of this one. It is the **Reply-to:** line. This is where replies are sent. One author claims this is "widely used by spammers to deflect criticism" (stopspam.org, 2008).

How it might be exploited: An illegitimate use could be to solicit replies in an email and have the responses sent to an "innocent" victim, making them the recipient of much spam, flooding their mail servers, and potentially causing a failure. The address could also be a simple garbage collection bin where responses get dumped and are never even examined.

Line [3] *Received:*

The **Received:** line indicates where the email came from and the path that it took to get to its final destination. A header may contain one or more of these lines, depending on the number of hops an email took to reach the recipient. In the case of multiple "Received:" lines, the top-most one is the most recent. To track the path taken by the email, the lines should be read in reverse order. Consider the following "Received:" line:

> **Received:** from hiyathere.com (alpha.truehost.com [111.208.141.212])
> **by** mail.truehost.com (8.10.2/8.10.2) with ESMTP id BQ0T3E2612;
> Wed, 6 Dec 2000 18:33:02 -0600 (CST)

In this example, the email was received from a server who claimed to be named "hiyathere.com" with IP address 111.208.141.212. The mail program did a reverse look-up and indicates in parentheses the true name of the sending machine associated with the IP, which is "alpha.truehost.com." Thus, it can be seen that forgery has occurred.

The values in parentheses are the "true" values; moreover, the IP address is always true. Thus, it is always possible to see where the email came from, and to decide potentially based on the content of the email whether it is spam or not. (For example, an email from your local bank that came from an IP address in Eastern Europe is probably a spoof.) Other available information in this header line includes the version of Sendmail being used (8.10.2/8.10.2), an internal ID number assigned to the message ("ESMTP id BQ0T3E2612") and the date and time the email was sent.

How it might be exploited: In Figure 3.7, an example received line is shown. Many of the email components may be spoofed; however, the IP address cannot. It indicates the IP address of the computer from which the email was sent. A spammer could try to mask this information by going through "relays" or hacking into machines and making the emails look like the hacked machine is the sender.

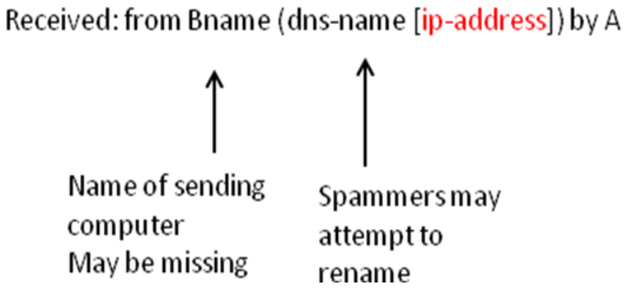
Received: from Bname (dns-name [ip-address]) by A
Name of sending computer
May be missing
Spammers may attempt to rename

Figure 3.7: A Sample “Received” Line from an Email Header

Line [6]: *From: “Edgar” someone@somewhere.com*

This is the “who the email is from” information, provided directly by the sender. In a program such as Microsoft Outlook, the user enters this information directly into their “profile.” As such, it is easy to falsify. For example, it is easy to send an email from “The President, iAmThePresident@whitehouse.gov” simply by entering this information before sending the email.

<u>How it might be exploited:</u> By falsifying this line, a user could be fooled into thinking they are receiving an email from someone they trust. As a result, they could choose to click on links or open attachments that they might otherwise have been wary of.

Line [8]: *Subject: crazy*

This is a sender-edited field and can be anything. It is usually one of the first clues other than the sender that the email might be spam or fake.

<u>How it might be exploited:</u> Similar to the “From” line, this could be used to lure the recipient into opening an email they might otherwise have put in the trash, as in junk mail received via the postal service.

Line [9] *Date:*

This specifies the date that the message was written (and sent).

<u>How it might be exploited:</u> The date itself may be forged, or the sending computer may simply be keeping incorrect time. Alternatively, a hacker or spammer may change the sending computer’s clock for devious purposes. This is how a message can be made to look as if it came “from the future.”

Line [10]: *Message-ID: 01c75cc9@somewhere.com*

The Message-ID is a unique identifier for each email of the format:

uniqueString@nameOfServerAssigedID

This could help identify spam or “hazardous” messages if the ID belongs to one domain or address (e.g., *me.com*), but the email’s sender belongs to another (e.g., *haha.com*). This would indicate that the sender is either using a fake address, or pretending that they own it. This line can also be forged. Another means of identifying forgery is if the ID contains an empty string or no @ sign.

Given the limitations presented in this section, available methods for analyzing the data are restricted to certain small pieces of the information provided by each email. Nonetheless, this work strives to provide as complete a characterization as possible of cyber adversaries who might use email as their attack vector. Further information on email headers and their spoofing potential can be found at stopspam.org (2008), the A3C connection (2000), and Digital Software Development (2009).

4. Characterization of Viruses Present in Attachments

Cyber adversaries choosing to use email as their attack vector can do this by exploiting one or more of the following venues: viruses (or other malware) in attachments and/or the email body, and phishing attempts in the email body. The first two are considered in this section, while section 5 explicitly considers threats present in URL content.

4.1 Methodology and Tools Used

Viruses, malware, worms and other methods of attack can be found both in email attachments and directly in the body of the email. In the latter case, simply previewing the email in a "preview window" can be sufficient to reconstruct (and possibly execute) the virus. (In this section, we use the word 'virus' in a generic sense, to encompass trojans and malware as well as traditional viruses.) The collected data was kept in text file format to avoid such problems.

To identify threats present in the emails or attachments, we employed a suite of (primarily free) tools, including the following products: Malwarebytes' Anti-malware, Super AntiSpyware, Spybot Search & Destroy, Windows Bitdefender, Norton Antivirus (Symantec), and AVG Free. Many other competent non-free commercial software products exist (e.g., Kaspersky, McAfee, etc.), which were not included in the test suite.

All emails were scanned for viruses, malware, Trojans, worms or other threats. Only the Norton and AVG Free software packages identified potential hazards. This may be due to the nature of the threats themselves, or the fact that much of the data came from previous years when these tools may not have existed, rendering them less likely to identify "old" threats. Figure 4.1 shows the types and frequencies of threats identified by Norton Antivirus for emails sent in the year 2007, and Figure 4.2 gives similar statistics for AVG free.

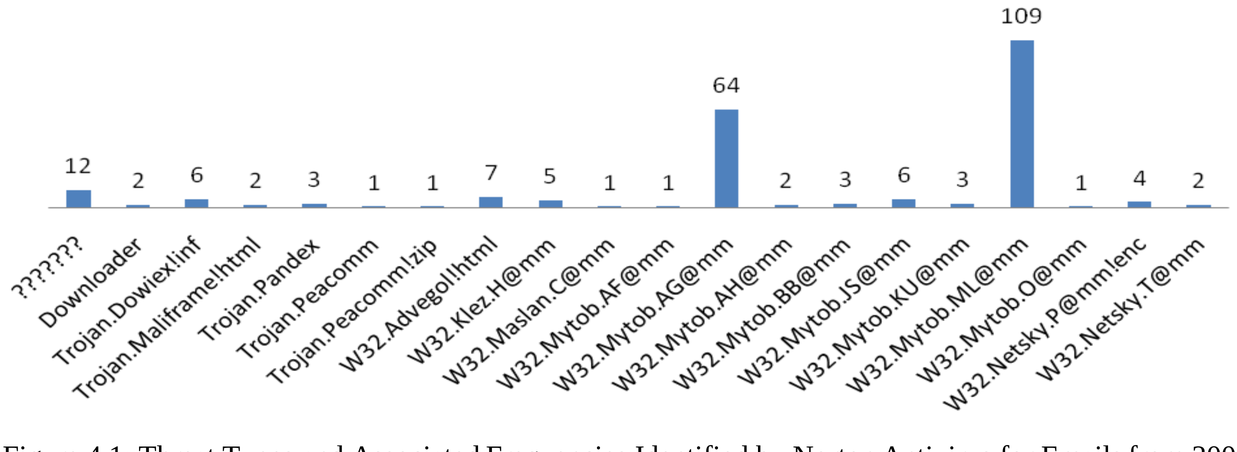

Figure 4.1: Threat Types and Associated Frequencies Identified by Norton Antivirus for Emails from 2007

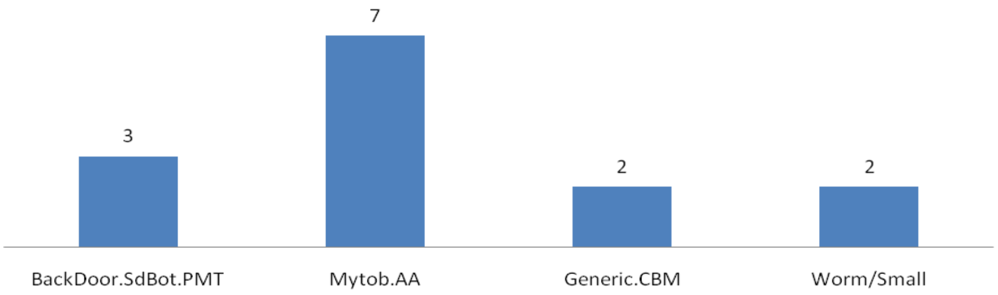

Figure 4.2: Threat Types and Associated Frequencies Identified by AVG Free for Emails from 2007

From these figures, we see that for the same data set (the year 2007), the two tools that detect threats have highly varying results. Norton Antivirus is a signature-based tool: it detects a virus if there is a defined bit sequence that identifies the virus (Symantec.com, 2009a). Note that product updates have recognized the limitations in this approach, and are now seeking to include behavioral-based recognition techniques (Symantec.com, 2009b). In contrast, AVG Free is an antivirus and antispyware tool, which does not incorporate download shielding or email threat detection measures; these are included in the commercial version (AVG.com, 2009). Since many of the threats found by Norton were present in attachments, this could be why AVG demonstrated a poorer performance at identifying them.

With respect to the other tools (Malwarebytes' Anti-malware, Super AntiSpyware, Spybot Search & Destroy, and Windows Bitdefender), either the age of the data or the fact that these tools are primarily concentrated around identifying spyware contributed to the lack of results. This could indicate that no threats of the sort these tools are set to detect were present, and not necessarily that the tools were faulty.

4.2 Threat Types and Distributions over Time

As detailed in Section 3.1, the email dataset obtained spans a period of 5 years (2004-2008). Figure 4.3 illustrates the trend over time in the number of viruses identified. A steady decrease is observed in the number of viruses received.

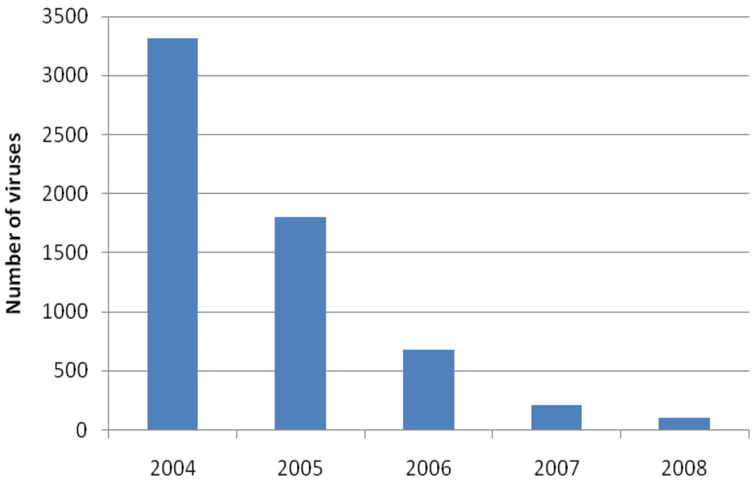

Figure 4.3: Number of Viruses Identified in Emails from the Dataset, over Time

A more detailed plot of the viruses identified by Norton over the 2004-2007 time periods is shown in Figure 4.4. Many of the identified viruses are variants of each other. We observe that the W32 prefix is present in many cases, which indicates that the virus is specifically targeting a Windows machine. It has been argued that Windows machines are more susceptible than Macintosh machines; however, since more individuals own Windows machines, it could simply be a matter of cyber adversaries choosing to tailor their threats to affect a larger population base (Schweitzer, 2009).

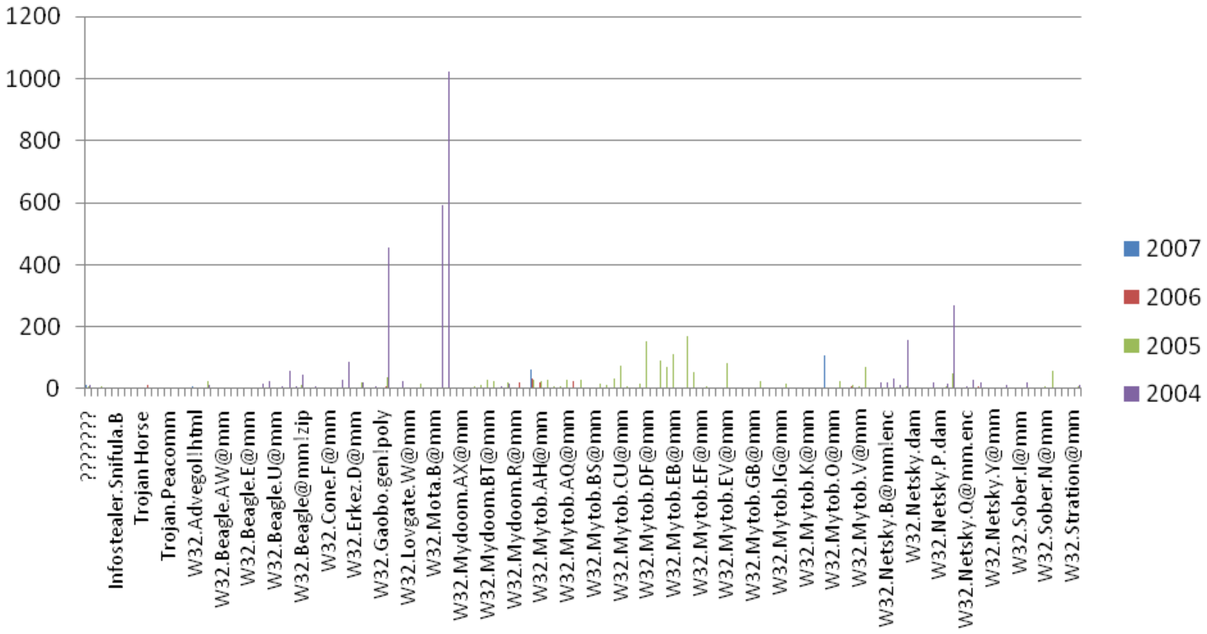

Figure 4.4: Number and Types of Viruses Detected by Norton Anti-Virus, by Year

If we adjust the data from Figure 4.3 for the number of emails received, we observe the trend shown in Figure 9. The results are relatively similar, but slightly more skewed to the extreme values.

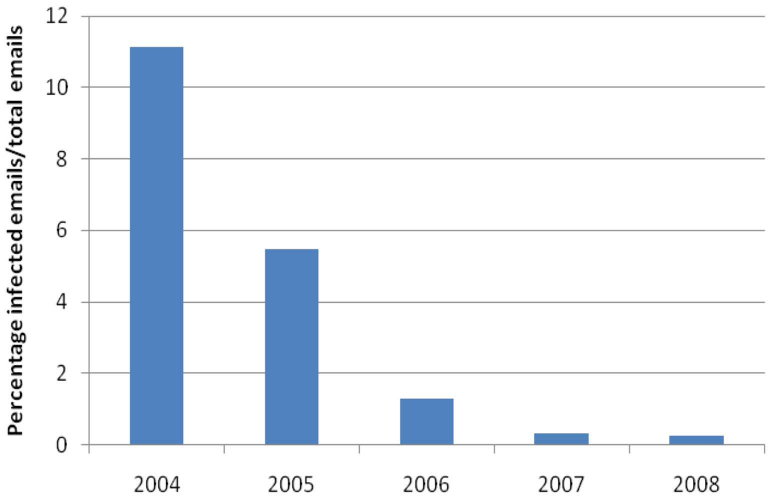

Figure 4.5: Ratio of the Number of Viruses Received to the Number of Emails Received, by Year

A direct comparison of the number of emails received and the number of infected emails is shown in Figure 4.6. Note that the number of infected emails represents a small fraction of the total; this could either be because most of the emails represent non-malicious spam, or because the cyber adversaries are using different attack vectors, such as malicious URLs.

