Understanding and Reversing the Profit Model of Spam (Position Paper)

Paul Judge, *Dmitri Alperovitch* and *Weilai Yang*

CipherTrust Inc.
4800 North Point Pkwy
Alpharetta, GA 30022 USA

Email: Firstname.Lastname@ciphertrust.com

March 6, 2005

1 Introduction

Spam, or unsolicited e-mail, has become a tremendous problem in recent years, evolving from being a minor nuisance as late as year 2000 to today comprising on average over 80% of all enterprise e-mail traffic and costing billions of dollars in lost productivity worldwide. It has become the parasite that infected the e-mail macrocosm and many now fear will lead to its destruction; the host becoming obsolete under the enormous harm done to it by the parasitic influence of this symbiotic relationship. While the news headlines continue to cite the increase in daily spam volumes, we suggest that the war against spam is reaching its final stages. Historically, before a war officially ends there is a key turning point which leads to victory. In the war on spam, we argue that we have reached that turning point.

Solutions have been proposed to "reverse the profit model" of spammers by charging for email[8]. We show that such a drastic overhaul of the e-mail system through the deployment of these solutions is not essential to accomplishing this goal. We begin by exploring the motivation of spammers and reviewing the detrimental effects of spam on the Internet. We then present a model of the profitability of spammers showing how various anti-spam initiatives affect it. We show that in fact, all the ingredients of an impending solution to the spam problem, such as near universal deployment of constantly technologically advancing anti-spam filtering technologies, improved user education, establishment of legal frameworks and aggressive enforcement of those frameworks, are already known and are at various stages of implementation.

2 Motivation and Methods of Spammers

In this section we explore the root cause behind the existence of spam and the continuous exponential increase in measured spamming activity worldwide. What causes one to become a spammer? Are these people looking to become a nuisance to society? Are they criminals and terrorists determined to inflict economic damage to the tune of billions of dollars on the entire developed world?

In reality, spamming is not a pastime or an ideological battle but an actual business process. Spamming is a commercial enterprise and, just like any other business, spammers one and only goal is to make a profit. This fact is useful in understanding the swift growth and spread of spam throughout the e-mail ecosystem, since knowing the enemy and understanding how they think and the factors that drive them is essential to winning this war.

As with any other business model, spammers must perform a few essential activities in order to create a profit:

1. *Find potential customers.* One cannot make money without a constant supply of people willing to send it to you. For spammers, finding these people involves obtaining lists of e-mail addresses of potential customers. Unfortunately, in the view of a spammer, a potential customer is anyone with an e-mail address residing on this planet. Since the distribution costs of spam are so low, there is little use in engaging in sophisticated market analysis to determine who is more likely to purchase spam-advertised products and designing campaigns to target only those people, as is frequently done with other traditional forms of direct marketing. As long as that transmission cost remains so insignificant, spammers will have no incentive to clean up and shrink their lists and cease to consider the entire world as their potential customer base.

2. *Offer a product or service to the potential customers.* Once a potential customer is found, one needs a product or service (or a scam, as is the case with phishing) to offer to that customer in exchange for payment. With that product or service in hand, the spammer can offer it to customers by broadcasting enticing e-mails to everyone on his compiled lists of addresses.

3. *Close the Deal*: Sell and deliver the product or service to some percentage of the recipients of the campaign. Once the spam has been sent, the spammer needs to only sit back and watch his orders accumulate over the next hours and days as a percentage of the people receiving the advertisement respond back and place their purchases.

The success of spam as a business is based on the very low cost of the first two activities, allowing even a minuscule response rate to still lead to substantial profit.

2.1 Task 1: Finding potential customers

Finding potential customers involves obtaining lists of email addresses and there are two main categories of methods that are used for this purpose: 1)purchase of lists and 2)address harvesting, the process of using tools to automatically collect email addresses on the Internet.

Address Harvesting:

- Spambots: A spambot is a program that crawls or traverses links on websites and newsgroups gathering email addresses. Pricing for such software ranges from free to high-end versions that cost up to $1000. Experiments conducted by the Federal Trade Commission showed that 86% of email addresses received unsolicited mail after being posted on websites and in newsgroups[4]. In some cases, email addresses posted in chat rooms received spam within 9 minutes of the posting. Today, the average time between the harvesting of the e-mail address by the spambot and spam starting to arrive at that address is 11 days, while the shortest recorded time is just 23 seconds[9].

- Viruses/Trojans: The practice of using viruses to harvest data from infected machines had first begun with the release of the ILOVEYOU e-mail virus in March 2000, which attempted to collect usernames and passwords from compromised machines and transfer them to the author of the virus. The outbreak of the Sobig virus in January 2003 marked the first detected collaboration between spammers and virus writers with the turning of the machines infected by this virus into spam-sending proxies very quickly after the initial infection. Nowadays most viruses and trojans include capability to mine hard drives of infected machines for e-mail addresses, as well as other personal information.

