Some Observations on the Occurrences of Phishing

Jaime Lara
jlaraz2@utpa.edu

Kris Rios
kriosz1@yahoo.com

Rafael Salazar
coolralph_s@hotmail.com

Kai S. Koong
koongk@utpa.edu

Lai C. Liu
liul@utpa.edu
Computer Information Systems and Quantitative Methods Department
The University of Texas Pan American
Edinburg, Texas 78539, USA

Abstract

Many malicious methods exist to steal information from individuals and organization. This study examines occurrences of phishing cases that were reported by the Anti-Phishing Working Group (APWG) between 2003 and 2008. Specifically, the items examined included the number of phishing attacks and the number of phishing sites identified by years and by quarters. The year 2007 had the highest average number of reports than in any other year, where September was the month that had the highest number of phishing attacks than any of the months and years studied. The year 2007 also had the highest number of new phishing sites reported for the years studied, where April was the peak for all the months and years. It can be concluded that the drop in number of sites reported from 2007 to 2008 could have played a big role in the decreasing rate in reports. Finally, other trends were analyzed and implications about future occurrences as well as recommendations for best practices are also provided.

Keywords: phishing, computer security, computer crime

1. INTRODUCTION

Individuals and organizations throughout the world use the Internet for a multitude of reasons: e-mail, instant messaging, document creation, e-commerce, online banking, credit card management, shopping, and the general sharing of information. The innovation of technology demands the fastest way of doing business or personal demands. E-mail and web browsers have become major ways of how people use Internet-connected devices. According to a Pew February-March

2007 survey, 91% of American Internet users send or read e-mail (Wang *et al.,* 2009; Pew, 2007). Personal and business information along with the ability to manage it in a safe manner is critical to a user's online experience and general well-being. Should the security of such information be compromised, a user could literally lose everything. Sometimes, information across these various services and websites are so intertwined that an attack or breach of security can extend and lead to a complete loss of identity. The scale of such a loss is broad in that one mistake can cause great chaos and stress for a person or an organization. To their detriment, the public is slowly realizing what is happening, with regard to identity thefts. An attacker can use a variety of malicious methods to exploit a user. This study will focus on phishing, a method of Internet identity theft.

Phishing is a form of Internet identity theft that employs both social engineering and technical subterfuge to steal consumers' personal identity data and financial account credentials. Social engineering schemes use 'spoofed' e-mails to lead consumers to counterfeit websites designed to trick recipients into divulging financial data such as account usernames and passwords. Hijacking brand names of banks, e-retailers and credit card companies, phishers often convince recipients to respond. Technical subterfuge schemes plant crimeware onto PCs to steal credentials directly, often using key logging systems to intercept consumers online account user names and passwords. Regularly, local and remote navigational infrastructures are corrupted to misdirect consumers to counterfeit websites and to authentic websites through phisher-controlled proxies that can be used to monitor and intercept consumers' keystrokes (APWG, 2008).

While the phishers develop ever more sophisticated attack vectors, businesses flounder to protect their customers' personal data and look to external experts for improving email security. Customers too have become wary of “official” email, and organizations struggle to install confidence in their communications (Ollmann, 2008). Each of these threats is driven by criminals looking for financial gain, not attention-seeking hackers (Anderson & Moore, 2007). Internet users are vulnerable to this scam due to all the types of hackers and 'black hats' out there waiting for someone to get online and bite down on the bait. With the click of a button, even a high-school hacker can get confidential marketing and strategic plans, financial statements, and customer information (McCune, 1998; Akerman, 2003; Koong *et al.*, 2008).

This subject is very hazardous for many corporations due to loss of important information that can lead the company down the drain. Security organizations and companies have done research and development on anti-phishing techniques and tools. These include basic changes in the E-mail infrastructure to help alleviate Spam, more widespread deployment of anti-Spam, anti-Malware, personal firewall products, privacy protection software, and stronger authentication for electronic transactions. Some of them have good effects on decreasing the number of phishing; unfortunately, phishing attacks are growing both in numbers and in complexity (Ding & Li, 2006).

2. STATEMENT OF THE PROBLEM

Billions of dollars are lost every year by individuals and organizations via phishing. Phishing attacks in the United States soared in 2007, as $3.2 billion was lost to these attacks, according to a survey by Gartner, Inc (Gartner, 2007). Phishing is an innovative way of scamming individuals and businesses because of the demand for faster uses of technology for e-commerce. People become victimized through e-mail and websites that appear to show the true content of a particular company, when it really is just an engineered scam created by a criminal just to acquire users' financial credentials. Individuals who use the Internet to perpetrate crimes are called 'black hats' or carders. Since security systems from businesses engaged in m-commerce and e-business tend to lag behind on new technology, black hats go 'phishing' for potential victims over the web (Koong *et al.*, 2008). The total number of unique phishing reports submitted to APWG in January 2008 was 29,284, an increase of over 3,600 reports from the previous month (APWG, 2008).