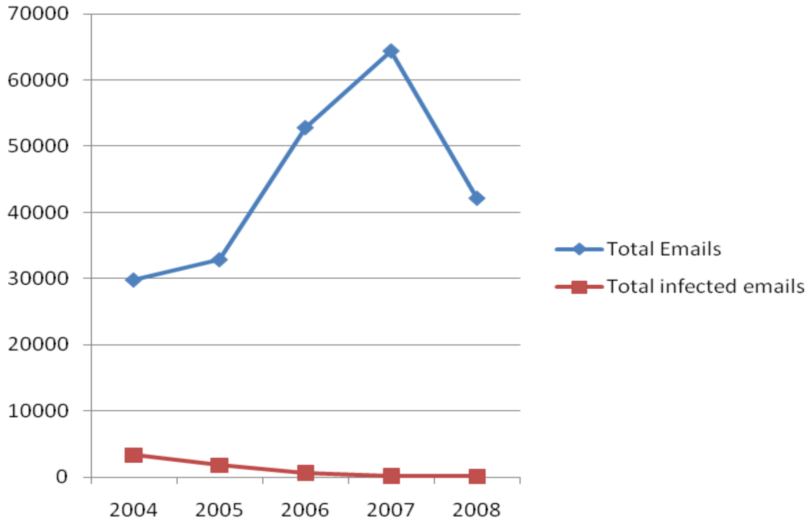

Figure 4.6: Comparison of the Number of Emails Received and the Number of Infected Emails, by Year

A finer-scaled version of this plot, using monthly increments, is shown in Figure 4.7, which can also be compared with the overall email traffic plot in Figure 3.1. In this plot, the downward trend in the number of viruses received in conjunction with the increase in the number of emails is even more visible.

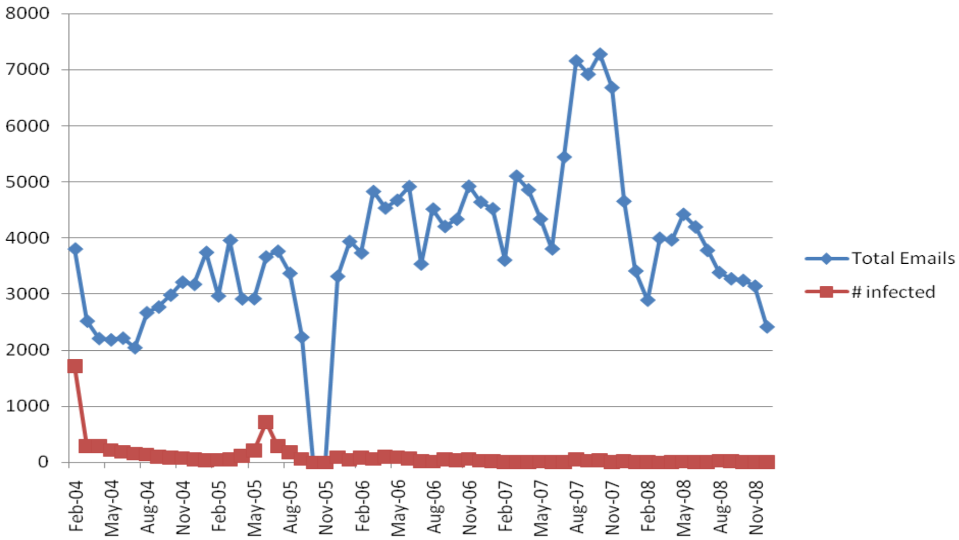

Figure 4.7: Comparison of the Number of Emails Received and the Number of Infected Emails, by Month

All of our analyses thus demonstrate a downward trend in the number of viruses over time. This is likely due to the fact sending viruses via email is a fairly unsophisticated attack method, and the overall sophistication level of attacks has greatly increased over time (Lipson, 2002). Moreover, the amount of user knowledge required to commit a sophisticated attack has decreased, which further accelerates the shift by adversaries to different methods of attack (Lipson, 2002). At the same time, email client capabilities at intercepting virus-laden attachments and emails have greatly improved, creating a further deterrent to this method of attack.

The viruses identified in our analysis were present in the emails in two different forms, either in the body of the email itself or in an attachment. Figure 4.8 shows the distribution of the viruses received in these two forms over time. We observe that as time advances, attachments continue to be the preferred method of attack. A finer version of this plot, using monthly increments, is shown in Figure 4.9.

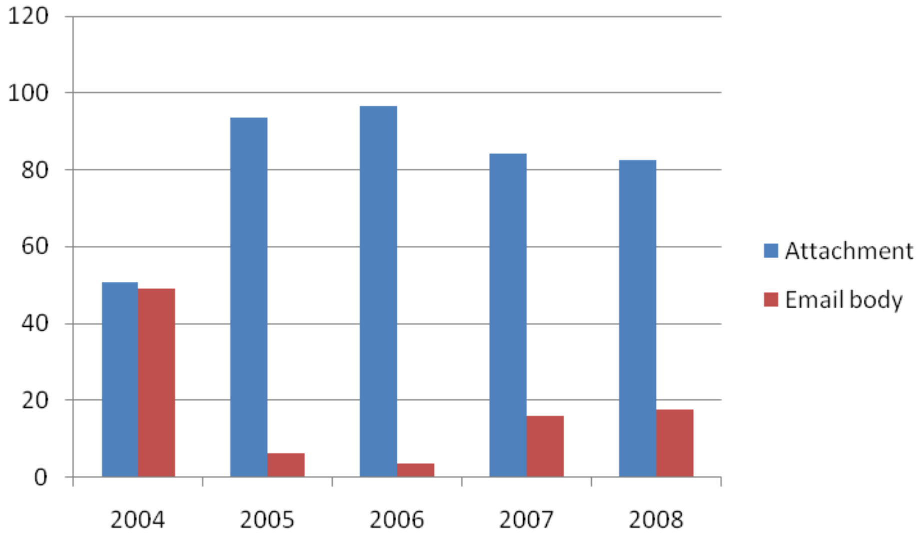

Figure 4.8: Comparison of the Number of Emails Sent via Attachments and in the Email Body, by Year

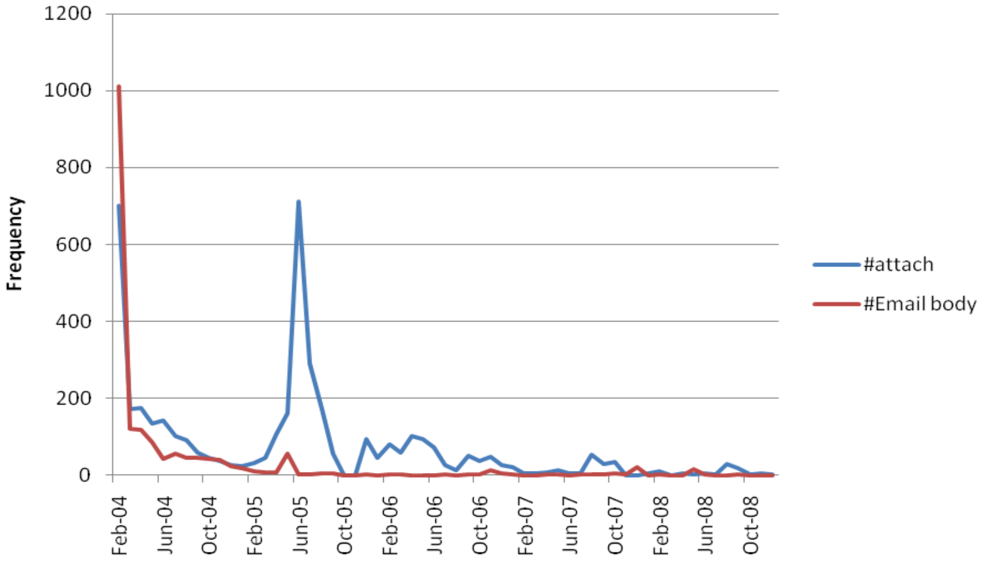

Figure 4.9: Comparison of the Number of Emails Sent via Attachments and in the Email Body, by Month

4.3 Implications on Adversary Behavior

We have seen in Sections 4.1 and 4.2 that the email data received by CIAC during the 2004-2009 time period contained many attached and embedded viruses. Over time, the raw quantity of viruses as well as the ratio of viruses to overall email traffic has decreased.

As described in the work of Lipson (2002) and others, we can infer from this portion of the analysis that while emails and their attachments must still be monitored with care, this is no longer the attack method of choice for many adversaries. In particular, given the financial reasons identified previously for hacking, phishing etc., this method of attack may no longer be seen as worthwhile, especially since many individuals and companies have a significant amount of protection in place with respect to email. For this reason, as will be discussed in the following section, it appears that attackers may have shifted their methodologies to more deceptive means: sending malicious URLs in the email body, for instance, which can pass undetected through email scanners.

5. Characterization of Malicious URL Content

This section addresses the maliciousness of URLs included in emails from the dataset. To assess the degree of danger of the URLs, four different online malicious website detection tools were employed. These online services each provided their own analysis on the URLs and indicated whether or not they were safe to visit. Since each of these services operates differently, we describe them each in turn below.

5.1 Tools Used

The four malicious website detection tools used are as follows:

- McAfee Site Advisor
- Norton SafeWeb
- Web of Trust
- Google SafeBrowse

Each of these services offers a web browser add-on tool. As a user attempts to visit a URL deemed dangerous, the browser will either block access or indicate that the website is dangerous, requiring an extra click to accept the risk before the page is loaded. All of the services also provide a website that gives richer details on the nature of the associated threat.

McAfee SiteAdvisor uses test computers to test the submitted URL. Feedback to the user indicates either that they have determined the site is safe; that they have not found any significant problems with the site; that they have not tested the site; or that they have found a specific danger associated with the site. McAfee registers an email address with the site if possible and captures the amount of email received. They also perform download tests and rate each one. Finally, McAfee allows users to post ratings and comments on any of the tested URLs.

Norton's SafeWeb closely resembles McAfee SiteAdvisor in its operation. However, they also indicate to the user the *number* of computer threats, identity threats, e-commerce safety threats, and annoyance factors associated with the website. Based on those results, Norton either issues a warning or allows the user access to the site. Norton also lists the number of each specific type of threat based on a classification among 18 different categories.

Web of Trust (WOT) is a Finnish service, which differs from McAfee and Norton in that test machines are not used. Instead, the service relies entirely on partners, blacklist sites, whitelist sites, open web information, and a user-contributed information base. A proprietary algorithm produces ratings for each site on trustworthiness (0-100) and confidence (1-5); vendor reliability (0-100) and confidence (1-5); privacy (0-100) and confidence (1-5); and child safety (0-100) and confidence (1-5). For example, a URL that appears on several malware blacklists, is rated poorly by the user community, and has no positive ratings will score low on trustworthiness with a medium to high confidence. A URL that is bookmarked on digg.com, appears in Wikipedia in several languages, is included in the Open Project Directory, and receives good user reviews will likely score high with a high confidence. URLs with very little source information have low confidence scores. The algorithm used by Web of Trust is not provided on their website, nor are definitions for the factors of trustworthiness, vendor reliability, privacy, and child safety. Altogether, WOT has rated over 24 million sites, 5.6% of which are dangerous based on “trustworthiness,” 5.9% dangerous based on “vendor reliability,” and 7.4% dangerous based on “privacy.”

Google’s SafeBrowse operates fairly differently from the other three services. Google does use test machines to visit the site; however, in addition it also reports 90-day *historical information*. The output includes the following components: whether a site is currently suspicious (and how often it was suspicious in the past); what happened when Google visited the site (when it was last scanned, up to three kinds of attacks found, up to three domains hosting malware, and up to three intermediary domains for distributing malware); whether the site has acted as an intermediary resulting in the further distribution of malware (possibly inadvertently, based on past scanning); and whether the site has hosted malware (how often it has done so, and up to three victim sites that initiated the distribution of malware).

Table 5.1 compares the features and operations of these four malicious website detection tools.

	McAfee	**Norton**	**Web of Trust**	**Google**
uses test machines to visit URLs	yes	yes	no	yes
provides user comments and/or ratings	yes	yes	yes	no
indicates the type of attack associated with site	yes	yes	no	yes
lists the specific name and url of attack	yes	yes	no	no
leverages other services, blacklists, or whitelists	no	no	yes	no
provides temporal history information	no	no	no	yes
provides country information	yes	yes	yes	no

Table 5.1: Comparison of the Four Malicious Website Detection Tools

Because each service operates differently and provides different information to the user, it is challenging to combine the information from the four sources to obtain a single overall ‘badness’ score. On the other hand, due to these differences, each service allows the problem to be analyzed from a different angle. For example, while Google has the least coverage among the URLs we tested, their service indicates how many times the URL was suspicious in the last 90 days and when it was last scanned. Web of Trust performs no analysis on the specific attack, but it has the most coverage by far of the sites we tested, due to the leveraging of other sources. McAfee has the richest information based on user reviews, while Norton provides excellent information on the specific attack type.

5.2 Querying Methodology

A database was formed containing all of the URLs received in the email dataset. Many emails had multiple URLs, and many URLs were present in several emails, producing a many-to-many relationship. The time and date of each email was also captured for each associated URL.

A Java program was then written to submit every unique URL to each of the four services (McAfee, Norton, Web of Trust, and Google) and process the natural language results into a tabular format. A delay was implemented between queries to avoid overloading networks on either side of the transaction. All of the URLs were sent to the McAfee, Norton and Google tools, and the corresponding results were processed. Due to querying limitations enforced by Web of Trust, only unique domains from emails sent in June and July of 2004-2005 and 2007-2009 were scanned.[1]

Due to the deceptive and adversarial nature of the problem we are studying, we do not have access to data representing the ground truth. If we had knowledge of the false positive rates and/or false negative rates associated with the malicious website detection tools, we would be able to perform a much richer analysis. The analysis contained herein is limited by the fact that each of the four services have tested only a fraction of the domains received, and only a fraction of those are determined to be malicious. We thus restrict ourselves to analyses where sufficient data is present for the results to be significant.

5.3 Malicious URL Summary Statistics

A total of 46,150 unique domains were contained in the email dataset. Of these, 29,867 domains were present in only one email. The mean number of emails a link was found in was 5.2, and the median was 1. The most common domains received are listed in Table 5.2, along with the number of associated emails.

Domain	Number of Emails Containing Link
www.xnabalada.com	14,233
www.w3.org	5,686
pics.ebaystatic.com	1,615
pages.ebay.com	1,581
cnlinfo.net	1,526
us.rd.yahoo.com	1,068
sites.google.com	1,002
cgi4.ebay.com	882
www.mailx.cf.st	854
cafe.daum.net	819
bbs.cnlinfo.net	804
awxyz.com:112	801

Table 5.2: Most Common Domains Received Across All Emails

[1] We use the term "domain" in our analysis to refer to the full root of a web address, disregarding subsidiary pages: for instance, www.us.rd.yahoo is considered a different domain from www.cn.rd.yahoo.com, but the same domain as www.us.rd.yahoo.com/pages. We do this because each unique root is often identified separately by the four search tools, while subsidiary pages inherit ratings from the associated parent page.)

Due to the limitation in querying the Web of Trust tool, much of our analysis (including all of the statistics in Section 5.3) is based on emails received in the months of June and July. Fortunately, these months are roughly representative of the whole year. Figure 5.1 shows the results of an odds ratio test, comparing the likelihood of a URL receiving a warning (by at least one of the tools) across all of the months, using June and July as a base. There are roughly the same number of months on either side of the midpoint (ratio = 1), implying that June and July are representative of the sample as a whole.