- Directory Harvesting Attacks: This is the process of connecting to a mail server to obtain information about addresses that it accepts mail for. There are two motivations for using this technique:
  1) Discovering new valid e-mail addresses and
  2) Verification of existing lists of addresses
  To obtain new addresses, the tool connects to the mail server and attempts to send email to many different random addresses created by traversing a dictionary of common names and combining it with the list of domains that are hosted by that mail server. The server typically acknowledges or denies the existence of each account, giving the tool the crucial feedback that it needs to determine the validity of

each address. Thus, this tool can be run against mail servers at any number of organizations to obtain new lists of potential “customers”.

The same technique is also used to verify purchased lists and scrub them of undeliverable addresses. The use of such tools, however, has waned over the last year, as the practice of using the spambots to crawl the public Internet and the use of viruses/trojans to do the same for the hard drives of computers all over the world has greatly eclipsed it in popularity.

Purchase Lists:

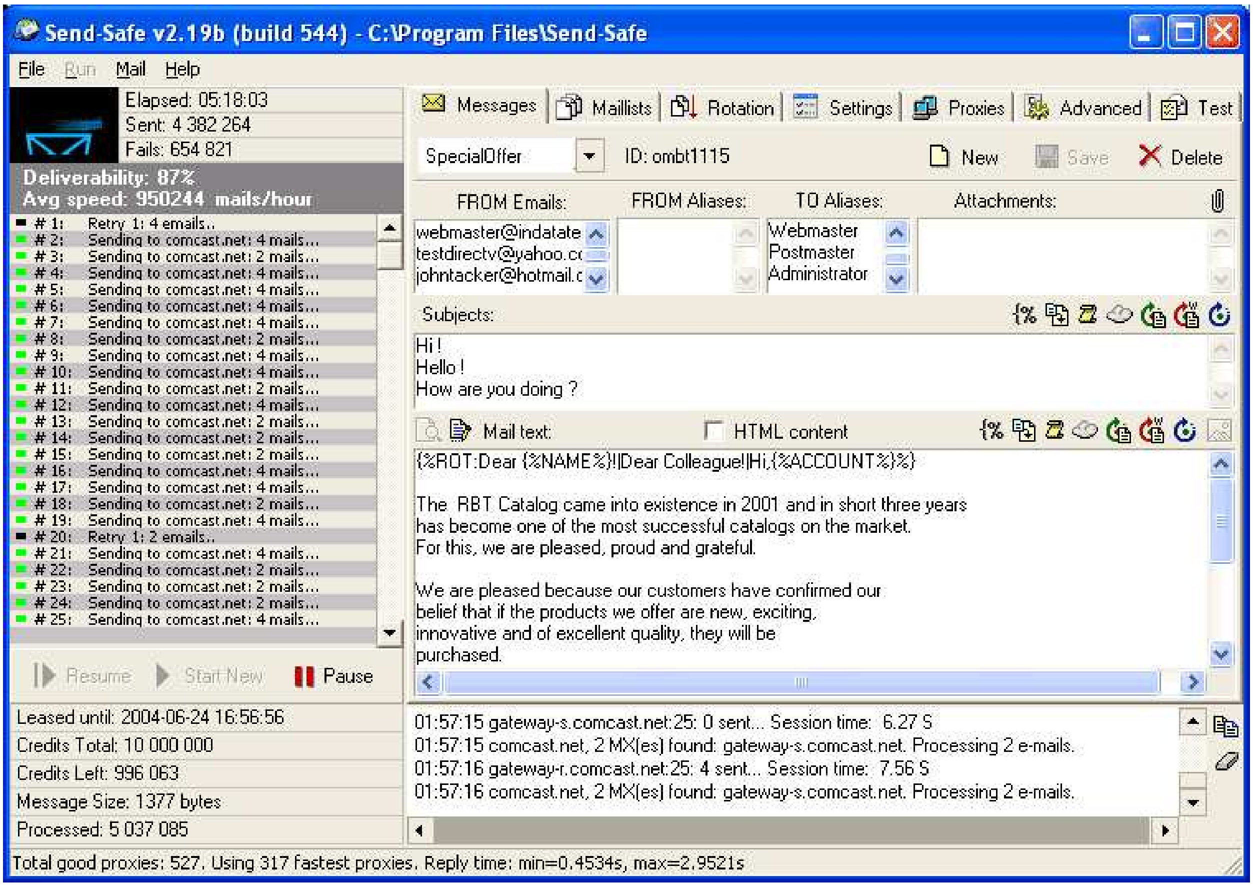
- Low-quality email address lists: These lists may cost as low as $19.95 for millions of addresses. They are cheap and widely available. They are frequently sold on underground spammer forums, as well as legitimate online auction sites. Usually, the addresses on these lists have been harvested or obtained using one of the illegitimate methods described above.
- Opt-in email address lists: These lists can cost up to $150 per thousand addresses or $150,000 for 1 million addresses. They are compiled from addresses of users that have at some point opted-in to hear about online product and service offers. These lists are typically used by legitimate bulk marketing companies. The cost of these lists is several orders of magnitude higher than the lists of the former category.
- Past customer address lists: Spammers frequently reuse the addresses of customers who had purchased products from them in the past, correctly predicting that having replied and purchased a product once through an unsolicited e-mail, these people are likely to come back and buy something else in the future. These lists are much smaller than those of the other categories and are significantly more expensive, since spammers are quite reticent to share such lucrative data with their competitors without adequate compensation.