Society is losing tremendous amounts of money due to phishing and its scams. Black hats also target foreign citizens who have been affected severely by phishing. Criminals, on the other hand, are improving their methods and taking in the rewards as they come. Although, in the past, most criminals only aimed their attacks at consumers in English-speaking countries, phishers have also launched attacks against citizens of Germany and Brazil (Sullins, 2006; Blau, 2004). This scheme is happening everywhere in the world and many individuals and organizations are taking notice. Due to phishing attacks across international borders, broad cooperation among various law enforcement authorities and Internet service providers (ISPs) are required to effectively deal with the issue. Greater effort must be put forth in the education of individuals and organizations about the problem (Vijayan, 2004). People have become increasingly aware of the pervasive threats to information security. There are a variety of solutions now available for solving the problem of information insecurity such as improving technologies, including the application of advanced cryptography, or techniques, such as performing risk analyses and risk mitigation (Bresz, 2004; Sasse, Brostoff & Weirich, 2004; Workman, 2008).

3. STATEMENT OF OBJECTIVES

The Anti-Phishing Working Group (APWG) website located at http://www.antiphishing.org/report_phishing.html takes e-mails from anybody who has or had a problem dealing with phishing and creates reports on those cases on a monthly, semi-annually, and annually basis. This research analyzed the number of phishing attacks and new phishing websites reported to the APWG for the period of 2003 to 2008. It is believed, that since some phishing websites are reported as a result of a phishing attack, successful or not, it can be concluded that these two metrics are linked.

Individuals will find this research useful when dealing with the amount of phishing reports that were received in the span that we will cover. It should also help organizations see the outcomes of the attack through email and websites and the level of general vulnerability. Computer information system scholars that specialize in malware or simply phishing, will find this study to be valuable in helping them understand the affects of phishing.

4. METHODOLOGY

The targeted population of this study consists of phishing attack and new phishing website reports made throughout the world for a period of late 2003 through the end of 2008. The data used in this study was obtained from the Anti-Phishing Working Group (APWG) via the organization's website at http://www.antiphishing.org. This organization obtained their results via various methods, including email submissions to reportfishing@antifishing.org during these respective years. The data extracted from these reports are:

1) Number of phishing attacks reported from November 2003 to December 2008.
2) Number of new phishing websites identified from July 2004 to December 2008.

First, yearly totals of phishing attacks reported and new phishing websites identified were generated. With this, we can examine the yearly bar graphs and quarterly line graph and how they are skewed for each of the information gathered. We are able to study what the correlation is between the number of phishing attacks reported and how many new phishing websites were identified to link observations to hypothesize what might have caused such trends. This information is presented in Figures 1 and 2 (in Appendix A).

Second, the actual monthly reports were grouped by quarters, for every year, and plotted into a graph to examine the general correlations and trends that may exist in the information collected. Phishing attack reports from January through March, April through June, July through September, and October through December were added to make the first quarter (Q1), second quarter (Q2), third quarter (Q3), and fourth quarter (Q4), respectively, for each of the years. The results are shown in Figure 3 (in Appendix A). Figure 4 (see Appendix A) is similar to Figure 3, in that same analytical technique was applied to the plotted and graphed data, however, the data represented in Figure 4 are new phishing sites that were detected and reported during the same period. Averages are included to illustrate a middle ground that exists between each year for each quarter.

Finally, raw monthly data for phishing attack reports were overlaid for each of the years studied. The average was also calculated for the reports for each entire year. After that, the average was calculated from both of the averages previously computed. With these calculations, we can observe where the reports peaked and where they leveled off for every month of each year. This information is both, plotted and sketched into a graph to provide two different perspectives of the same data which is highlighted in Figure 5 (see Appendix A). The same systematic method employed in Figure 5, was also used to calculate the average number of new phishing sites identified for each month of each year, as well as the average for each year at its entirety. Then the average of the averages was taken and shown. A table and a graph were also created where we can see how the websites numbers increased and decreased in the different months of the different years. This information is detailed in Figure 6 (see Appendix A).

5. FINDINGS

Our findings are mostly drawn from the information and data presented in the Phishing Activity/Attack Trends Report created by the Anti-Phishing Working Group (APWG). The data presented here was extracted from several of the trend reports created by the APWG. Most notably, various reports from 2004 and reports from January of each year after that happened to contain the appropriate figures required by this study. The scope of this finding extends from November 2003 to December 2008 and will analyze the phishing trends between months of each year, quarters of each year, as well as an overall year over year trend. When referring to phishing trends as a whole, this study is referring to new phishing incidents reported and new phishing websites identified. It is very important to note that the findings do not start at the same time period. To be more specific, the data for phishing attacks reported start on November 2003, and yet, data for the number of new phishing websites identified starts in July 2004. The APWG employed a new methodology around this time, which added the number of phishing websites identified as a new metric for their reports. As of July 2004, "we are introducing a new methodology, which provides a measurement of phishing activity based on the number of fraudulent 'baiting' websites extracted from phishing email messages, in lieu of counting the email messages themselves as presented in previous reports" (APWG, 2004). For this reason, information prior to the third quarter of 2004 is not taken into account when formulating a hypothesis. Furthermore, all information, with respect to phishing attacks and websites reported, is included for completeness.