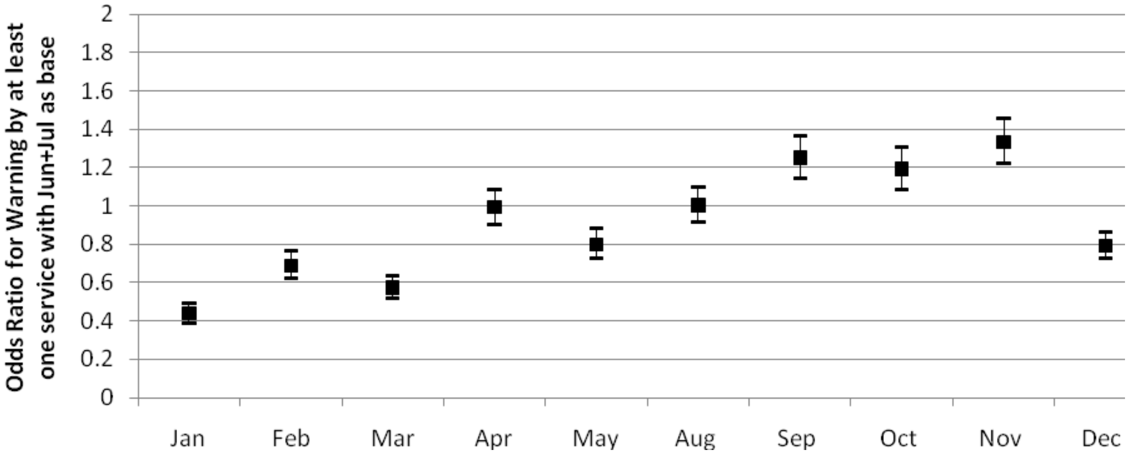

Figure 5.1: Odds Ratio and 95% Confidence Bars Comparing Maliciousness of June-July to Other Months

Statistics on the number of domains per email that produce a warning by at least one of the four sites can be found in the Appendix, along with an in-depth analysis of the top five domains encountered.

One way of comparing the maliciousness of different domains is to compute an aggregate score based on all four services. Since not all four services tested each domain, we associate the score with a confidence level. For example, if all four services tested the domain to be suspicious or malicious, we give a score of 1 with a confidence of 1; here, each service has a weight of ¼ to both the score and the confidence. If two services tested the domain and only one found it to be suspicious, then we assign a score of ½ with a confidence of ½. More sophisticated classification methods could further refine these weights, but in the absence of the “ground truth” this analysis was considered to be the most reasonable.

Figure 5.2 below displays a histogram of the number of domains tested by each of the tools.

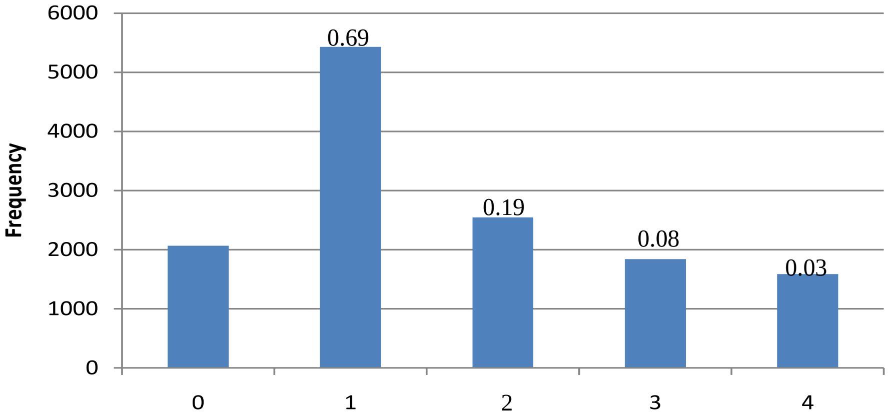

Figure 5.2: Histogram of the Number of Services that Tested Each of the Domains in the June-July Dataset

For example, 5,428 domains were tested by only one service, while 2,547 domains were tested by 2 services. The value above indicates the average score in that category; e.g., the average score for domains tested by only one service is 0.69. Note that domains with a greater web presence, which presumably are less likely to be malicious, are more likely to have been tested by all or most services.

With regard to countries of origin, Figure 5.3 displays a histogram of the countries of the domains in the emails. The country of origin was reported by the Web of Trust, McAfee and Norton tools. When these datasets report conflicting country information, we indicate that as "conflicting"; when none of the sites report a country of origin, we indicate that as "unknown."

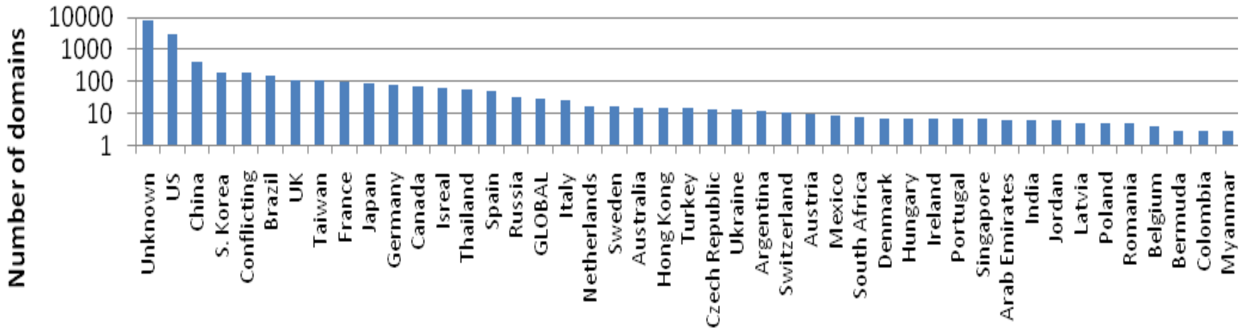

Figure 5.3: Frequency of Countries Associated with Domains in the June-July Dataset

Summary statistics on the percentage of domains in which either a warning or "safe to proceed" is issued by at least one site are listed in Table 5.3. We also list the percentage of domains untested by any service, and where the services disagree with each other (i.e. Norton declares a domain is safe and McAfee issues a warning for the same domain). In this table, "False" means that the described event did not occur, and "True" means that it did.

	False	**True**	**%**
warning by at least one service	8,155	5,316	39%
not tested by any service	11,405	2,066	15%
at least one service tested safe	5,901	7,570	56%
services disagree on warning	11,990	1,481	11%

Table 5.3. Summary Statistics on Domain Warnings Issued by the Four Services

We observe that at least two of the services disagree on giving a warning for 11% of the domains. That is, one service declares that a domain is safe, while another declares that it is dangerous. In future work, it would be interesting to utilize this subset of data to better understand the false positive and negative rates of the tools. While the true false positive and false negative rates will not be known, this information could be used to give a more accurate weighting for the aggregated score described previously. Moreover, the cases in which the services disagree could potentially represent attacks that are more sophisticated.

Figure 5.4 illustrates the change over the years in the total and percentage of domains with a warning by at least one service. We observe that more recent domains are more likely to be determined malicious, which is reasonable since all domains were tested in 2009.

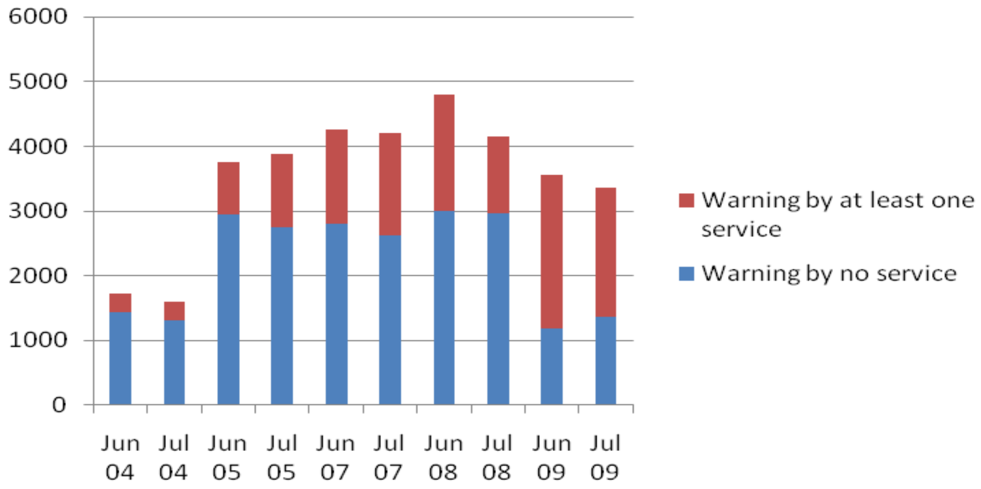

Figure 5.4: Warnings by at Least One Service, Over Multiple Years

5.4 Comparison of Malicious URL Detection Tools

Web of Trust has tested more sites than the other three, as shown in Figure 5.5. Of the tested domains, Web of Trust also issues warnings for a larger percentage of sites, followed by McAfee, Norton, and Google.

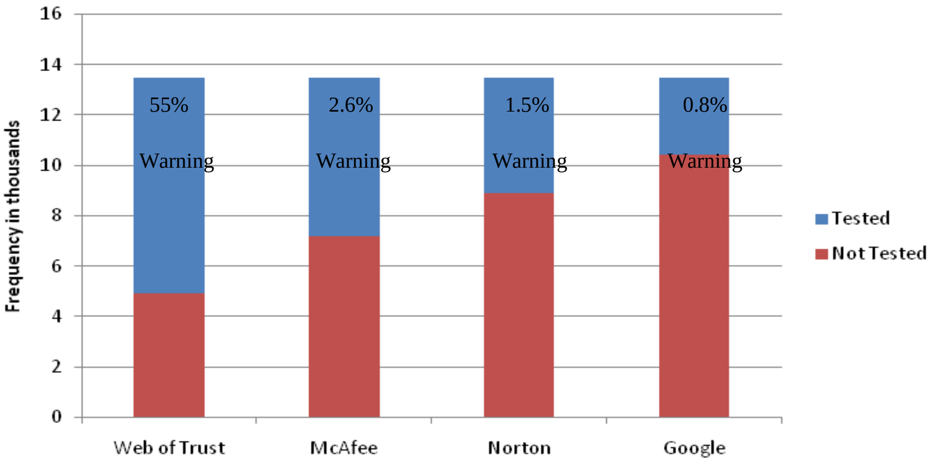

Figure 5.5: Histogram of Number of Domains Tested by Each of the Services in the June-July Dataset

5.4.1 Web of Trust

As mentioned in Section 5.1, Web of Trust rates sites based on trustworthiness, vendor reliability, privacy, and child safety; each of these ratings is associated with a confidence factor, from 1 (low confidence) to 5 (high confidence), indicating the strength of the result. Figure 5.6 shows a histogram of the “trustworthiness” ratings when the confidence level is arbitrary, versus when the confidence level is relatively high (>3). This is based on the June-July dataset.

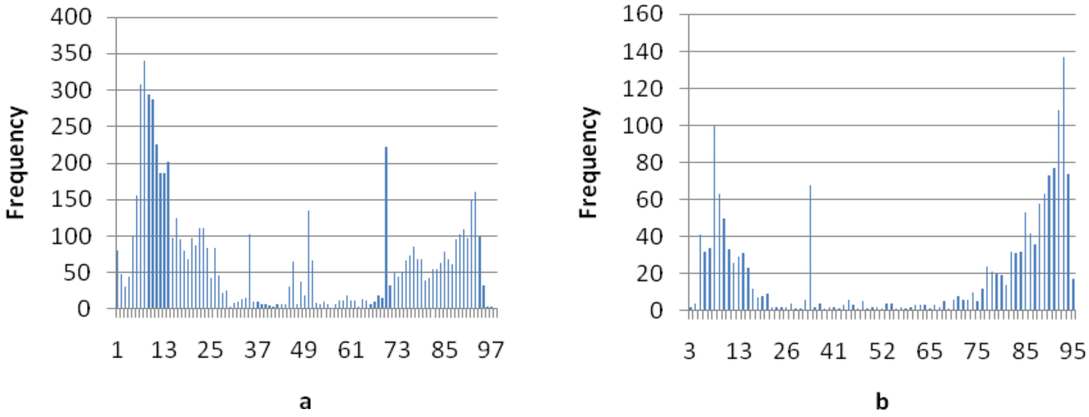

Figure 5.6: Histogram of Trustworthiness Scores for Arbitrary Confidence (a) and High Confidence (b)

Histograms of the Web of Trust ratings for vendor reliability and privacy by confidence level can be found in the appendix.

In addition to rating the websites, Web of Trust also allows users to provide ratings and comments. Figure 5.7 is a histogram of the categories of the user comments.

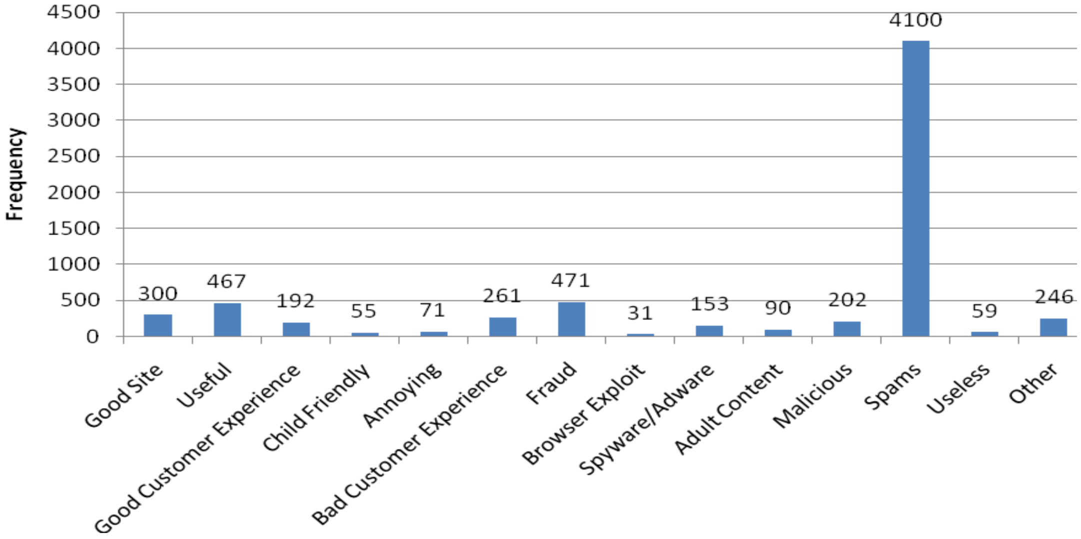

Figure 5.7: Categories of User Comments in Web of Trust

It is worth noting that we cannot currently examine attack types over time, because the domains were all tested during the same relatively short time period. We do not know what attack type was present when the domain was sent, only what attack type is present today.

5.4.2 McAfee Site Advisor

McAfee's user base provides an excellent source of data. Figure 5.8 shows the proportion of sites with at least one user report of each attack type, across all months. Spam is the most commonly cited method, followed by phishing and other scams. A full description of the McAfee results, along with the proportions of such threats among tested domains, can be found in the appendix.

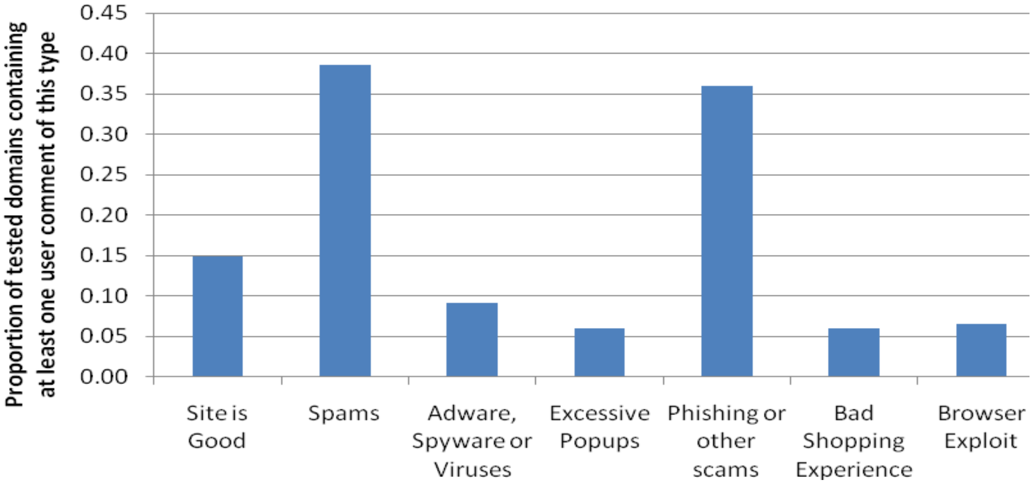

Figure 5.8: Proportion of Tested Domains Containing Threats, as Cited in User Comments

Each of these categories can be further decomposed by country of origin, leading to a rich set of data. A histogram for the "adware, spyware, or viruses" country is shown in Figure 5.9; from it we see that the United States and China are the largest sources of such malware. Similar graphs for each of the other categories can be found in the appendix.