Phishing is a form of online identity theft that uses spoofed emails designed to trick recipients into clicking on e-mail links that lead to fraudulent web sites and input their personal financial data and other sensitive information, such as credit card numbers, financial account user names and passwords, and social security numbers, which are typically resold and later used in financial and identity theft crimes. Phishers typically use the same methods to obtain lists of addresses of potential victims. They do, however, usually go through much more trouble to scrub the lists of any addresses that may belong to a government or an anti-spam organization. This is done out of necessity to stay under the radar for as long as possible to prolong the availability of the fraudulent website advertised in the e-mail, since law enforcement and good samaritans on the Internet are very proactive at notifying the usually oblivious host of the phishing site and getting it shut down within hours of the first deployment of the attack.

2.2 Task 2: Offering a Product

The process of spammers offering a product involves sending messages to the compiled list of email addresses. Spammers use sophisticated mass-mailing programs to send e-mails to the millions of addresses they have compiled on their lists. These programs are able to randomize the subject and sender identifying information, as well as virtually every part of the message in an attempt to bypass duplicate-detecting anti-spam tools and filtering techniques that are dependent on content-filtering. The price of such programs ranges from free to several hundred dollars. One of the most sophisticated of these packages, a program by the name of SendSafe, uses a subscription model, charging anywhere from $50 to $3000 per month from spammers wishing to use this software to send their mailings. Many of these more expensive software packages contain very impressive list management functions and are able to remember which users have opted-out of the mailings to allow the spammer, should she wish to do so, to stay in compliance with CAN-SPAM and other similar legislation. They are also able to track how many people have read the message and how many responded to it by inserting spying bugs in the HTML code of the message.

Another way that spammers reduce their costs and improve the chances of spam getting through detection filters is through the use of botnets. A botnet is a network of compromised machines that are running a trojan program called a 'bot' that connects to an Internet Relay Chat (IRC) channel that is operated by the spammer (or person supplying the spammer with botnets) and accepts commands from the channel operator, essentially ceding complete control of the compromised machine to the criminal. Early on in the infection process, the virus and trojan writers use machines in the botnet to scan other random computers on the Internet for known vulnerabilities and infect them with their new bot. Once the botnet expands to include thousands of machines and is deemed to be sufficiently large, the bots are commanded to turn on SOCKS proxies on the infected machines in the IRC channel and are, thus, instantly turned into an army of anonymous spam-sending proxies distributed all over the world. Spamware software, such as SendSafe, is able to make use of these proxies to optimize the speed and deliverability of a spam campaign.



Elapsed: 05:18:03
Sent: 4 382 264
Fails: 654 821
Deliverability: 87%
Avg speed: 950244 mails/hour

```
# 1:   Retry 1: 4 emails..
# 2:   Sending to comcast.net: 4 mails...
# 3:   Sending to comcast.net: 2 mails...
# 4:   Sending to comcast.net: 4 mails...
# 5:   Sending to comcast.net: 4 mails...
# 6:   Sending to comcast.net: 4 mails...
# 7:   Sending to comcast.net: 4 mails...
# 8:   Sending to comcast.net: 2 mails...
# 9:   Sending to comcast.net: 4 mails...
# 10:  Sending to comcast.net: 4 mails...
# 11:  Sending to comcast.net: 2 mails...
# 12:  Sending to comcast.net: 4 mails...
# 13:  Sending to comcast.net: 2 mails...
# 14:  Sending to comcast.net: 2 mails...
# 15:  Sending to comcast.net: 2 mails...
# 16:  Sending to comcast.net: 4 mails...
# 17:  Sending to comcast.net: 4 mails...
# 18:  Sending to comcast.net: 2 mails...
# 19:  Sending to comcast.net: 4 mails...
# 20:  Retry 1: 2 emails..
# 21:  Sending to comcast.net: 4 mails...
# 22:  Sending to comcast.net: 2 mails...
# 23:  Sending to comcast.net: 2 mails...
# 24:  Sending to comcast.net: 2 mails...
# 25:  Sending to comcast.net: 4 mails...
```

Leased until: 2004-06-24 16:56:56
Credits Total: 10 000 000
Credits Left: 996 063
Message Size: 1377 bytes
Processed: 5 037 085

SpecialOffer
ID: ombt1115

FROM Emails: webmaster@indatate, testdirectv@yahoo.cc, johntacker@hotmail.c
TO Aliases: Webmaster, Postmaster, Administrator

Subjects:
Hi !
Hello !
How are you doing ?

Mail text:
{%ROT:Dear {%NAME%}!|Dear Colleague!|Hi,{%ACCOUNT%}%}

The RBT Catalog came into existence in 2001 and in short three years
has become one of the most successful catalogs on the market.
For this, we are pleased, proud and grateful.