A phishing incident report is made when an individual or organization has noticed that their information security has been compromised. Please note, that the base finding clearly states that the phishing scams are often carried out by malicious websites via malicious emails, and that not all attacks are reported. Many individuals or organizations do not know where to report such an attack. Consider that just because an attack occurred, it does not mean that it was reported. It is unknown how many phishing attacks go unreported each year. The same goes for the number of phishing websites that go undocumented each year as a result of phishing attacks not reported.

The remainder of this section examines the data presented in three forms: by year, by quarter, and by month, over the six years. The figures, grouped in twos, each provide different pictures where one can gather a different set of assumptions based on the way they flow, illustrating different viewpoints or perspectives.

More phishing websites exist now, more than ever. Phishing websites, despite a slight 23.45% reduction in growth, seem to have had no effect on the rising number of phishing attack reports. From Figures 1 and 2 we can assume that there is a correlation between the number of phishing attacks reported and the number of new phishing websites identified. There seems to be a leveling off of new phishing attack reports seen in the 2007 to 2008 period, rather than a decline. The decline in 2008 of new phishing websites identified, seems to be the catalyst for this drop off. The low level of phishing websites identified in 2004 and 2005, seen against the much higher level of new phishing attack reports around the same time, suggests that people were less informed about phishing attacks. The steep 338.53% increase in the number of new phishing websites identified from 2005 to 2006, along with another increase from 2006 to 2007, continue to suggest that these malicious attackers realized that people were generally uneducated about phishing attacks. The leveling off of phishing attack reports could also suggest that people were starting to become more informed about phishing. For these reasons, this study will not place too much weight on the numbers from 2005.

It is also interesting to note that the Anti-Phishing Working Group, an industry, government and law enforcement association focused on eliminating the identity theft and frauds that result from the growing problem of phishing, crimeware, and email spoofing was founded in 2003 by leading software security and financial industry companies (AWPG, 2008). This could have also impacted the initial numbers, as this organization was still in its infancy. It was also mentioned earlier, that in July 2004, a shift in methodology took place, with regard to metrics presented in their reports. Coincidentially, the numbers before the second half of 2004 are not high enough to be reasonable or suitable for this study, but are included for the sake of presenting the reported data in its entirety.

In Figures 3 and 4 we divide each year into four quarters. The purpose of this is to look into more detail at the trends of each quarter and how they relate to the other corresponding quarters of each year. The idea for this came from the financial industry's method of comparing the current quarter with that of the corresponding quarter of the previous years. In Q3 to Q4 of 2006, there is a large increase in the number of phishing websites identified; however, the number reported phishing attacks seemed to stabilize. By this time, newer versions of popular web browsers like Microsoft's Internet Explorer and Mozilla's Firefox and email clients like Microsoft's Windows Live Mail and Mozilla's Thunderbird, were being released with new and/or improved anti-phishing features that helped protect users from phishing scams. It seems that this continued to help users of the web, as the phishing attack report rates took a dip in the second quarter (Q2) of 2007 again. The new phishing website creation rates seemed to surge. In 2008, the rate of new phishing sites identified appears to decline while the rate of phishing attack reports mimicked the 2007 rates almost perfectly.

Figures 5 and 6 are present to show monthly trends across a year, how each month compares with past and future months, and the averages of both. On average, there is an increase in the number of phishing sites created, month over month, each year. With the expectation of late 2006 and most of 2007, a bad year all-together for phishing attacks, both websites and attacks seem to be in line with each other. Spikes in phishing attack reports seem to occur during the summer and early fall of each year. Perhaps it is not too far reaching to point out that this is the time when young "entrepreneurs" have time off from school and other duties. Even though demographics of the individuals that operate and create phishing websites are not part of the scope of this study, it makes sense that maybe a different breed of "script kiddie" may be roaming the Internet with their own malicious website. On average, one of the lower points of the year for both attacks reported and websites identified seems to show as the end of December approaches. Perhaps, this is when phishers prefer to take a holiday rest.

6. CONCLUSION

It is evident that phishing reports kept increasing from 2005 through 2008, although there were minor decreases in reports in several months. The years 2003 to 2005 had the least phishing attacks and least phishing sites reported. The years 2006 and 2007 had the biggest jump of phishing reports and phishing sites reported than the rest of the years. It can be assumed that in the year 2005, people were not so much aware of the problem. Phishing, from widespread public view, was fairly new at the time and was barely starting to catch the eyes of crime perpetrators.

Moreover, the increasing level of phishing attacks across the years is real, but a decreasing rate of the reports is also evident. Perpetrators are always innovating phishing technology to facilitate crime or to keep up with new security systems implemented by computer systems users. With this said, ways of phishing improved from 2005 through 2007. The year 2007 had the highest average number of reports than in any other year, where September was the month that had the highest number of phishing attacks than any of the months and years studied. 2007 also had the highest number of new phishing sites reported for the years studied, where April was the peak for all the months and years. It can be concluded that the drop of number of sites reported from 2007 to 2008 plays a big role in the decreasing rate in reports.