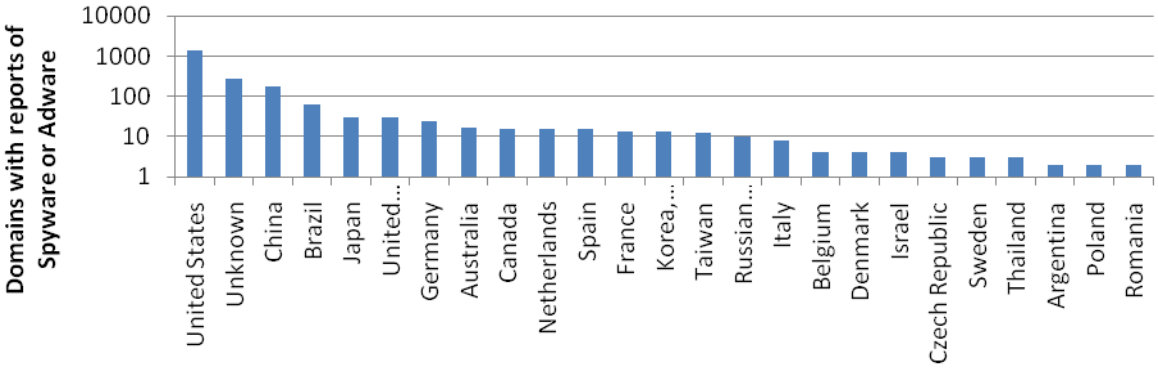

Figure 5.9: Countries of Origin for Domains with McAfee Site Advisor Reports of Adware or Spyware

5.4.3 Norton SafeWeb

Norton SafeWeb is similar in operation to McAfee Site Advisor, with an additional feature that calculates the number of different threats associated with a particular site. Figure 5.10 shows a histogram of the number of different computer threats present in the tested domains, across all months. A histogram of the number of different identity threats found in the domains is located in the appendix.

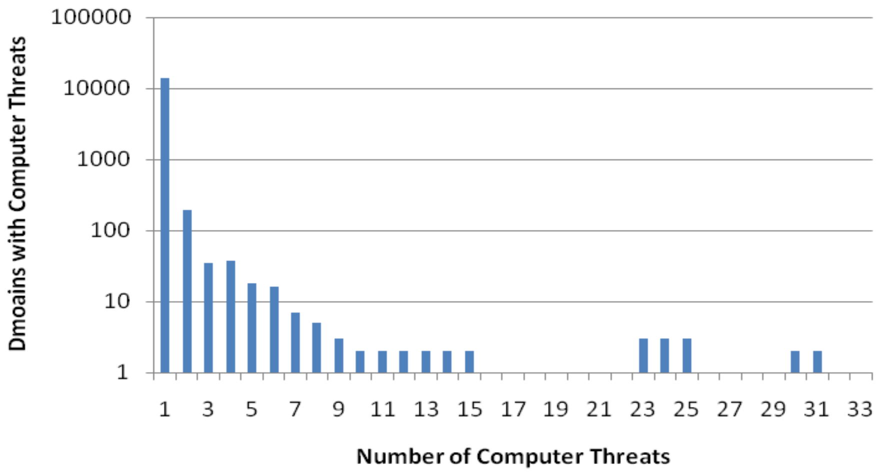

Figure 5.10: Number of Different Computer Threats Presented in Domains Tested by Norton SafeWeb

Norton's test machines also capture detailed information on the type of associated attack. Table 5.4 displays the number of domains with at least one incident of each attack type. Viruses are the most common method, followed by drive-by-downloads and phishing attacks.

Attack Type	Number of domains with at least one incident
Viruses	248
Drive By Downloads	95
Phishing Attacks	31
Adware	23
Information stealers	14
Security Risks	12
Downloaders	10
Heuristic Viruses	9
Trojans	9
Suspicious Applications	9
Worms	7
Dialers	3
Backdoors	1
Spyware	1
Malicious Downloads	1

Table 5.4: Types of Incidents Detected by Norton SafeWeb, with Frequencies

5.4.4 Google SafeBrowse

Although Google tested the fewest domains among the four services, it provides information along the most interesting dimensions, including a temporal dimension. For example, only 18 of the 72 (25%) currently suspicious domains hosted the malware themselves. In addition, a total of 132 of the 7,689 tested domains were found to have hosted some malware within the past 90 days.

Figures 5.11-5.13 display histograms of the kinds of information captured by Google, across all months. These results further illustrate the degree of complication in understanding the malware network. For example, Figures 5.11 and 5.12 below demonstrate that a particular domain can switch between being suspicious and not suspicious many times within a 90 day period and can do so with many days in between. Figure 5.13 shows that a small number of infected domains can be used to infect a large number of others.

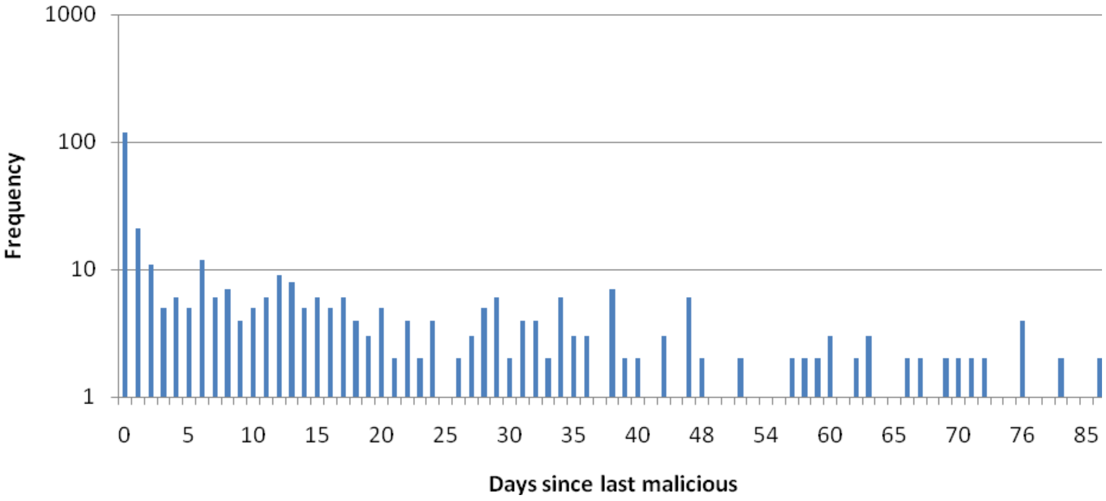

Figure 5.11: Days since Google SafeBrowse Determined a Domain to be Malicious

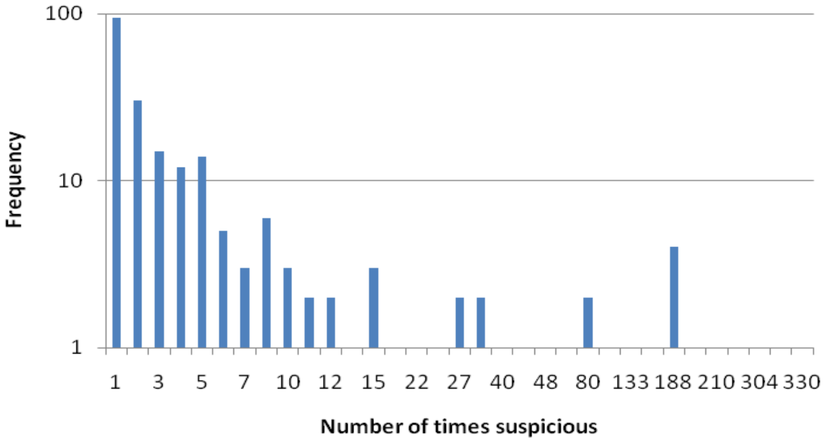

Figure 5.12: Number of Times Google SafeBrowse Found a Domain to be Suspicious

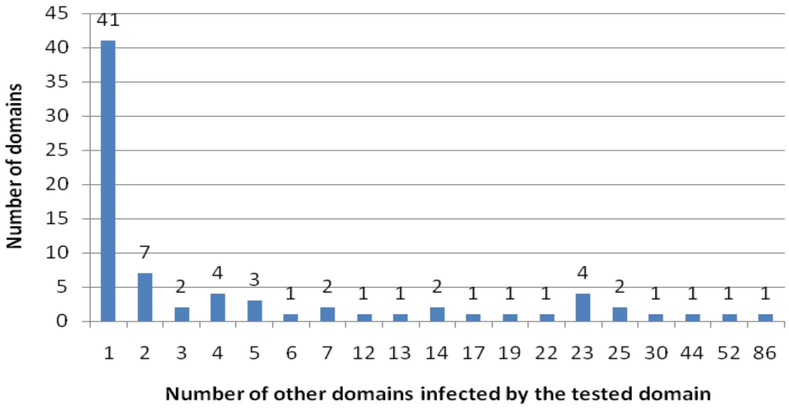

Figure 5.13: Number of Other Domains Determined to be Infected by the Tested Domain

In addition, other analyses from Google SafeBrowse show that malware is often not hosted on the domain's network host, or it may be hosted on many different networks. The domain may also be acting as an intermediary and knowingly serving malicious content to other domains, or the domain may unknowingly be serving malicious content through one of its intermediary sites. Figures demonstrating each of these relationships are included in the appendix.

5.5 Data Trends and Implications on Adversary Behavior

In this section, we analyze various trends present in the data, both in terms of the different kinds of adversaries and their behavior. Where possible, we try to draw conclusions about adversarial origins and motivations. In what follows, we consider only the domains that were tested by at least one of the tools.

We use two types of statistical tests to assess trends in the June-July dataset. For binary data (such as warning versus no warning), we use an odds ratio test computed by median-unbiased estimation; the associated confidence interval is computed using exact methods. For continuous data (such as Web of Trust scores) we use box plots to display different categories side by side with the bootstrapped means of each category. The horizontal line inside the box indicates the median of the bootstrapped data. The top and bottom of the boxes indicate the 25th and 75th percentiles, and the dashed lines indicate a 95% confidence interval of the data. Both methods were implemented in the statistical package R.

The first analysis concerns the trustworthiness of the emails received, by country. The data include the Web of Trust trustworthiness scores and the aggregated country data presented in Figure 5.3. Only scores with a confidence level of greater than 1 out of 5 are used. Figure 5.14 is a box plot of the bootstrapped mean of the trustworthiness scores by country. We constructed the box plot using the 20th percentile instead of the mean for bootstrapping, and found that we get a similar result (the same countries score low or high respectively) as a slightly different ranking of the countries by median. Considering the 20th percentile focuses the analysis on the lowest scores. Here, outliers are represented by small circles.

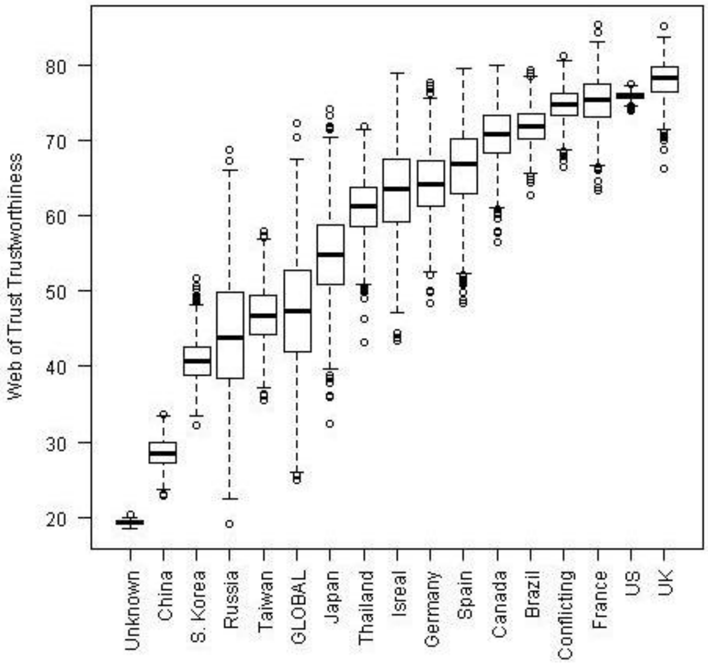

Figure 5.14: Bootstrapped Mean Trustworthiness Scores with a 95% Confidence Level, by Country

From this analysis we observe that the lowest trustworthiness scores come from domains where the country of origin could not be ascertained, followed by China, and South Korea. In general, the lowest trustworthiness scores are associated with Southeast Asian countries, and the highest scores are associated with the United States and Western Europe. A similar analysis using the Web of Trust vendor reliability scores instead of trustworthiness scores produces comparable results, and is included in the appendix.

The next analysis addresses the relative likelihood that a site will generate a warning by at least one of the four tools. The data in this case is the same as that used to generate Figures 5.2-5.5. An odds ratio test is performed to compare the likelihood of domains from different countries being flagged with the likelihood associated with the US. Figure 5.15 shows the results; in this case the median ratio is denoted with a box, and the whiskers on either side represent a 95% confidence interval.

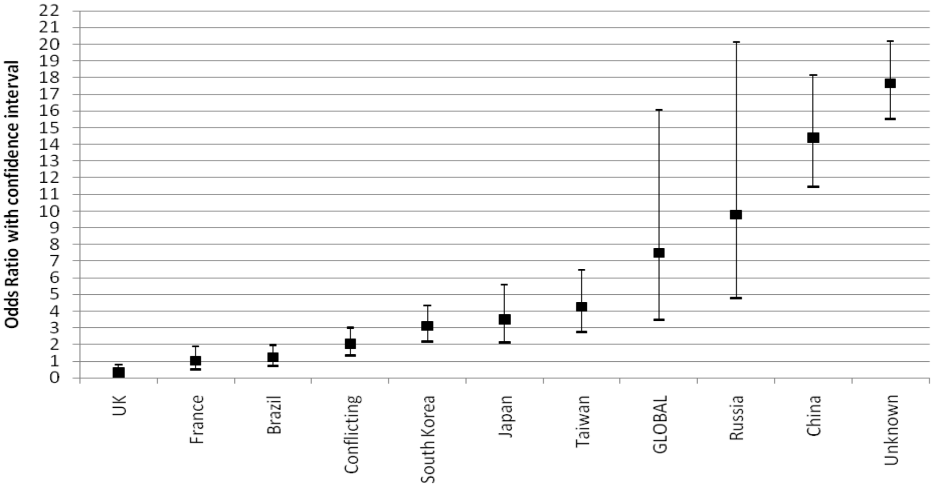

Figure 5.15: Odds Ratio and 95% Confidence Intervals for Odds of Generating a Warning Relative to the US

In this graph, an odds ratio value below 1 suggests that a domain from that country is less likely to generate a warning than one from the US. There is only one such country in this analysis, which is the United Kingdom. An odds ratio above 1 suggests that a domain from that country is more likely to generate a warning; the higher the ratio, the more likely the difference. We see from this analysis that domains from China are much more likely to generate a warning than any other country; it can be inferred that spammers from China are more likely to have a malicious intent than those in other places.

As a companion to this analysis, Figure 5.16 compares the likelihood of generating a warning in the year 2004 versus the other years. As might be expected, domains from 2009 have the most warnings, possibly both because malicious domains tend to be short-lived and because our analysis was done in 2009.

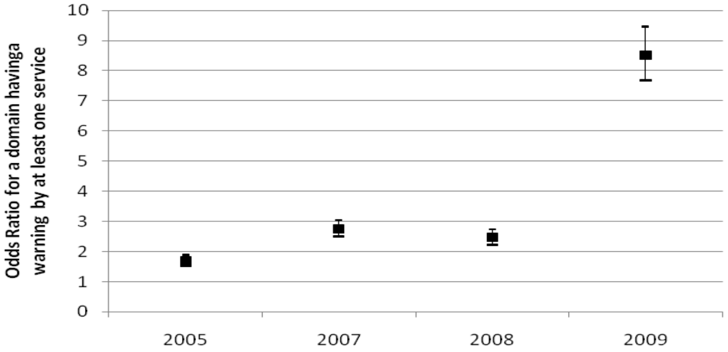

Figure 5.16: Odds Ratio and 95% Confidence Intervals for Odds of Generating a Warning Relative to 2004

Two different analyses considered the effect of time on when a threat was received. Figure 5.17 gives a box plot of the weighted maliciousness score (as described in Section 5.3) versus the day of the week that an email was sent. As in Figure 5.14, the dashed lines indicate the 95% confidence interval and the dots represent data outliers.