We are pleased because our customers have confirmed our
belief that if the products we offer are new, exciting,
innovative and of excellent quality, they will be
purchased.

```
01:57:15 gateway-s.comcast.net:25: 0 sent... Session time: 6.27 S
01:57:15 comcast.net, 2 MX(es) found: gateway-s.comcast.net. Processing 2 e-mails.
01:57:16 gateway-r.comcast.net:25: 4 sent... Session time: 7.56 S
01:57:16 comcast.net, 2 MX(es) found: gateway-s.comcast.net. Processing 2 e-mails.
```

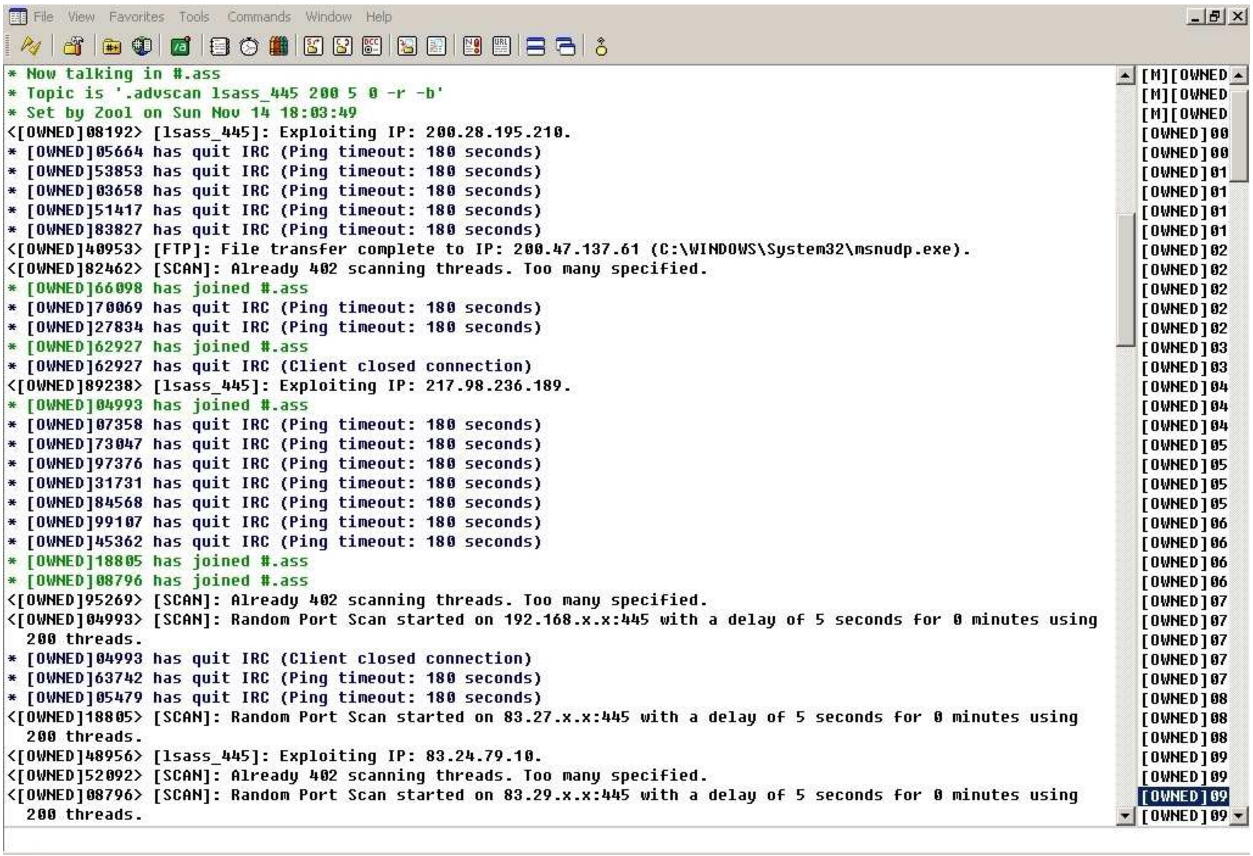
Total good proxies: 527. Using 317 fastest proxies. Reply time: min=0.4534s, max=2.9521s

Figure 1: Screenshot of SendSafe's use of proxies

Thus, with a dial-up or broadband connection and the click of a button, an average person is able to reach millions of potential customers, all while staying relatively anonymous from those that may wish to find her.

The low cost of placing product offers in front of many prospects makes spamming more attractive than traditional approaches, such as advertising, telemarketing, and direct mail. Table 1 compares the cost of offer placement among various common approaches. It shows that traditional approaches, such as direct mail and advertising, have per recipient costs ranging from $1.39 per recipient to less than $0.07. Nevertheless, the cost of spam is orders of magnitude lower costing only 1/20 of a cent per recipient.