All phishing reports from 2005 through 2008 had a maximum point on June, decreased in July, increased in August with the exception of 2005, then had a decreasing slope from October through December, and then increased again in the beginning of the next year's first month. With this, we can surmise that people are highly vulnerable to attacks mostly in the month of June and October, which can also be proven as they have the two biggest average monthly reports. Individuals should be more careful in their lookout of scams during these months to avoid becoming a victim of crime.

7. LIMITATIONS AND DIRECTIONS FOR FUTURE RESEARCH

A major limitation encountered while harvesting information about phishing are the numerous amounts of undocumented reports. This limits the amount of data that can be analyzed to only what is available through the Anti-Phishing Working Group website's data source. Many victims don't know where to report such attacks or simply are unresponsive to the situation. Sometimes, people ignore (or just delete) "phishy" or suspicious-looking emails due to the fact that they are becoming more knowledgeable. Others are not aware that their information has been compromised; at least they have not awakened any suspicions, yet. The outcome of such limitation means that the study may not reflect results to be accurate to its entirety; still, results should depict a robust and trustworthy revelation.

Future research and development is needed to create an anti-phishing program that will detect if the website you are accessing is fraudulent. As for now, we should stick to our current resources and methods to combat such scams. By deploying anti-phishing techniques and tools into e-mail systems and browsers, individuals and organizations can help thwart the flow of valuable information into the hands of the 'black hats'. Becoming educated about phishing and the repercussions of stolen information makes people understand the importance of protection. The APWG has setup the APWG Public Education Initiative (PEI), located at http://education.apwg.org/, in the hopes of educating the public on phishing, its threats, and other related online theft methods. After all, this should have good effects on decreasing the number of phishing occurrences; otherwise, they will increase if absolutely no attempts are made to try to stop them.

8. REFERENCES

Akerman, N. (2003) "Protecting Yourself While Protecting Your Computer Data: Two Laws Make It More Important Than Ever." EDPACS: The EDP Audit, Control, and Security Newsletter, Vol. 30, pp.1–8.

Anderson, R. and T. Moore (2006) "The Economics of Information Security." Science, Vol. 314, pp. 610-13.

APWG (2004) "Phishing Attacks Trends Report, July 2004 Report." Anti-Phishing Working Group, available at http://www.antiphishing.org/reports/APWG_Phishing_Attack_Report-Jul2004.pdf

APWG (2008) "Phishing Activity Trends: Report for the Month of January 2008." Anti-Phishing Working Group, available at http://www.antiphishing.org/reports/apwg_report_jan_2008.pdf

Blau, John (2004, August 23) "Big German Banks Hit by Phishing Attacks." Computerworld, available at http://www.computerworld.com/softwaretopics/software/groupware/story/0,10801,95429,00,html

Bresz, F. P. (2004, July–August) "People—Often the Weakest Link in Security, but One of the Best Places to Start." Journal of Health Care Compliance, pp. 57–60.

Ding, Binxing & Ruifeng Li (2006) "Phishing and Anti-phishing." Department of Computer and Systems Sciences, Stockholm's University / Royal Institute of Technology, pp. 1-77.

Gartner (2007) "Gartner Survey Shows Phishing Attacks Escalated in 2007; More than $3 Billion Lost to These Attacks." Retrieved on July 7, 2009, from Gartner Newsroom, available at http://www.gartner.com/it/page.jsp?id=565125

Koong, K. S., L. C. Liu, S. Bai and B. Lin (2008) "Identity Theft in the USA: Evidence from 2002 to 2006." International Journal of Mobile Communications, Vol. 6, No. 2, pp.199–216.

McCune, J. C. (1998) "How Safe Is Your Data?" Management Review, Vol. 87, pp. 17–21.

Ollmann, Gunter (2008) "The Phishing Guide." Retrieved on June 30, 2009, from Anti-Phishing Working Group Website at http://www.antiphishing.org/

Pew (2007) "February 15 – March 7 2007 Tracking Survey." Pew Internet & American Life Project, 2007.

Sasse, M. A., S. Brostoff and D. Weirich (2004) "Transforming the Weakest Link - A Human/Computer interaction Approach to Usable and Effective Security." BT Technology Journal, Vol. 19, pp. 122—131.