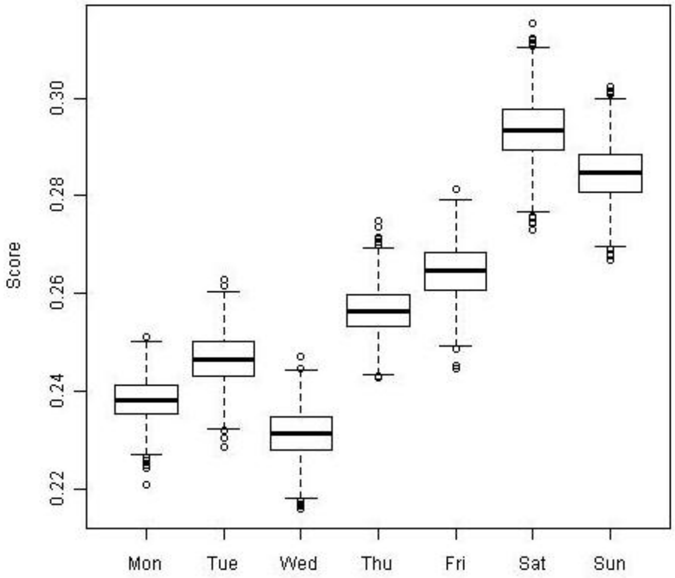

Figure 5.17: Bootstrapped Mean Maliciousness Scores with a 95% Confidence Level, by Day of Week

From these results we see that emails are more likely to be malicious if they are received on the weekend. This is somewhat plausible, as a higher percentage of legitimate work-related emails are likely to be received during the week, in particular during traditional business hours.

An identical calculation was performed to compare the level of maliciousness with the time of day that an email was sent. Figure 5.18 shows the results.

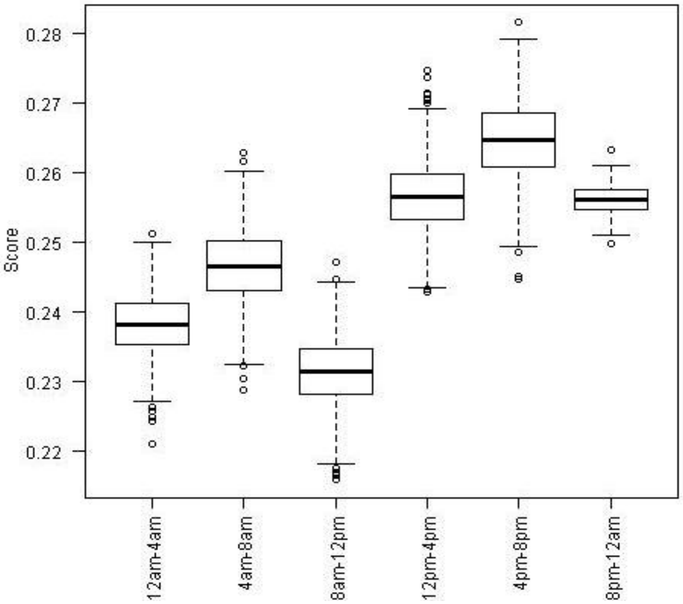

Figure 5.18: Bootstrapped Mean Maliciousness Scores with a 95% Confidence Level, by Time of Day

These results suggest that a smaller fraction of malicious emails are received in the morning, particularly during working hours. There is a greater likelihood of maliciousness associated with emails sent later in the afternoon and at night. Altogether, this suggests that adversaries are probably more active during the nighttime hours, either because they work more at night or because such hours correspond to daylight hours in some of the countries that represent the greatest threats, such as China.

The final analysis in this section compares the median maliciousness scores associated with different countries. This is similar to the analysis in Figure 5.14, using the same underlying data as Figures 5.17 and 5.18.

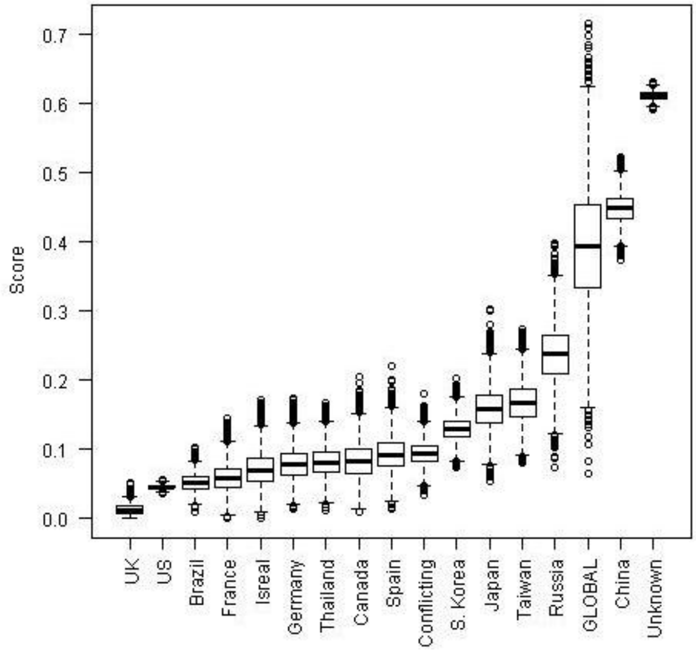

Figure 5.19: Bootstrapped Mean Maliciousness Scores with a 95% Confidence Level, by Country

As before, we observe that China and other far Eastern countries are associated with the highest levels of maliciousness; western Europe and the US are associated with the lowest levels. The subset of domains for which the country is unknown (either through obfuscation or other measures) is the most malicious.

6. Conclusions

6.1 Overall Implications on Adversary Behavior

The data for this study allows us to provide a characterization of the kinds of adversaries who use email as an attack vector. Because email is a generic platform from which many different kinds of attacks can be launched, we can reasonably assume that the adversaries represented in this sample span a range of different classes and skill levels represented in Table 2.1. In particular, those adversaries choosing to send their threats via attachments probably represent some of the less sophisticated groups (such as script kiddies), while those employing malicious URLs are likely to be more sophisticated (such as cyber punks or coders). It is also possible that some of the most sophisticated groups (such as black hat hackers) use such a platform as a stepping-stone to gain entry to systems from which to launch far more complicated and damaging attacks, particularly in conjunction with techniques such as social engineering. Moreover, since the traffic is associated specifically with a site sponsored by the Department of Energy, it is highly probable that the hacktivist and even possibly cyberterrorist adversary classes are represented in this sample as well. In the absence of "ground truth" data, we cannot claim any of these assertions with certainty, but these conclusions seem the most plausible given the data that we have observed.

In terms of the attack methods chosen, we observe that the use of email attachments as an attack vector has decreased sharply over time (Figures 4.3 and 4.5). This is likely due both to the fact that email servers have implemented stronger screening procedures to guard against such attacks, and also because the threat space itself has shifted, with less sophistication required to launch an attack using malicious URLs (Lipson, 2002). The number of attacks detected via malicious URLs actually increased during the years in the sample (see Figures 5.4 and 5.16), which supports this hypothesis. With regard to the attacks themselves, the majority of email attachments contained Windows viruses (Figure 4.4), while the largest portion of the malicious URLs were associated with viruses and drive-by downloads (Table 5.4).

Our analyses identify several traits about the adversaries and trends in their behavior. In Section 3, we observed that the top four countries of origin of emails in the sample are China, the United States, South Korea, and Brazil (Figures 3.3 and 3.5). These are identical to the top countries associated with domains embedded in the emails (Figure 5.3). Two of these countries (China and South Korea) also score very high on the maliciousness (Figure 5.19) and low on the trustworthiness (Figure 5.14) of associated emails, while the other two (the United States and Brazil) do not. We can therefore conclude that the largest number of malicious emails in the sample is connected with adversaries in southeast Asia. (Taiwan and Japan have results that are similar to China and South Korea (though less extreme), which bolsters this finding.)

We also observe that emails sent on weekends (Figure 5.17) are more likely to be malicious than emails sent on weekdays, and the time of day with the highest percentage of malicious activity is late afternoons and evenings (Figure 5.18). The first result could be due to the fact that fewer legitimate queries are sent on weekends, and the second result could be because late afternoons in California correspond to mornings in Asia.

The persistence of different names and domains within the dataset is addressed in Appendix A, along with measures of IP and "Send" deception used by adversaries. We see that the highest number of emails is received from senders with free e-mail accounts (particularly yahoo and hotmail), and the overall deceptiveness of sender information is highest in the year 2007. This analysis is still preliminary, but it does illuminate a different and interesting aspect of adversarial behavior.

Finally, we note that the different tools that we used to study cyber adversaries produced dramatically different results. In the case of email attachments, only two of the six tested tools (Norton Antivirus and AVG free) found any threats at all (Figures 4.1 and 4.2). With respect to malicious URL detection tools, Web of Trust tested many more of the domains than the other three, as well as detecting threats in a significantly higher percentage of websites (Figure 5.5). While the other tools contain some more interesting outputs (in particular, Norton SafeWeb's classification of attack types, Google SafeBrowse's temporal analysis, and McAfee SiteAdvisor's information on countries of origin), such outputs are only useful if the detection rate of the tool itself is high enough to draw statistical conclusions about the sample as a whole.

6.2 Areas for Future Study

The analysis capabilities that we have demonstrated lend themselves easily to future studies. The software that we used is entirely free, with the exception of Norton AntiVirus. Customized scripts for querying the four malicious URL detection tools and processing the results are available as one of the deliverables of this project.

In terms of areas where our work might be expanded, the most obvious extensions are associated with data in which the "ground truth" is known. Given such data, we could further grade the accuracy of all of the tools considered; for instance, we could estimate the false alarm and detection rates for each of the malicious URL detection tools and construct the corresponding ROC curves. Such findings could then be used to create a classification scheme utilizing the weighted results of all four of the tools. We could also study the response rates of different antivirus programs, and determine which services are best at identifying email threats.

The dataset we had was quite good, in that it represented truly unfiltered traffic; however, it also contained some limitations, most notably relating to the time periods of the data collected. As previously discussed (Figure 3.1), the email traffic for the CIAC addresses tapered off over time, corresponding with the shutdown of the associated website. Given a more current stream of emails, some of the tools might have performed better and a larger percentage of malicious emails might have been tagged.

Other areas of future study might address the deceptiveness of the emails themselves, as was started in the analysis in Appendix A. Certain email topics might be associated with higher levels of maliciousness, which could be used in conjunction with known header spoofing techniques to generate a classification scheme based on the deceptive characteristics of the email textual and header information. This scheme could be used in conjunction with a malicious URL classification scheme to give a much deeper view of the threat space.

Finally, this work could be expanded by additional data mining of the collected antivirus and malicious URL detection results. We have performed analyses and statistical tests on all topics that we thought were likely to yield interesting results, but the sheer volume of the data itself (hundreds of thousands of emails) suggests that there could be other trends that we might have overlooked in this first pass. We are happy to provide our raw data to anyone who is interested in continuing such analyses.

Acknowledgements

We would like to thank Bill Orvis first and foremost for giving us access to the CIAC data, for burning us CDs of all the newest emails as they arrived, and for collecting the data in the first place. This work would quite literally not have taken place without his helpfulness.

We would also like to acknowledge Tony Bartoletti and Bill Hanley for serving as mentors to this project from the beginning. Tony's enthusiasm for sharing his ideas and expertise in cyber security was extremely helpful, and our analyses benefitted greatly from his discussions. Bill served in both technical and managerial roles, suggesting interesting areas of study and making sure that we had the resources to complete the work. In addition, we thank Jason Lenderman for his Java classes that made malicious URL data collection easier; both he and Ryan Prenger also offered valuable statistical advice on work in Section 5.

Finally, we would like to thank everyone who talked with us about aspects of cyber security at LLNL and elsewhere: Chris Roblee, Celeste Matarazzo, Matt Myrick, Michael O'Hara, Brenda Ng, Bob Deri, John Nitao, John Dias, and David Burnett. Your input has been very valuable in shaping our work.

Appendix A. Supplemental Material on Email Persistence and Deceptiveness

A.1 Measures of Email Persistence

The data considered in this study is email data, which does not provide a means for measuring the sender's motivation. Additionally, it would be difficult to measure a sender's maliciousness. As an alternative measure, we consider the *persistence* of a sender. This can be measured through the name and/or domain from which the email originated. Although these fields are spoofable, this provides an initial measure. It can also indicate the "creativity" or "level of sophistication" of an adversary depending on the frequency of the same name or domain.

Name persistence (Name@something.com):

Figure A.1 shows the different "Names" and their distribution in the emails received. The most frequent cases are 'info', 'CDGH,' and 'root.' Blank cases have been excluded, but they receive the most counts.

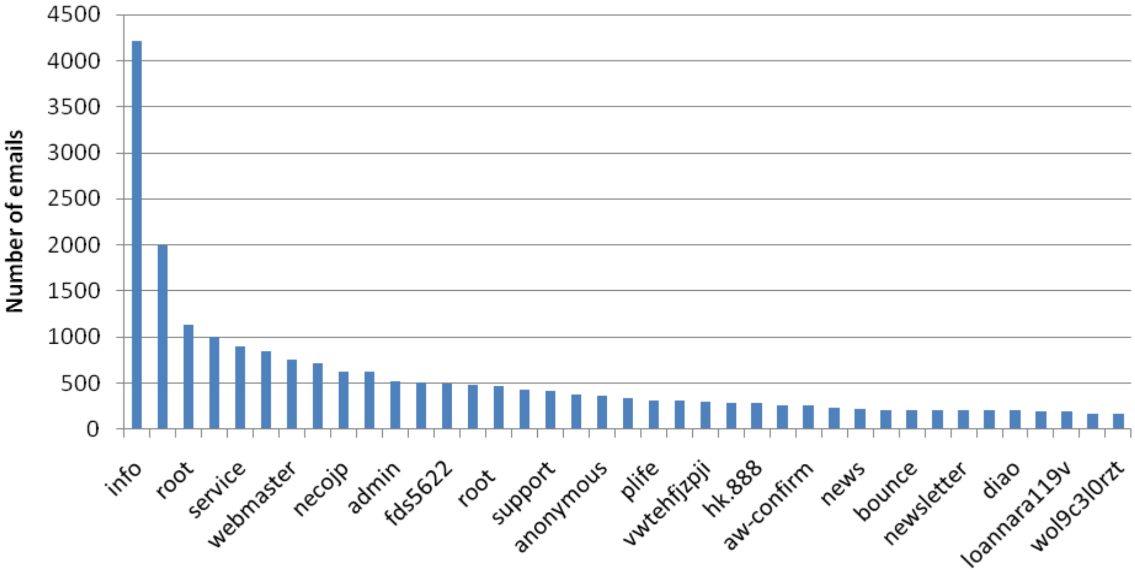

Figure A.1: Frequency of the Same "Name" Usage in Emails Received

Figure A.2 shows the breakdown of the top "names." All names except for 'info' have a significant chunk allotted to the same domain. This makes sense given the bland nature of many of them, which might allow them to pass through a spam filter. There is very little overlap between the domains associated with each of the top "names."

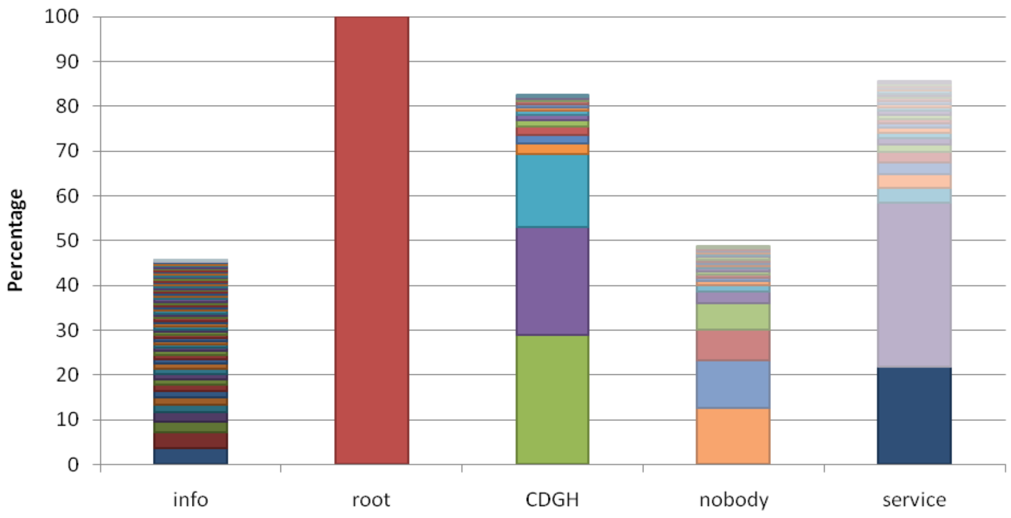

Figure A.2: Number of Different "Domains" Associated with Each of the Top "Names"

Domain persistence (Name@Domain):

Figure A.3 shows the distribution of the "domains" from which emails claim to have been sent. Figure A.4 provides a breakdown of the "Names" associated with the top 4 domains. The domains 'yahoo.com' and 'hotmail.com' have 1/3 to ½ of the data with the same "Name." It is not surprising that the top domains included popular free e-mail account domains. For a spammer, these services make it possible to obtain an almost endless supply of addresses and for them to be "relatively" anonymous.