In addition, due to the ignominious nature of the products and services that are frequently sold through spam advertisements, such as access to pornographic sites or sexually enhancing herbal medication, and the lack of knowledge on the part of many potential buyers about the real resale cost of these products and



```
* Now talking in #.ass
* Topic is '.advscan lsass_445 200 5 0 -r -b'
* Set by Zool on Sun Nov 14 18:03:49
<[OWNED]08192> [lsass_445]: Exploiting IP: 200.28.195.210.
* [OWNED]05664 has quit IRC (Ping timeout: 180 seconds)
* [OWNED]53853 has quit IRC (Ping timeout: 180 seconds)
* [OWNED]03658 has quit IRC (Ping timeout: 180 seconds)
* [OWNED]51417 has quit IRC (Ping timeout: 180 seconds)
* [OWNED]83827 has quit IRC (Ping timeout: 180 seconds)
<[OWNED]40953> [FTP]: File transfer complete to IP: 200.47.137.61 (C:\WINDOWS\System32\msnudp.exe).
<[OWNED]82462> [SCAN]: Already 402 scanning threads. Too many specified.
* [OWNED]66098 has joined #.ass
* [OWNED]70069 has quit IRC (Ping timeout: 180 seconds)
* [OWNED]27834 has quit IRC (Ping timeout: 180 seconds)
* [OWNED]62927 has joined #.ass
* [OWNED]62927 has quit IRC (Client closed connection)
<[OWNED]89238> [lsass_445]: Exploiting IP: 217.98.236.189.
* [OWNED]04993 has joined #.ass
* [OWNED]07358 has quit IRC (Ping timeout: 180 seconds)
* [OWNED]73047 has quit IRC (Ping timeout: 180 seconds)
* [OWNED]97376 has quit IRC (Ping timeout: 180 seconds)
* [OWNED]31731 has quit IRC (Ping timeout: 180 seconds)
* [OWNED]84568 has quit IRC (Ping timeout: 180 seconds)
* [OWNED]99107 has quit IRC (Ping timeout: 180 seconds)
* [OWNED]45362 has quit IRC (Ping timeout: 180 seconds)
* [OWNED]18805 has joined #.ass
* [OWNED]08796 has joined #.ass
<[OWNED]95269> [SCAN]: Already 402 scanning threads. Too many specified.
<[OWNED]04993> [SCAN]: Random Port Scan started on 192.168.x.x:445 with a delay of 5 seconds for 0 minutes using
  200 threads.
* [OWNED]04993 has quit IRC (Client closed connection)
* [OWNED]63742 has quit IRC (Ping timeout: 180 seconds)
* [OWNED]05479 has quit IRC (Ping timeout: 180 seconds)
<[OWNED]18805> [SCAN]: Random Port Scan started on 83.27.x.x:445 with a delay of 5 seconds for 0 minutes using
  200 threads.
<[OWNED]48956> [lsass_445]: Exploiting IP: 83.24.79.10.
<[OWNED]52092> [SCAN]: Already 402 scanning threads. Too many specified.
<[OWNED]08796> [SCAN]: Random Port Scan started on 83.29.x.x:445 with a delay of 5 seconds for 0 minutes using
  200 threads.
```

Figure 2: Screenshot of an IRC botnet that is enlarging itself through the use of the LSASS Windows vulnerability

services, the profit margins for the spammers can frequently be as high as 400%.

With phishing, the profit margins can be infinitely high as these criminals don't even need to spend money on purchasing inventory and providing a product or service to the customer. What phishers offer instead is a very sophisticated scam that consists of a falsely set-up and very realistic-looking website, typically using free hosting by residing on a hacked webserver or compromised zombie, which can fool victims into believing it to be a genuine website of their financial institution and that they have to provide their most personal and sensitive information on this site at once. After phishers receive such information, they frequently resell it to professional carders, who have great expertise at completely emptying out checking accounts via wire transfers to banks typically located in the Baltic countries, such as Latvia and Estonia, as well as other countries in Eastern Europe and former Soviet Union. Moreover, on top of this initial devastating financial loss, victims often find themselves victimized again months and years later through other forms of identity fraud committed by other criminals that have purchased their Social Security Number and other personal identifying information from the phisher who scammed them in the first place.

2.3 Task 3: Closing the Deal. Improving the Response Rate

Spam is so cost effective for the sender that even with very low response rates, the sender is still able to make obscene profits. Jeremy Jaynes, the recently convicted North Carolina-based spammer, had amassed a net wealth of $24 million in the course of just a couple of his years in his spamming profession[7]. These days, a good experienced spammer can easily clear a profit of $800,000 or more each month.

Table 1: The Cost Per Recipient of Various Marketing Tools

	Total Cost	**Number of Recipients**	**Cost Per Recipient**
Direct Mail	$9,700	7,000	$1.39
Telemarketing	$160	240	$0.66
Print - targeted	$7,500	100,000	$0.075
Print - general	$30,000	442,000	$0.067
Fax	$30	600	$0.05
Online Ads	$35	1,000	$0.035
Spam	$250	500,000	$0.0005

Based on the numbers from Table 1, direct mail is 2800 times more costly than sending unsolicited e-mail. Similarly, others have shown that direct mailers usually require a response rate of about 2% to make money on their campaigns[6]. Spammers, on the other hand, can break even with response rates as low as 0.001%. This shows that direct mail campaigns require a response rate at least 2000 times higher than spam. Thus, a spammer can send 500,000 messages and still be pleased and stay profitable with just 5 responses.