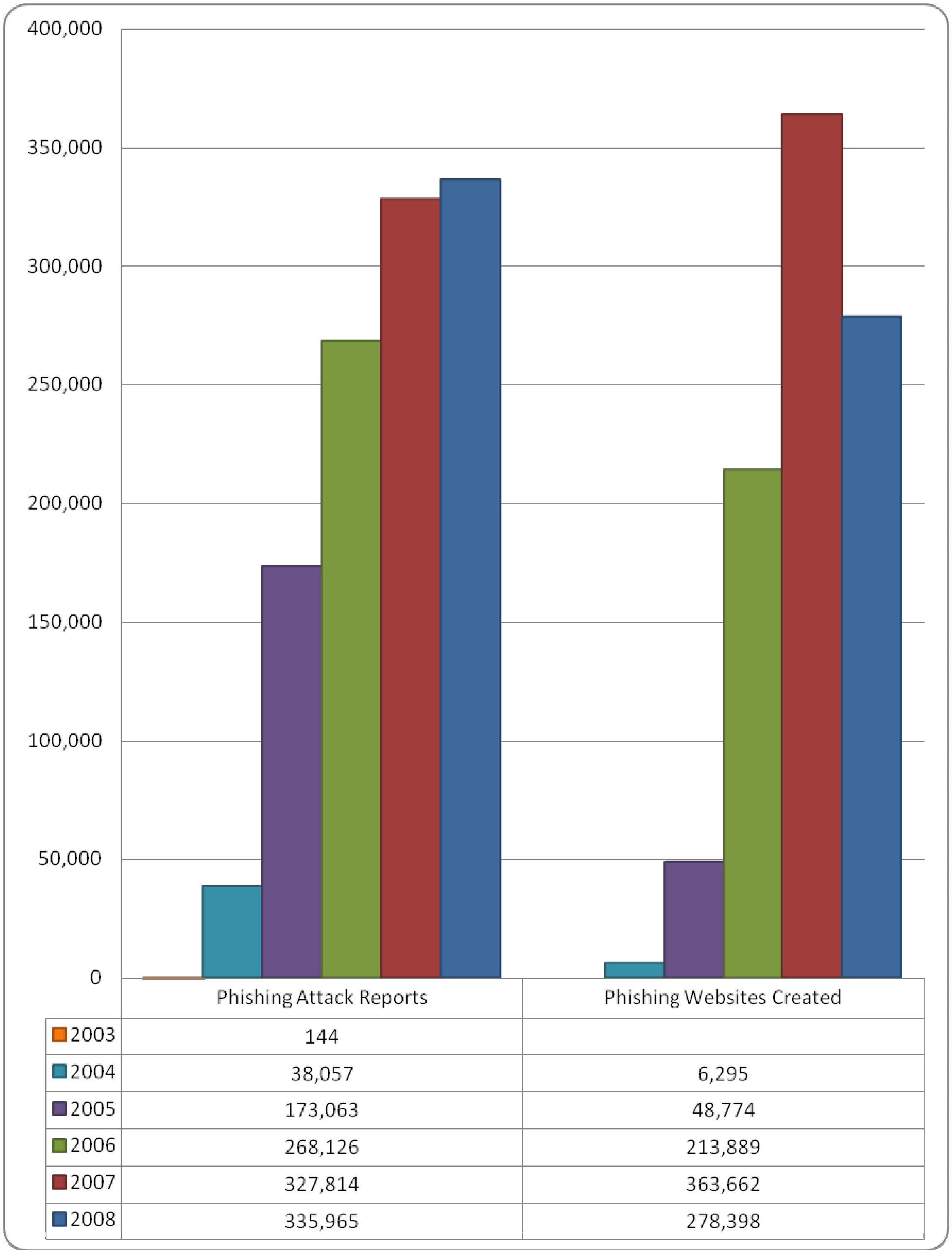
Sullins, L. (Spring, 2006) "Phishing for a Solution: Domestic and International Approaches to Decreasing Online Identity Theft." Emory International Law Review, Vol. 20, No. 1, pp. 397-433.

Vijayan, J. (2004, October 11) "Companies Fight Back against Phishing Scams." Computerworld, Vol. 38, No. 41, pp. 12-12.

Wang, Jingguo, Rui Chen, Rui, Tejaswini Herath, H. Raghav Rao (2009) "Visual E-mail Authentication and Identification Services: An investigation of the Effects on E-mail Use', Decision Support Systems, pp. 1-35.

Workman, M. (2008, February 15) "Wisecrackers: A Theory-Grounded Investigation of Phishing and Pretext Social Engineering Threats to Information Security." Journal of the American Society for Information Science & Technology, Vol. 59, No. 4, pp. 662-674.

Appendix A



	Phishing Attack Reports	Phishing Websites Created
🟧 2003	144	
🟦 2004	38,057	6,295
🟪 2005	173,063	48,774
🟩 2006	268,126	213,889
🟥 2007	327,814	363,662
🟦 2008	335,965	278,398

Figure 1: Phishing Attack Reports and New Websites identified (Yearly Trends)