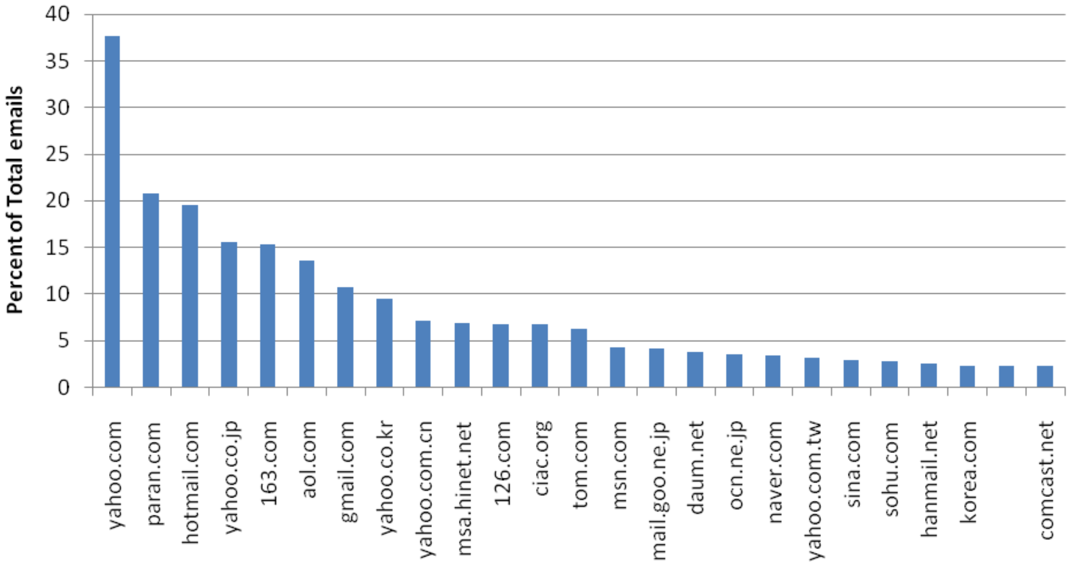

Figure A.3: Frequency of the Same "Domain" Usage in Emails Received (values > 2%)

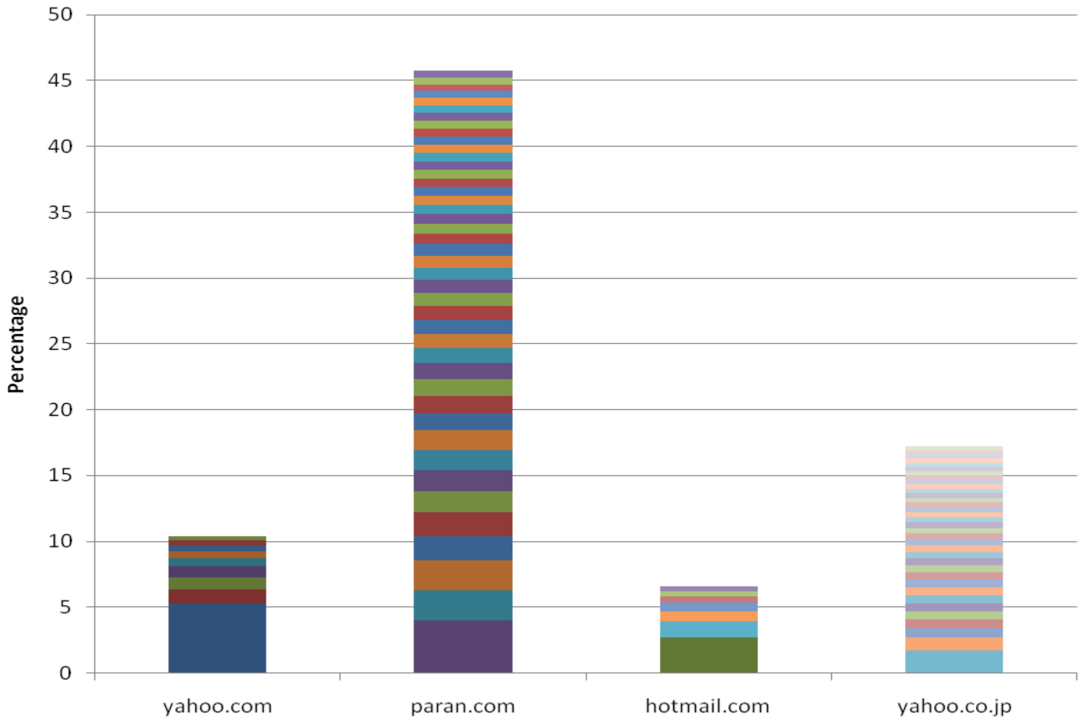

Figure A.4: Number of Different "Names" Associated with Each of the Top "Domains"

A.2 Measures of Deceptiveness

IP Deception

Although the IP from which an email has been sent can generally be trusted (since it is verified on the receiving end), adversaries can "spoof" this information using for example an open proxy server (see section 3.3).

By tracing the email's history through the 'Received' lines, one can detect if the sender tried to mask the sending location by matching the IP indicated by the authenticating servers and the sender. Alternatively, this could be used to see if the sender hacked another computer, used an open relay or a proxy server (see, for instance, https://iihelp.iinet.net.au/Understanding_and_dealing_with_spam_email). We denote this as "IP deceptiveness." A spammer could certainly send an email directly masking their IP using a dial-up connection or other means, which would be quickly identifiable due to the receiving server authentication. Additionally, we are concerned with breaks in multi-received line cases.

$$\text{“IP deceptiveness”} = \begin{cases} 0, & \textit{break in IP history} \\ 1, & \textit{no break in IP history} \end{cases}$$

Figure A.5 shows the trend in "IP deceptiveness" from month to month, with values varying between 0 and 1 where 0 is deceptive and 1 is reliable (based on IP traced history from the 'Received' lines). From the plot, it can be inferred that in 2007 when there is an increase in the number of emails received, the deceptiveness actually increases. Overall, the trend seems to be consistent from year to year.

Figure A.5: TO BE ADDED

From Deception

Another form of deception involves faking the sender of the email by forging the “From:” line. In the emails received, the mail transfer software inserted another “From” line. By comparing the senders in these 2 lines, one can see if the sender attempted to mask their identity. The “deceptiveness” variable is designed such that:

$$\text{“From deceptiveness”} = \begin{cases} 0, & From1 \neq From2 \\ 1, & From1 = From2 \end{cases}$$

Figure A.6 shows the trend in the mean “From deceptiveness” from month to month, with values varying between 0 and 1. From the plot, it can be inferred that in 2007 when there is an increase in the number of emails received, the deceptiveness actually increases. Overall, the trend seems to be consistent from year to year.

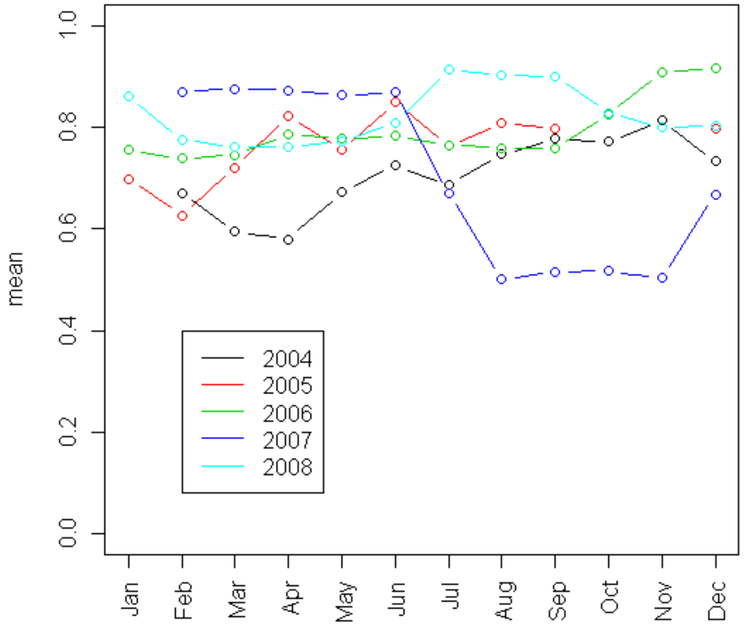

Figure A.6: Mean of “From” Deceptiveness by Month (0 = No Match, 1 = Match)

Figure A.7 shows the trend by day of month for the year 2005. Each month appears fairly similar within a reasonable amount of expected deviations.

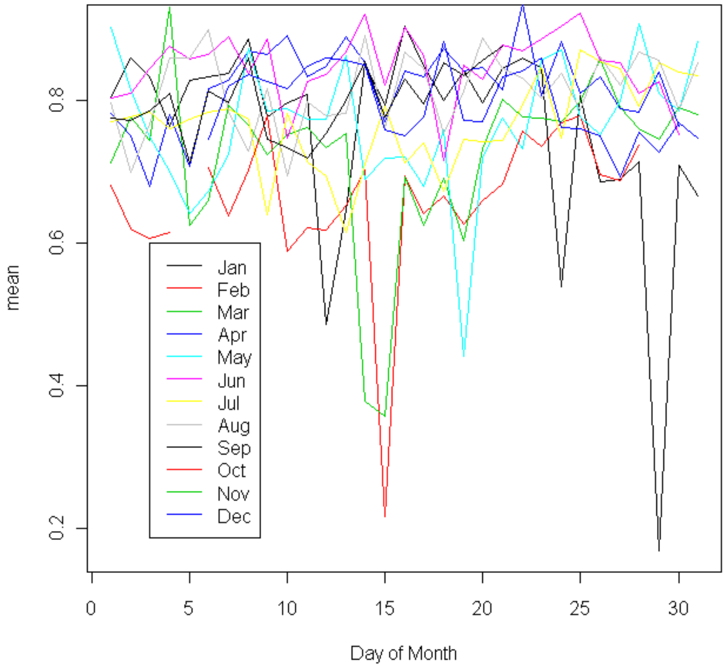

Figure A.7: Validity of IP Links by Day of Month, for the Year 2005 (Mean by Day)

Appendix B. Supplemental Material on Malicious URL Detection Tools

Supplemental Email Analyses

Figure B.1 displays the number of domains per email with a warning by a least one of the three services (McAfee, Norton, Google) for the entire time period considered. That is, at least one service listed the site as dangerous. For example, in the month of Nov of 2008, 22 emails contain 1 domain with a warning by at least one service (red), 103 emails contain 2 domains with a warning by at least one service (green), 62 emails contain 3 domains with a warning by at least one service (purple), and 2 emails contain 4 domains with a warning by at least one service (blue). This produces a total of 189 emails that contain at least one domain with a warning by at least one of the services in that month.

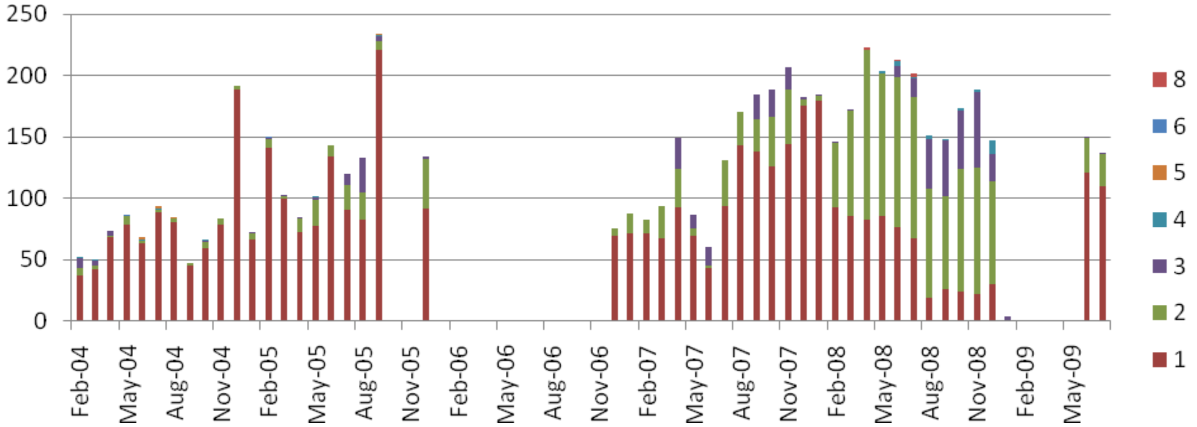

Figure B.1: Number of Domains per Email with a Warning by At Least One of the Three Services

In the same month, Figure B.2 shows that 1,825 emails were received that contained domains with no warnings by any site. Note that the CIAC website containing the email addresses used in this study was taken offline in 2008.

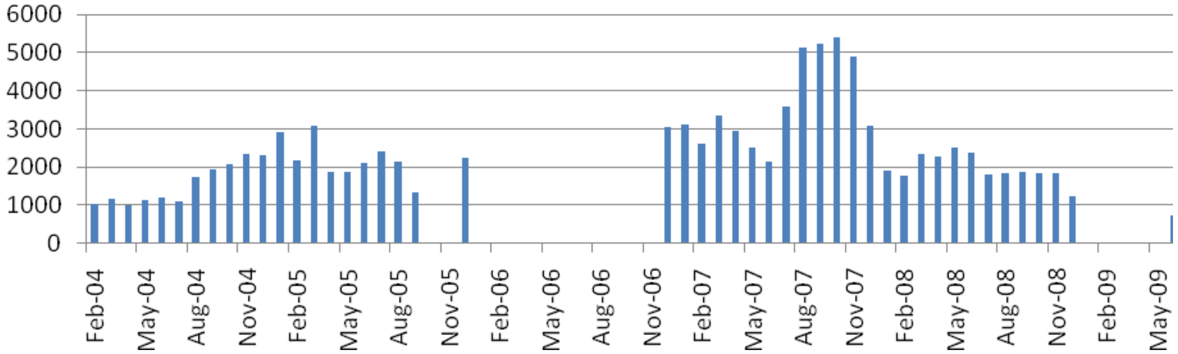

Figure B.2: Number of Domains Containing No Warning by Any of the Four Services

Supplemental Domain Analyses

The following are detailed descriptions of the top five web domains appearing in the email dataset.

- **www.xnabalada.com**
    - 1,394 emails contained this domain. Appeared only during July 17-31, 2007.
    - This domain has not tested by any of the 4 services.
    - In 13 of the emails containing this domain, another domain in the same email was determined malicious according to at least one service. The domain appeared with the following suspicious domains:
        - bule777.com:112
        - gdslys.3653h.com
        - goodcity4.com
        - jow.7cv.com
        - kity123.anytome.com,
        - pure-love.biz
        - serebu.biz
- **www.w3.or**g
    - 1,276 emails contained this domain. Received 2-25 times per day during entire test period.
    - Tested by all services to be safe.
    - Appears to be a resource for improving web pages.
- **cnlinfo.net**
    - 280 emails contained this domain during the test period.
    - Received in 2007 and 2008 only.
    - Appeared as cnlinfo.net/reg.aspx and job.cnlinfo.net frequently)
    - Determined to be malicious by Web of Trust and Site Advisor
    - Tested + determined safe by Norton and not suspicious by Google.
    - McAfee says some downloads (including the Generic PWS.y Trojan download qq2005sp1_PConline.exe) on the site tried to change system settings.
    - Web of Trust warning with low confidence
        - Appeared on an automatically composed list of spamvertised websites
    - Listed as located in China with “Lots of Users”
- **zx-zx-zx.com**
    - 267 emails contained with domain during the time period of June and July of 2008.
    - Hosted in US.
    - The only malicious domain each email.
    - Web of Trust rates it 6 for Trustworthiness, Vendor and Privacy with Confidence 2. Child Safety score is 1 with confidence 2.
        - Appeared on 3 blacklists.
    - McAfee tested and did not find any problems with the site, although 2 users reported phishing attacks (8 users reported spam).
    - Untested by Norton and Google.