The people behind phishing e-mails are truly master scam artists. By hijacking the brands and website layouts of trusted and well-known banks, online retailers and credit card companies, phishers are able to convince up to 5% of recipients of their campaigns to respond and provide their personal and financial information to them. As a result of these scams, an increasing number of consumers are suffering from credit card fraud, identity theft and massive financial loss. The cause of such stratospheric response rates compared to those of regular spam is the social engineering aspect of phishing attacks. There are few people that would purchase a bottle of Viagra pills from an unsolicited e-mail that arrives into their mailbox with much of the words of the e-mail obfuscated beyond any readability to get passed content filters. Yet there are plenty who, without giving it a second thought, would fill out an online form with all their sensitive banking information in response to a branded and professionally written e-mail that appears to come from their bank and asks of them to confirm their identity online.

3 Detrimental effects of spam on the Internet

In this section we explore the size of the spam problem and its effects. Years ago, spam was considered a minor nuisance. More recently, it has evolved into a major issue impacting just about any company that does business online or any individual that has a public e-mail account. Today, the scale and effect of the spam epidemic leads us to suggest that spam is no longer simply an annoyance, but is a type of information security problem. The properties of the spam epidemic cause it to be classified along with well-known types of attacks such as denial-of-service, theft, and invasion of privacy.

3.1 Size of the Problem

Spam first appeared with a few rare occurrences in the early days of the Internet. The very first recorded unsolicited mailing occurred in 1978 on a system called ARPANET, the precursor to what is now known as e-mail[15]. Over time, more people were drawn to the ease of using spam to offer goods and services, legitimate or not. That has been coupled with what has been called "friendly fire", the influx of seemingly harmless emails that are circulated between friends such as chain letters, jokes, and online petitions. Consequently, the volume of spam has increased at an exponential rate. As of early 2004 about 65% of the mail received by an enterprise was comprised of spam, compared to about 80% today[2]. This ratio is typically even worse at universities and for personal email accounts. Some users no longer scan their mailbox searching for spam to delete, but instead they delete all messages and then scan for the rare good email among the junk. According to IDC market research firm, a typical organization of 1000 people will receive and circulate 2.1 million spam messages every year[5]. We have dealt directly with a number of organizations that have lost email service in the past because the email infrastructure was brought to its knees by an incoming spam flood.

Phishing is now exhibiting a similar exponential rate of increase as spam did in its early years. In January 2005, there were 12,845 new unique phishing email messages reported to the Anti-Phishing Working Group, which is a substantial increase of 42% over the unique reports for December, and represents an average monthly growth rate of 30% since July 2004[12].

3.2 Costs of the problem

There have been a number of approaches proposed to calculate the cost of the spam problem. We categorize the expenses into four categories: the cost of human resources, the technical resources costs, the intangible costs, as well as a separate direct loss caused by phishing crimes.

3.2.1 Costs of human resources

According to Gartner, a company with 10,000 employees loses $13 million worth of productivity annually because of "friendly fire" spam alone[13]. By 2004, the National Technology Readiness Survey conducted by the University of Maryland estimated that the productivity cost of spam had reached $21.5 billion[11]. And this does not account for technical costs incurred due to spam. The most common cost calculation used to estimate productivity loss is the total of wasted wages due to employees sorting through spam in their inbox. This is typically calculated by the amount of mail received by the organization, amount of time dealing with each spam message, and the average employee hourly rate.

3.2.2 Technical costs

As spam messages now account for a lion's share of total email traffic and the average spam message is 4 kilobytes in size, the bandwidth usage of spam is significant. A report by the European Commission in 2000 estimated the global bandwidth costs of spam at $9 billion annually[14]. Additionally, once those messages are received, they are processed by mail gateways and mail servers utilizing processing resources. Then, these messages typically reside on the mail storage devices on the servers and/or the desktops.

The processing and storage costs of spam are significant. Many organizations have had to double spending on their email infrastructure to cope with the increased amount of messages. This includes the cost of new hardware, operating systems, and mail server software licenses. Furthermore, many industries, such as the healthcare and financial sectors, have regulations that require that all email communications be archived for a certain period of time, up to seven years in some cases. So, not only does the spam consume current storage capacity, but also once received, organizations must pay for storage of spam for several years to come.

We have discussed the costs in the context of organizations such as universities, enterprises, and ISPs. These costs are almost always passed on to end-users. For example, some major ISPs estimate that 10-30% of subscriber fees are used towards dealing with spam.

3.2.3 Intangibles

A less common cost that has been proposed is the effect that spam has on "pollution" of the work environment. Over 25% of all spam attacks today are adult-oriented and this number doubled in the last year. This influx of explicit pornographic e-mails causes many parents to also feel a loss of parental control over their children. This has become another reason that children are no longer allowed to use the Internet unsupervised in many households.