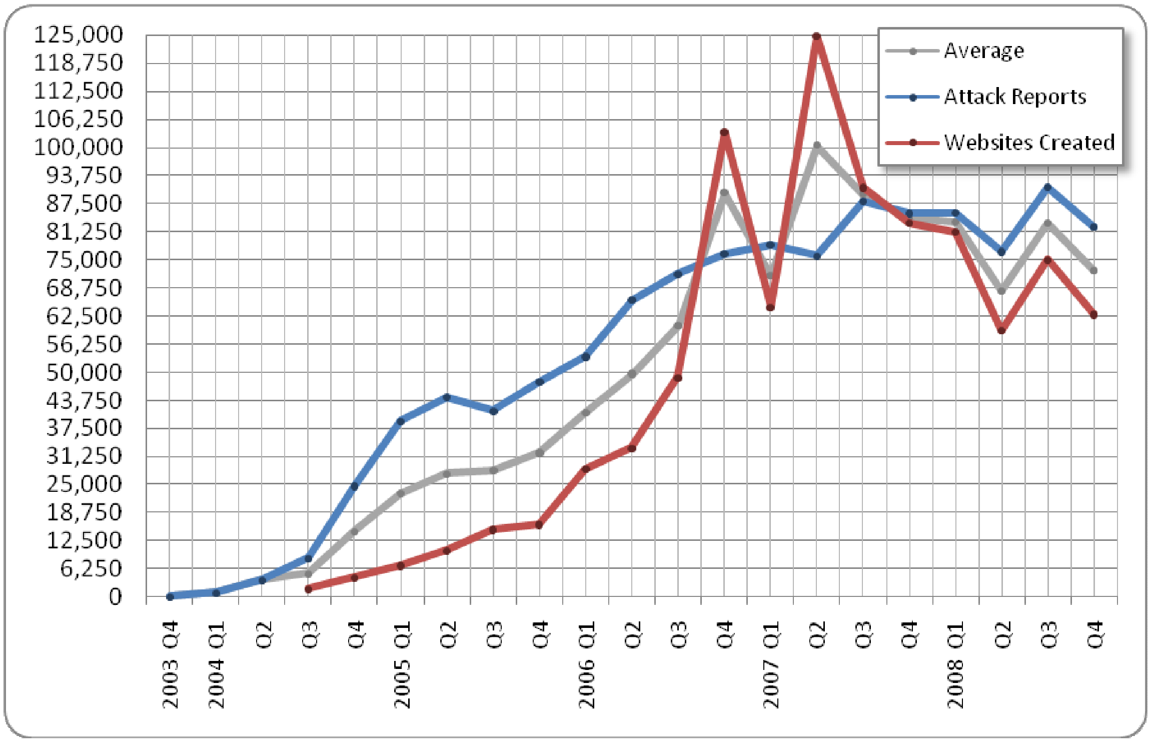
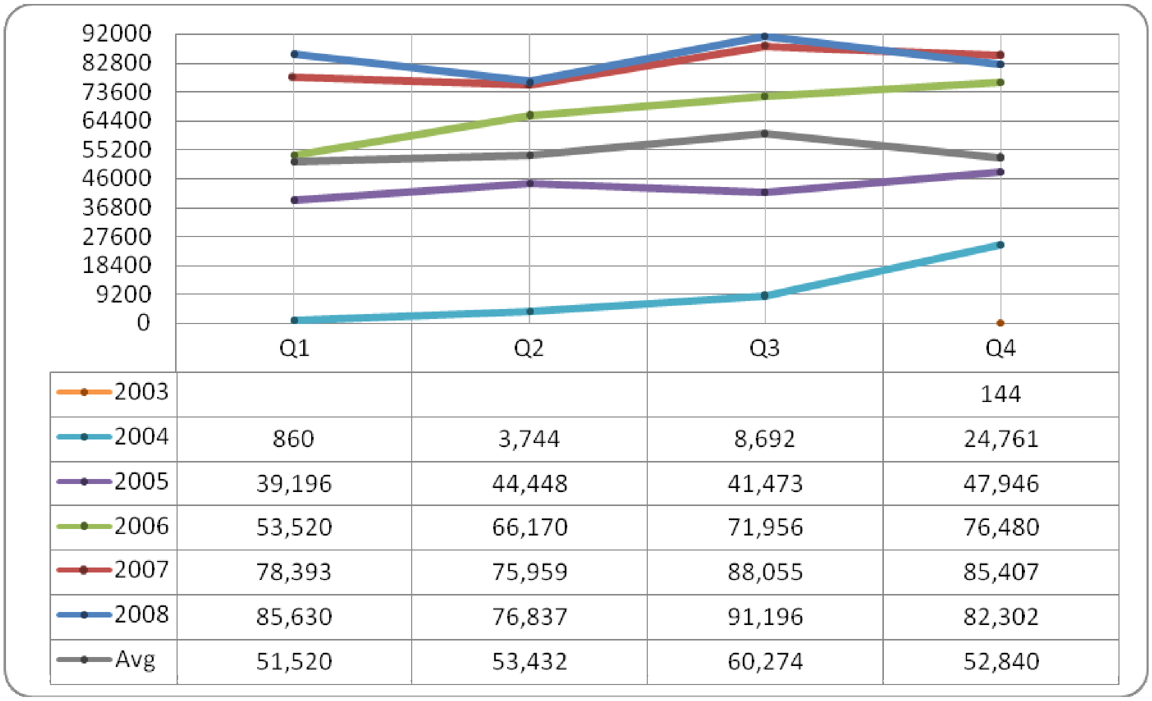
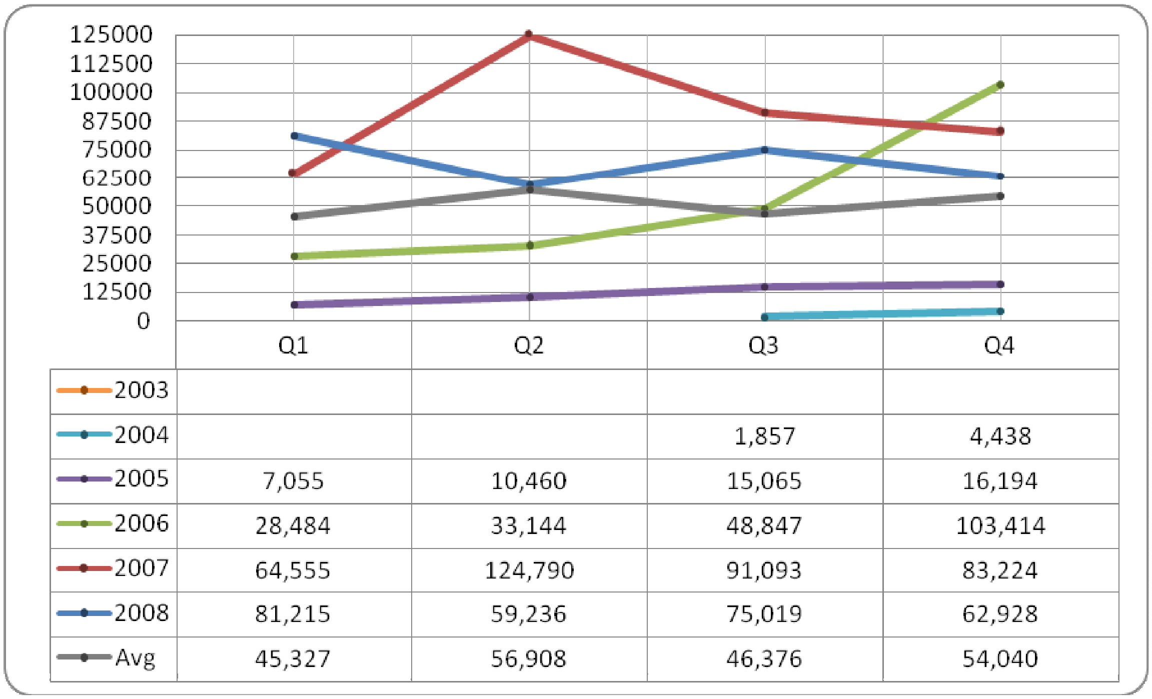