Supplemental Web of Trust Analyses

Web of Trust histograms for vendor reliability and privacy, delineated by confidence level, are as follows.

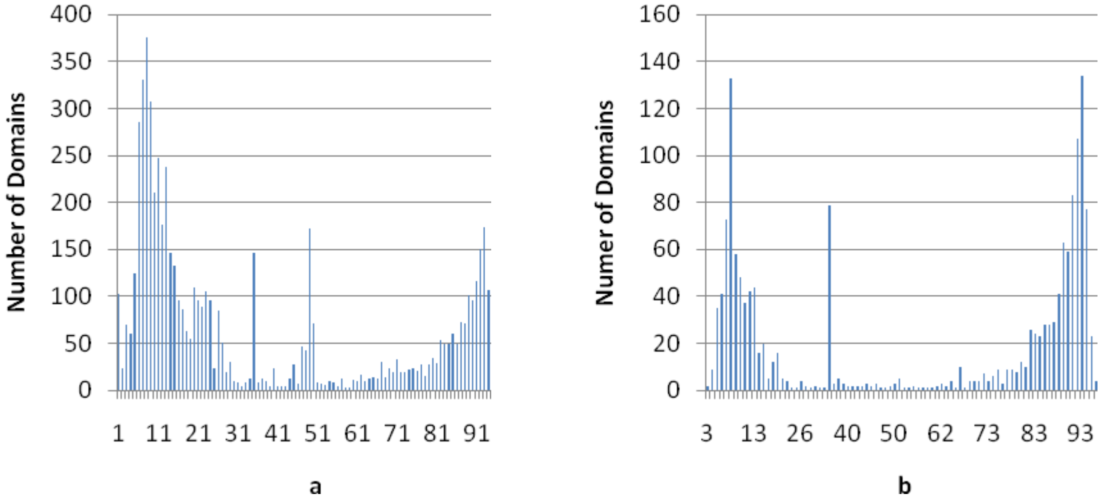

Figure B.3: Histogram of Vendor Reliability Scores for Arbitrary Confidence (a) and High Confidence (b)

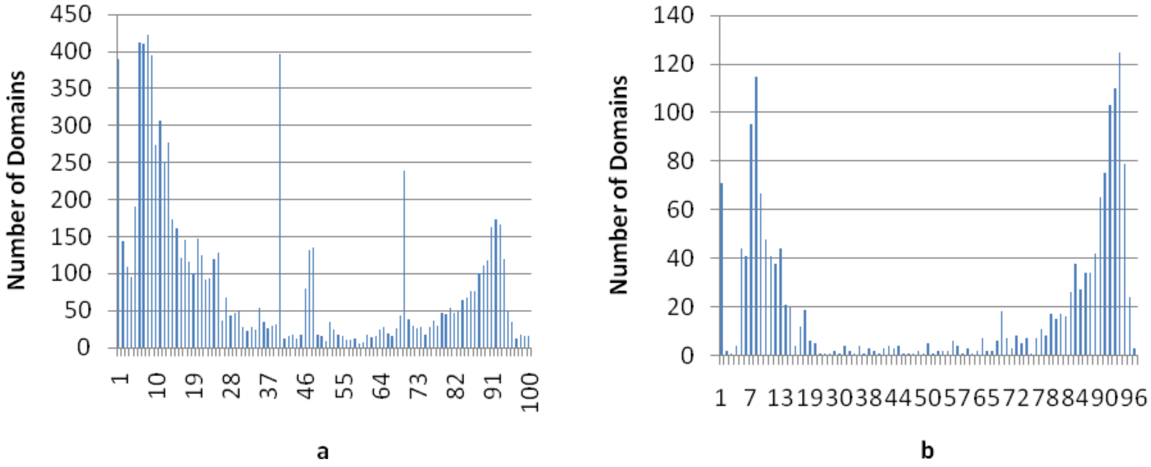

Figure B.4: Histogram of Privacy Scores for Arbitrary Confidence (a) and High Confidence (b)

Supplemental McAfee SiteAdvisor Analyses

A full description of the McAfee test results is shown below. The most common result is that no significant problems were found. Results specific to China and the US are provided, since they make up most of the interesting cases.

	Total	Proportion among all domains in emails	Proportion among tested domains	China	US
Didn't find any significant problems	19,965	0.495	0.871	784	7395
Not Tested	15,448	0.383	NA	NA	NA
Safe to Use	1,260	0.031	0.055	0	1221
promoted through spam	978	0.024	0.043	0	0
links to malware distributor	207	0.005	0.009	71	4
contains malware downloads	156	0.004	0.007	59	34
contains a little malware	127	0.003	0.006	1	1
tries to change browser's homepage	110	0.003	0.005	102	2
several popups	30	0.001	0.001	1	5
change system settings	23	0.001	0.001	14	1
links to browser security breach	10	0.000	0.000	0	6
unauthorized changes to PC	10	0.000	0.000	0	1
tricks you into providing financial info	9	0.000	0.000	0	0
engaged in negative activities	7	0.000	0.000	0	5
spams if sign up on site	7	0.000	0.000	0	2
affiliated with red sites	3	0.000	0.000	0	0
change homepage and popups	3	0.000	0.000	0	2
captures mistyped url	2	0.000	0.000	0	1
promotes malware	2	0.000	0.000	0	0
distributes others software	1	0.000	0.000	0	0
downloads are free of malware	1	0.000	0.000	0	0
misleading claims of work-at-home	1	0.000	0.000	0	1
misleading offers	1	0.000	0.000	0	1
some popups	1	0.000	0.000	0	0

Table B.1: A Full Description of the McAfee Test Results

Histograms for the countries of origin of different kinds of threats detected by McAfee SiteAdvisor are given in Figures B.5-B.8.

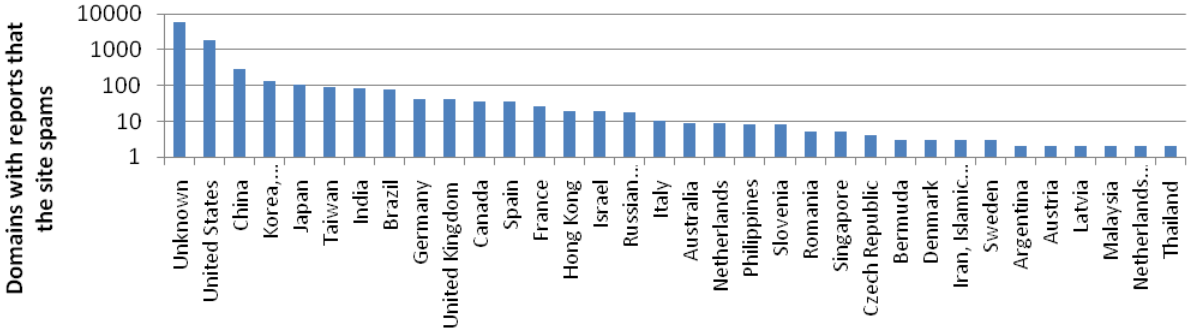

Figure B.5: Countries of Origin for Domains with McAfee Site Advisor Reports of Spams

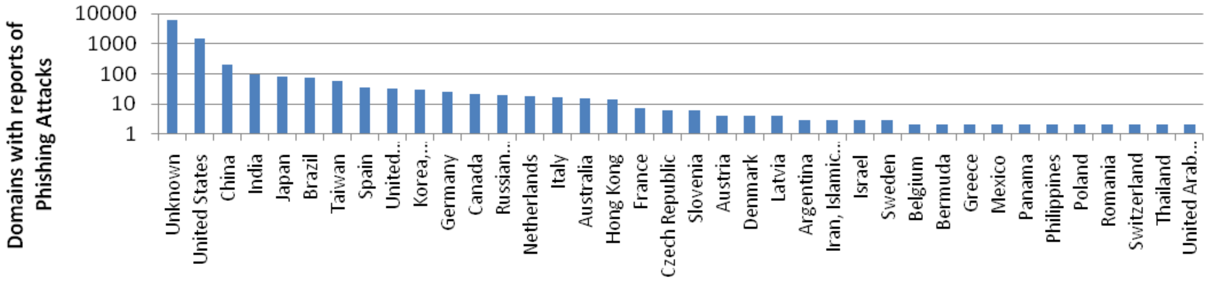

Figure B.6: Countries of Origin for Domains with McAfee Site Advisor Reports of Phishing Attacks

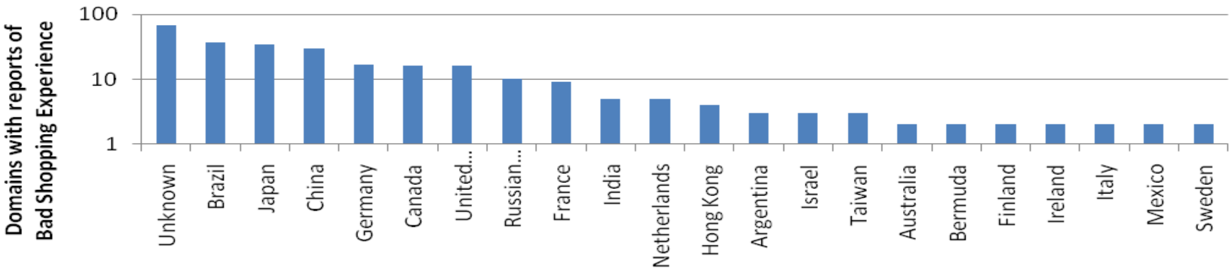

Figure B.7: Countries of Origin for Domains with McAfee Site Advisor Reports of Bad Shopping

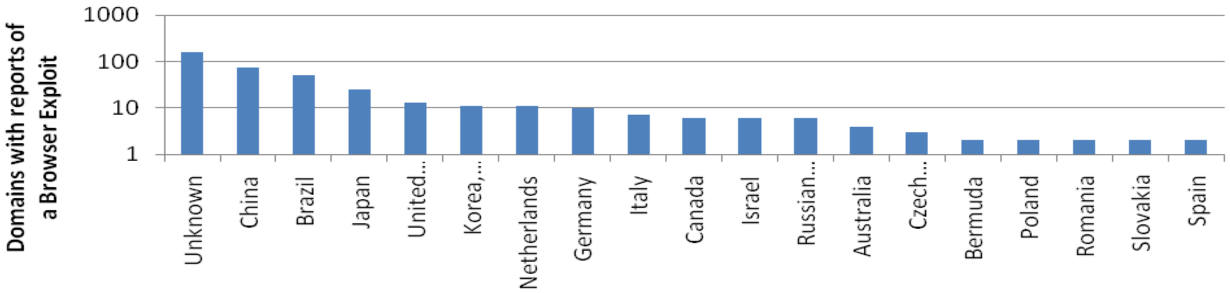

Figure B.8: Countries of Origin for Domains with McAfee Site Advisor Reports of a Browser Exploit

Supplemental Norton SafeWeb Analyses

A histogram for the number of identity threats found in domains tested by Norton SafeWeb is as follows.

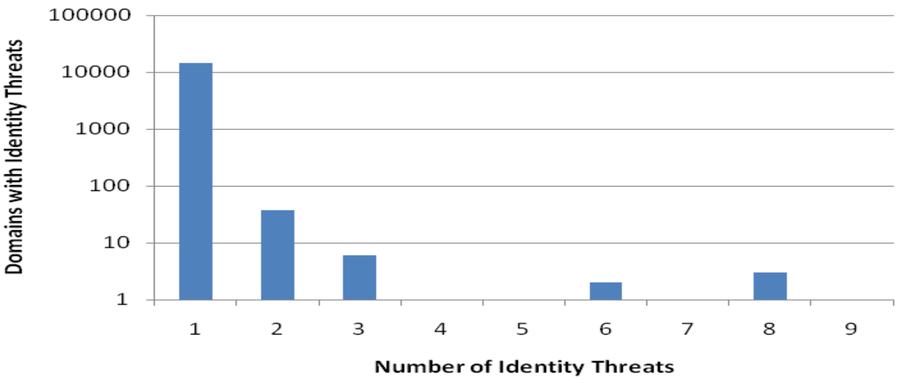

Figure B.9: Number of Different Computer Threats Presented in Domains Tested by Norton SafeWeb

<u>Supplemental Google SafeBrowse Analyses</u>

Figures B.10-B.15 show more of the analyses that are possible via Google SafeBrowse. These establish that there are a varying number of routes by which malware can be hosted and distributed.

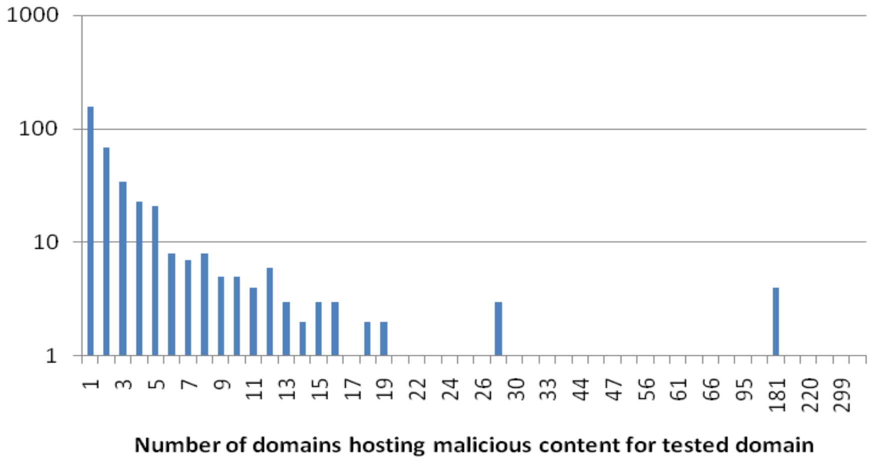

Figure B.10: Number of Domains Found by Google SafeBrowse Hosting Malicious Content for Domain

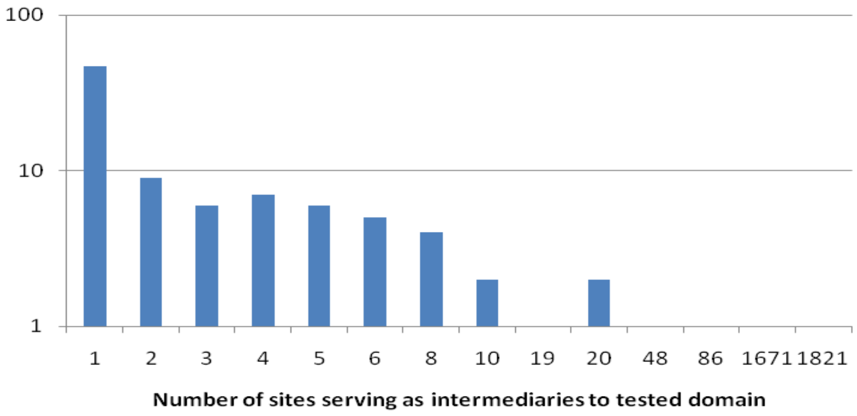

Figure B.11: Number of Sites Found by Google Safebrowse Serving as Intermediaries to Tested Domain

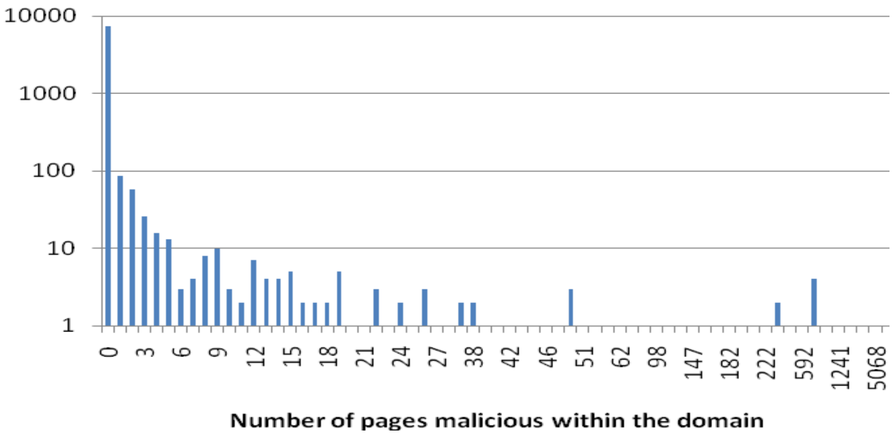

Figure B.12: Number of Malicious Pages Found by Google Safebrowse within Infected Domains