Some have suggested that spam has also caused a change in the way people fundamentally use email. For instance, it is common these days for people to have multiple email addresses, usually revealing certain private ones only to well-known correspondents. People are also reluctant to make their email address publicly known and can reduce their participation in newsgroups and message boards in the fear of the address being harvested by spammers. In addition, some are hesitant to use web pages and auction sites that require registration and revealing of one's email address. Therefore, it can be argued that the spam problem is disrupting the positive use of the Internet and has a profound effect that goes well beyond e-mail.

Finally, there is an issue of false positives, or mail that is accidentally deleted by the spam filter or the recipient due to its incorrect identification as spam. It is clear that the cost of missing an important email message is substantial and can have severe economical impact on a business, even though the precise costs are difficult to calculate, as businesses are often hesitant to reveal such incidents publicly.

3.2.4 Phishing Loss

Along with the human resource and technical losses described above, phishing creates an extra direct financial loss for the victim of the attack through the emptying out of her credit cards and bank accounts. In addition, the loss of time that is required to put one's life together and stress endured by an individual after being victimized by this crime can dwarf all the indirect costs related to dealing with regular spam.

While the industry consensus is that phishing attacks are on the upswing, there are disagreements about the severity of the damage. TRUSTe, an online privacy nonprofit organization, arrived at the $500 million figure for the total direct phishing fraud damage to date by applying the survey's $115 average loss per victim to the general Internet population[10]. A little more than 2 percent of all the people surveyed said they had lost money, in most cases within two weeks of being phished. A Gartner study calculated that a whopping $2.4 billion was lost to phishing fraud in the last 12 month-period[13].

4 The Profitability Model of Spamming

Many agree that the route to solving the spam problem is to reverse the spammer profit model. Rather than directly charging for emails, which is a very controversial and likely impractical solution, we show that there are a number of factors that can affect the profitability of spammers. We suggest that the spammer profit can be measured with the following formula:

Spammer Profit = [(Chance of Getting Away with Crime) * (Number of Delivered Messages) * (Response Rate) * (Profit per Item)] - [(Number of Sent Messages) * (Cost of Address Amortized over Use + Cost to Send)] - [(Risk of Getting Caught) * (Cost of Punishment)]

This formula can be further generalized into:

$P = [(1\text{-}p) * (N_d) * (R_r) * (P_i)] - [(N_s) * (C_a + C_s)] - [(p) * (C_p)]$

where

P - Spammer Profit
p - Probability of getting caught sending spam
N_d - Number of sent messages that are delivered to the intended recipient
$N_d = N_s * (AS_d) * (AS_b)$
AS_d - Anti-spam deployment rate
AS_b - Anti-spam block rate
R_r - Response rate
P_i - Profit per item
N_s - Number of sent messages
C_a - Cost of acquiring each address amortized over the useful life of that address
C_s - Cost to send each message
C_p - Cost of punishment

The first way to affect the profitability is to reduce the amount of spam messages that reach inboxes. The two ways of controlling this are through maximizing the following terms of the equation: 1)effectiveness rate of the anti-spam technology (AS_b) and 2)deployment rate of anti-spam technologies (AS_d). With the average effectiveness of most industry tools already at 90% and market leading solutions reaching rates of 99%, as well as the now virtually universal deployment of at least some sort of anti-spam technology, both of these terms are close to reaching their global maximum.

Even when some number of spam messages do reach inboxes, profitability can be limited by reducing the response rate (R_r). User education and awareness are key to making this a reality. User education can also help in reducing the supply of fresh e-mail addresses to spammers and phishers alike by concealing them from public view or obfuscating them to avoid detection by Web crawlers and spambots. Thus, we can substantially increase the costs involved in obtaining email addresses (C_a). Much still needs to be done in this area, but with constant and overwhelming media coverage of this issue, things are likely to get better in time.

Finally, anti-spam legislation not only introduces spammer overhead (C_p) in the form of the expenses of litigation but also provides a strong deterrence (p). The relationship between technology and legislation is a familiar one. For example, consider the problem of computer intrusions: the technology available to protect resources, such as firewalls and intrusion-detection systems, is supported by legislation, allowing the prosecution of those who are determined to circumvent the technology. The same relationship is seen in something as basic as property theft, with burglar alarms and door locks providing protection, as well as being supported by appropriate property laws that provide a strong deterrence.

5 Transformations in spammer activities aimed at maximizing profitability

As the anti-spam community works on these fronts to reduce the overall spam profit model, spammers are aggressively responding to squeeze the last drops of profitability out of this business. Here we show how many of the spammer trends that have received much attention over the last year are not arbitrary, but are directly focused at variables in the profit equation.