Figure 2: Phishing Attack Reports, Websites Identified, and Averages (Quarterly Trends)



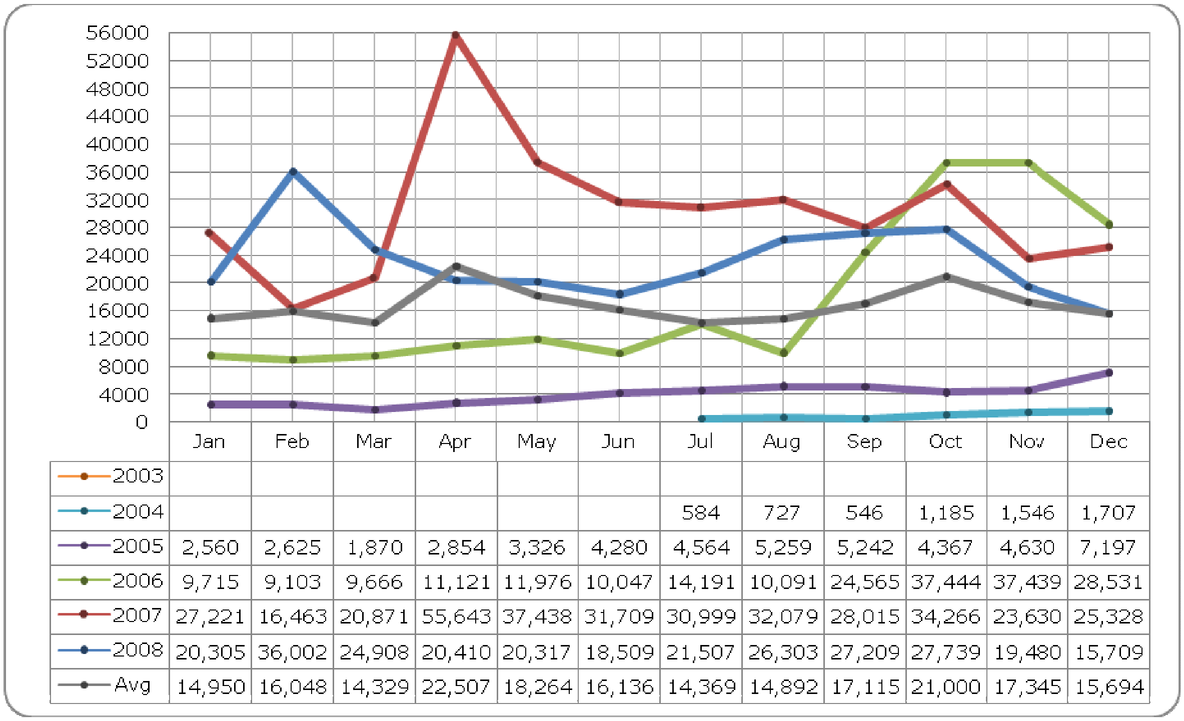
	Q1	Q2	Q3	Q4
2003				144
2004	860	3,744	8,692	24,761
2005	39,196	44,448	41,473	47,946
2006	53,520	66,170	71,956	76,480
2007	78,393	75,959	88,055	85,407
2008	85,630	76,837	91,196	82,302
Avg	51,520	53,432	60,274	52,840