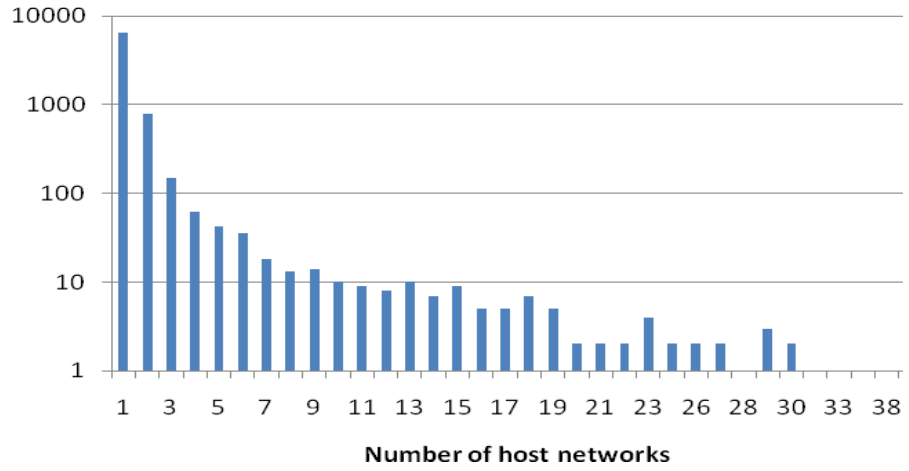

Figure B.13: Number of Host Networks Determined by Google Safebrowse for All Domains

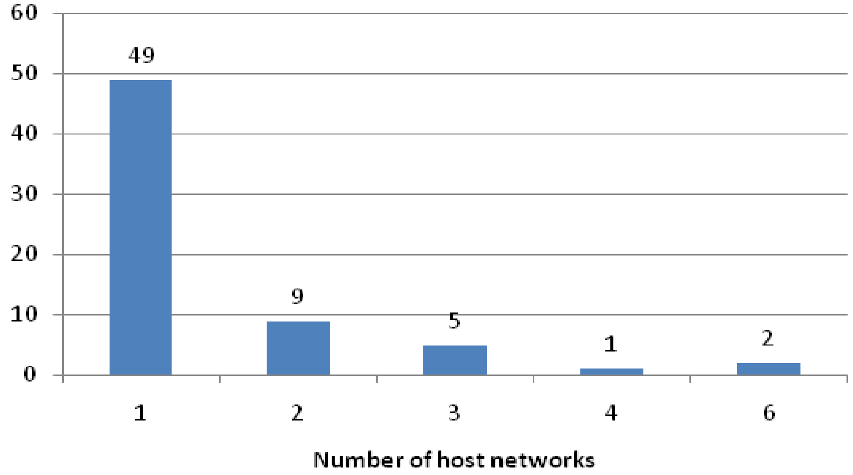

Figure B.14: Number of Host Networks Determined by Google Safebrowse for Suspicious Domains

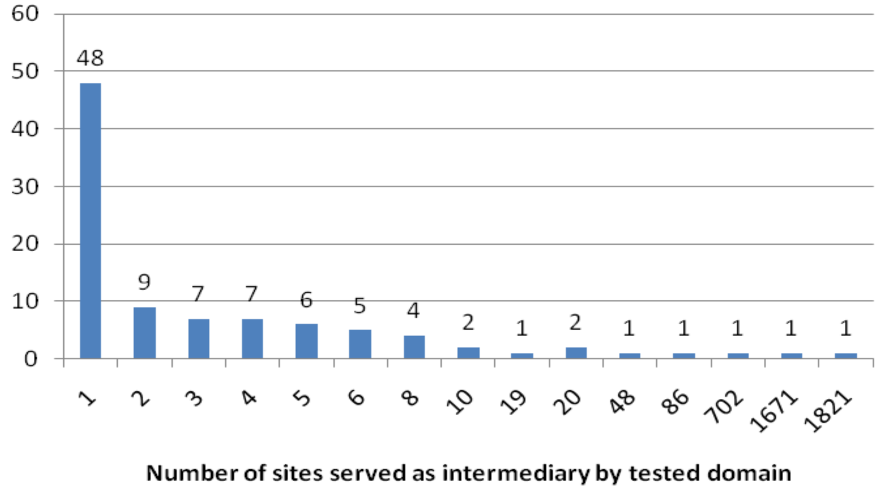

Figure B.15: Number of Sites Serving as Intermediary to a Domain Tested by Google SafeBrowse

Supplemental Data Trend Analyses

Figure B.16 represents a similar bootstrapping analysis to Figure 5.14, using the vendor reliability scores instead of trustworthiness scores. Only scores with a confidence level of greater than 1 out of 5 are used. The results are comparable.

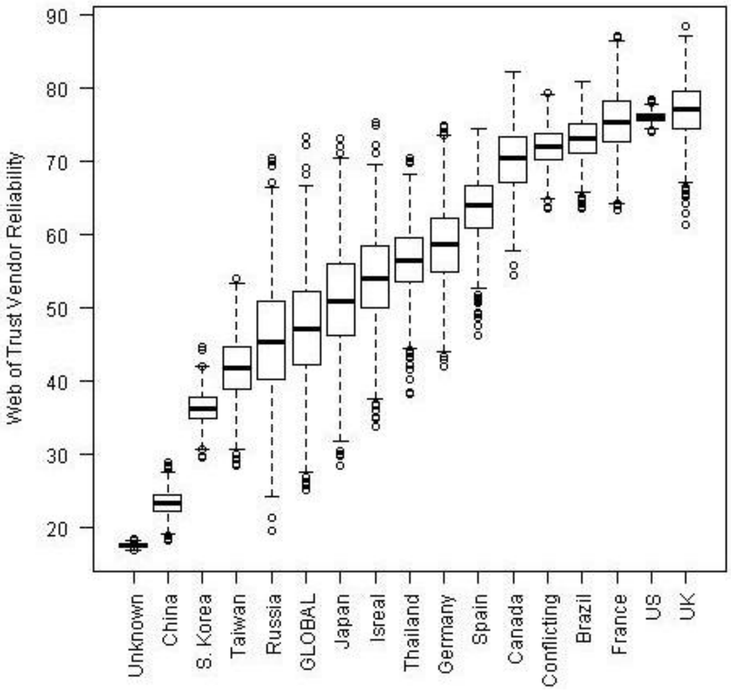

Figure B.16: Bootstrapped Mean Vendor Reliability Scores with a 95% Confidence Level, by Country

References

1. A3C Connection. Reading email headers. Technical Report, Academic Computing and Communications Center, University of Illinois at Chicago, October-December 2000. Accessed at http://www.uic.edu/depts/accc/newsletter/adn29/headers.html.

2. AVG.com. AVG anti-virus and security software. Press release, 2009. Accessed at http://www.avg.com/product-avg-anti-virus-free-edition.

3. N. Chantler. *Profile of a Computer Hacker*. Infowar, 1996.

4. T. Chen and J. Robert. Worm epidemics in high-speed networks. *Computer*, 37(6):48-53, 2004. Accessed at http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=1306386.

5. Digital Software Development. How to read email headers. Technical Report, Digital Software Development, 2009. Accessed at http://www.emailaddressmanager.com/tips/email.html.

6. S. Eltringham (editor), Computer Crime and Intellectual Property Section, US Department of Justice. *Prosecuting Computer Crimes*. Office of Legal Education Executive Office for US Attorneys, 2007. Accessed at http://www.usdoj.gov/criminal/cybercrime/ccmanual/index.html.

7. A. Emigh. Online identity theft: phishing technology, chokepoints, and countermeasures. Technical Report, Infosec Technology Transition Council, Department of Homeland Security, 2005. Accessed at http://www.cyber.st.dhs.gov/docs/phishing-dhs-report.pdf.

8. J. Evers. Report: net users picking safer passwords. *ZDNet News*, December 16, 2006. Accessed at http://news.zdnet.com/2100-1009_22-150640.html.

9. L. Garber. Denial-of-service attacks rip the internet. *Computer*, 33(4):12-17, 2000. Accessed at http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=839316.

10. Government Accountability Office (GAO). Cybercrime: public and private entities face challenges in addressing cyber threats. Technical Report GAO-07-705, US Government Accountability Office, 2007. Accessed at http://www.gao.gov/products/GAO-07-705.

11. S. Hansman and R. Hunt. A taxonomy of network and computer attacks. *Computers and Security*, 21:31-43, 2005. Accessed at http://linkinghub.elsevier.com/retrieve/pii/S0167404804001804.

12. R. Hollinger. Computer hackers follow a Guttman-like progression. *Sociology and Social Research*, 72:199-200, 1988. Accessed at http://www.phrack.com/issues.html?issue=22&id=7.

13. J. Howard. An analysis of security incidents on the internet, 1989-1995. Ph.D. Thesis, Carnegie Mellon University, 1997. Accessed at http://www.cert.org/archive/pdf/JHThesis.pdf.

14. J. Howard and T. Longstaff. A common language for computer security incidents. Technical Report SAND98-8667, Sandia National Laboratories, 1998. Accessed at http://www.cert.org/research/taxonomy_988667.pdf.

15. Internet Crime Complaint Center. 2008 Internet crime complaint report. Technical Report, Internet Crime Complaint Center, 2008. Accessed at http://www.ic3.gov/media/annualreports.aspx.

16. G. Jones. The 10 most destructive PC viruses of all time. *VARBusiness Magazine*, July 7, 2006. Accessed at http://www.crn.com/it-channel/190301109.

17. B. Kehoe. *Zen and the Art of the Internet: a Beginner's Guide*. Prentice Hall, 1992. Accessed at http://www-rohan.sdsu.edu/doc/zen/zen-1.0_toc.html.

18. H. Kikuchi, M. Terada, N. Fukuno, and N. Doi. Estimation of increase of scanners based on ISDAS distributed sensors. *Journal of Information Processing*, 16:100-109, 2008. Accessed at http://www.jstage.jst.go.jp/article/ipsjjip/16/0/16_100/_article.

19. M. Kjaerland. A taxonomy and comparison of computer security incidents from the commercial and government sectors. *Computers and Security*, 25:522-538, 2006. Accessed at http://linkinghub.elsevier.com/retrieve/pii/S0167404806001234.

20. E. Kowalksi, D. Cappelli, and A. Moore. Insider threat study: illicit cyber activity in the information technology and telecommunications sector. Technical Report, National Threat Assessment Center, United States Secret Service, 2008. Accessed at http://www.secretservice.gov/ntac.shtml.

21. B. Landreth. *Out of the Inner Circle: a Hacker's Guide to Computer Security*. Microsoft Press, 1985.

22. M. Landler and J. Markoff. Digital fears emerge after data siege in Estonia. *The New York Times*, May 29, 2007. Accessed at http://www.nytimes.com/2007/05/29/technology/29estonia.html.

23. C. Landwehr, A. Bull, J. McDermott, and W. Choi. A taxonomy of computer program security flaws, with examples. *ACM Computing Surveys*, 26(3):211-254, 1994. Accessed at http://chacs.nrl.navy.mil/publications/CHACS/1994/1994landwehr-acmcs.pdf.

24. E. Levy. The making of a spam zombie army: dissecting the Sobig worms. *IEEE Security and Privacy*, 1(4):58-59, 2003. Accessed at http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=1219071.

25. H. Lipson. Tracking and tracing cyber attacks: technical challenges and global policy issues. Technical Report CMU/SEI-2002-SR-009, Carnegie Mellon University, 2002. Accessed at http://www.sei.cmu.edu/pub/documents/02.reports/pdf/02sr009.pdf.

26. J. Markoff. Before the gunfire, cyberattacks. *New York Times*, August 12, 2008. Accessed at http://www.nytimes.com/2008/08/13/technology/13cyber.html.

27. C. Meyers, S. Powers, and D. Faissol. Taxonomies of cyber adversaries and attacks: a survey of incidents and approaches. Technical Report, Lawrence Livermore National Laboratory, 2009.

28. K. Mitnick. *The Art of Deception: Controlling the Human Element of Security*. Wiley, 2002.

29. J. Murphy, P. Elmer-Dewitt, and M. Krance. The 414 gang strikes again. *TIME Magazine*, August 29, 1983. Accessed at http://www.time.com/time/magazine/article/0,9171,949797,00.html.

30. R. Rantala. Bureau of Justice Statistics special report: Cybercrime against businesses, 2005. Technical Report NCJ 221943, US Department of Justice, 2008. Accessed at http://www.ojp.usdoj.gov/bjs/abstract/cb05.htm.

31. E. Raymond. *The Art of UNIX Programming*. Addison-Wesley Professional Computing Series, 2003. Accessed at http://www.faqs.org/docs/artu/.

32. M. Rogers. A new hacker taxonomy. Technical Report, University of Manitoba, 1999. Accessed at http://homes.cerias.purdue.edu/~mkr/hacker.doc.

33. M. Rogers. Psychological theories of crime and hacking. Technical Report, University of Manitoba, 2000. Accessed at http://homes.cerias.purdue.edu/~mkr/crime.doc.

34. M. Rogers. A social learning theory and moral disengagement analysis of criminal computer behavior: an exploratory study. Ph.D. Thesis, University of Manitoba, 2001. Accessed at http://homes.cerias.purdue.edu/~mkr/cybercrime-thesis.pdf.

35. M. Rogers. A two-dimensional circumplex approach to the development of a hacker taxonomy. *Digital Investigation,* 3:97-102, 2006.

36. D. Russell and G. Gangemi. *Computer Security Basics*. O'Reilly, 1991.

37. W. Scherlis. DARPA establishes computer response team. Press Release, Defense Advanced Research Projects Agency (DARPA), 1988. Accessed at http://www.cert.org/about/1988press-rel.html.

38. E. Schultz. DOE's Computer Incident Advisory Capability (CIAC). Technical Report UCRL-JC-105099, Lawrence Livermore National Laboratory, 1990. Accessed at http://www.osti.gov/bridge/product.biblio.jsp?osti_id=6054719.

39. W. Schwartau. *Information Warfare: Cyberterrorism: Protecting Your Security in the Electronic Age.* Thunder's Mouth Press, 1996. Accessed at http://www.winnschwartau.com/resources/IW1.pdf.

40. D. Schweitzer. Why I don't want you to buy a Mac. *PC World*, July 9, 2009. Accessed at http://www.pcworld.com/article/168133/why_i_dont_want_you_to_buy_a_mac.html.

41. E. Shaw, K. Ruby, and J. Post. The insider threat to information systems: the psychology of the dangerous insider. *Security Awareness Bulletin*, 2:1-10, 1998. Accessed at http://www.pol-psych.com/sab.pdf.

42. A. Smith and W. Rupp. Issues in cybersecurity: understanding the potential risks associated with hackers/crackers. *Information Management and Computer Security*, 10(4):178-183, 2002. Accessed at http://www.emeraldinsight.com/Insight/viewContentItem.do?contentType=Article&contentId=862828.

43. Stopspam.org. Reading email headers. Technical Report, stopspam.org, 2008. Accessed at http://www.stopspam.org/index.php?option=com_content&id=45.

44. M. Soper. Digital picture frames- now with free malware! *MaximumPC Magazine,* February 16, 2008. Accessed at http://www.maximumpc.com/article/digital_picture_frames_now_with_free_malware.

45. C. Stoll. Stalking the wily hacker. *Communications of the ACM*, 31(5):484-497, 1988. Accessed at http://pdf.textfiles.com/academics/wilyhacker.pdf.

46. Symantec.com. Symantec gateway security archive. Technical report, Symantec.com, 2009a. Accessed at http://www.symantec.com/avcenter/security/Content/Product/Product_SGS.html.

47. Symantec.com. Symantec unveils new model of consumer protection codenamed "Quorum". Press release, July 9, 2009b. Accessed at http://www.symantec.com/about/news/release/article.jsp?prid=20090706_02.

48. C. Taylor, J. Alves-Foss, and V. Freeman. An academic perspective on the CNSS standards: a survey. In *Proceedings of the 10th Colloquium for Information Systems Security Education*, pages 39-46, Adelphi, MD, 2006. Springer. Accessed at http://www.cisse.info/colloquia/cisse10/proceedings10/pdfs/papers/S02P01.pdf.