Figure 3: Phishing Attack Reports (Quarterly Trends)



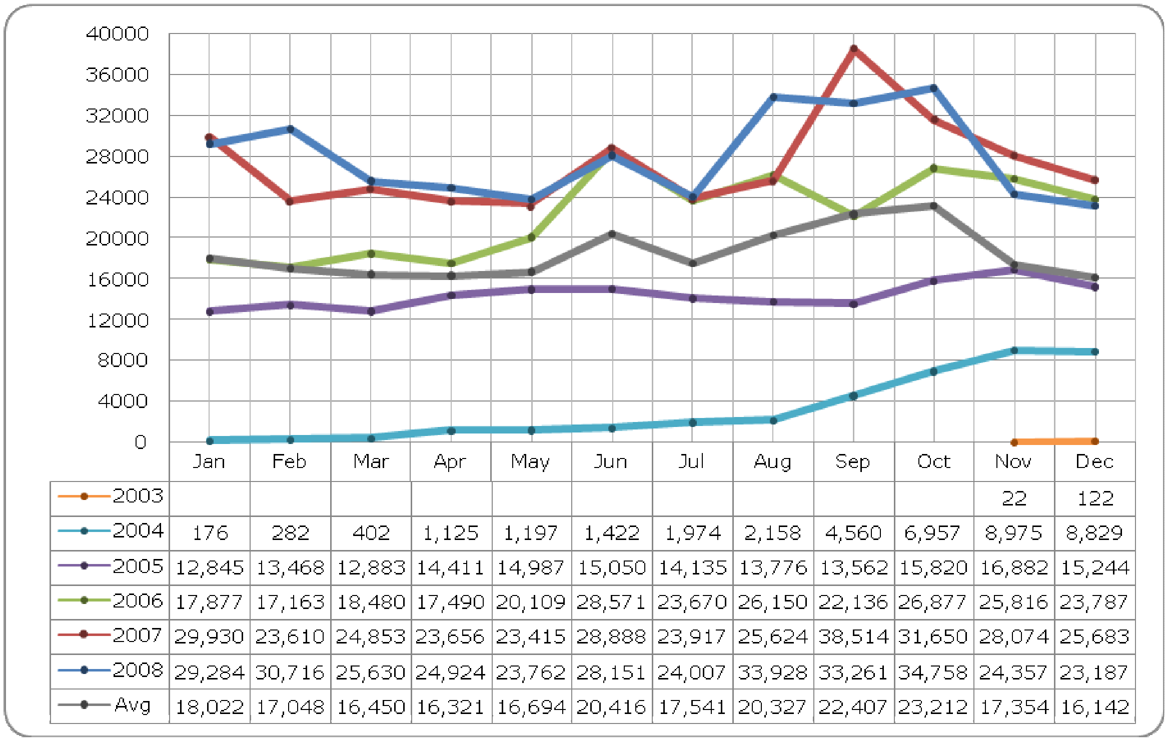
	Q1	Q2	Q3	Q4
2003				
2004			1,857	4,438
2005	7,055	10,460	15,065	16,194
2006	28,484	33,144	48,847	103,414
2007	64,555	124,790	91,093	83,224
2008	81,215	59,236	75,019	62,928
Avg	45,327	56,908	46,376	54,040

Figure 4: New Phishing Websites Identified (Quarterly Trends)



	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
2003												
2004							584	727	546	1,185	1,546	1,707
2005	2,560	2,625	1,870	2,854	3,326	4,280	4,564	5,259	5,242	4,367	4,630	7,197
2006	9,715	9,103	9,666	11,121	11,976	10,047	14,191	10,091	24,565	37,444	37,439	28,531
2007	27,221	16,463	20,871	55,643	37,438	31,709	30,999	32,079	28,015	34,266	23,630	25,328
2008	20,305	36,002	24,908	20,410	20,317	18,509	21,507	26,303	27,209	27,739	19,480	15,709
Avg	14,950	16,048	14,329	22,507	18,264	16,136	14,369	14,892	17,115	21,000	17,345	15,694

Figure 5: New Phishing Websites Identified (Monthly Trends)



	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
2003											22	122
2004	176	282	402	1,125	1,197	1,422	1,974	2,158	4,560	6,957	8,975	8,829
2005	12,845	13,468	12,883	14,411	14,987	15,050	14,135	13,776	13,562	15,820	16,882	15,244
2006	17,877	17,163	18,480	17,490	20,109	28,571	23,670	26,150	22,136	26,877	25,816	23,787
2007	29,930	23,610	24,853	23,656	23,415	28,888	23,917	25,624	38,514	31,650	28,074	25,683
2008	29,284	30,716	25,630	24,924	23,762	28,151	24,007	33,928	33,261	34,758	24,357	23,187
Avg	18,022	17,048	16,450	16,321	16,694	20,416	17,541	20,327	22,407	23,212	17,354	16,142

Figure 6: Phishing Attack Reports (Monthly Trends)