There are three classes of mutation techniques that we will discuss:
1) Spam volume increase
2) Filter evasion and
3) Gravitation toward phishing scams

The most natural first reaction against improved filtering techniques was for spammers to simply send more messages. If they knew that some portion of their messages was being blocked by spam filters, they could simply aim to send that much more to compensate for the loss. The basic economics of email allows this to be accomplished with very little if any incremental cost to the spammer. As discussed earlier, the nature of email causes the majority of the cost to be borne by the recipient rather than the sender as is the case with most equivalent systems in the physical realm. Additionally, since spammers are now almost always using illegally acquired resources, such as botnets and compromised relays, there is often no actual corresponding increase in cost with increasing mail volumes.

The second type of mutation is filter evasion, a wide array of techniques used by spammers to avoid certain spam detection approaches. For example, as a reaction to blacklisting, spammers began to zombie botnets to have an enormous and constantly refreshed set of mail-sending IPs. As a reaction to content filtering techniques, spammers started to misspell and obfuscate words. The anti-spammers countered by deploying more intelligent adaptive learning techniques that learn the difference between spam text and good text. Spammers countered by adding large amounts of good text to their messages to attempt to weigh the filters in their favor. As spam signatures became popular, spammer began using message randomization techniques to create unique copies of each message in a campaign. Thus, spammers are forced to constantly innovate and react to constantly improving filtering techniques in order to maintain their current profit margins.

Finally, spam is slowly but surely being rendered less and less profitable by a number of technological, societal and legal factors. Improved filtering techniques and better user education have led to a reduction in response rates to spam messages. The enactment of new laws, such as CAN-SPAM in United States[3] and similar initiatives in other countries around the world, in conjunction with a significant increase in law-enforcement activity and prosecutions in this area, have resulted in a sharp rise in the probability of getting caught and the cost of punishment. In combination, these factors have led many spammers to migrate to more overt illegal activities, such as phishing.

In the past year, there has been a sharp rise in the number of phishing scams, which can cause considerably more consequential and longer-term damage than ordinary spam. The term ”phishing” originated to describe the attempts by crackers to steal AOL accounts from users by getting them to give up account information over instant messaging. Today, however, e-mail is the preferred communication medium for most phishing attacks [1]. Such e-mails purport to come from legitimate and well-known financial organizations, such as CitiBank and PayPal, or online auction sites, such as eBay. The typical theme that is used in the vast majority of phishing scams is the alerting of the user that their account is about to be suspended unless they immediately click on the link provided within the message and update their account information on that site, which is typically hosted on a hacked server and has no connection to the organization it pretends to affiliate with. The statistics given by Anti-Phishing Working Group show that the average monthly growth rate in phishing sites since July 2004 through January 2005 is 28%[12]. The number of brands hijacked by phishers also continues to increase at an exponential rate. Figure 3 shows the sharp rise in the number of reported phishing sites in recent months.

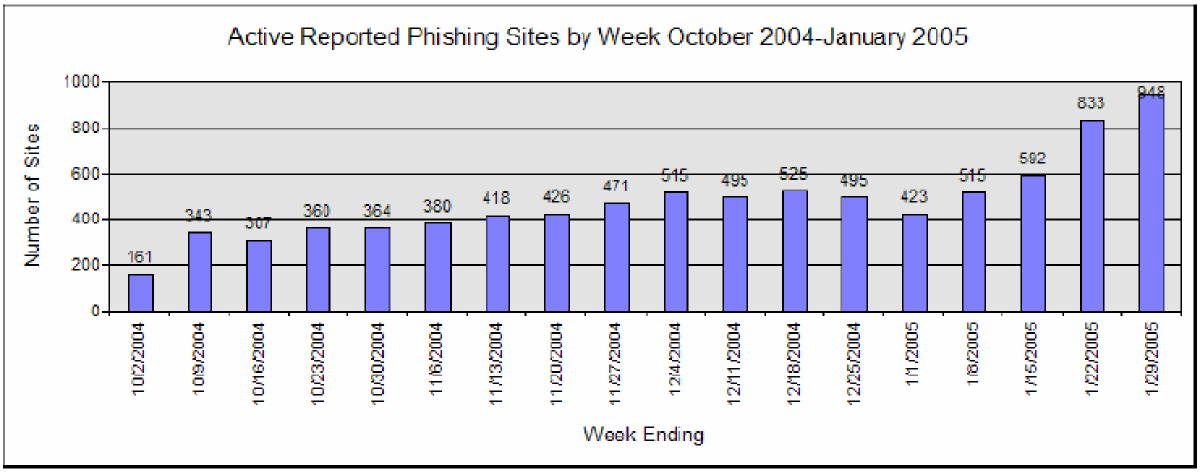


Figure 3: Active Reported Phishing Sites

The rational for spammers to migrate to phishing is simple - if the risks of capture are increasing and the profits keep dwindling, one needs to do something drastically different to improve the profit model or drop out of this business entirely. Phishing offers a perfect opportunity to do just that. With astonishingly high response rates, infinitely high profit margins, as well as significantly better penetration through spam filters due to the very legitimate-looking format and text of phishing messages, phishing is appearing significantly more lucrative to more and more spammers. On top of that, the risk of capture due to phishing does not appear to be much higher than that of spam just yet, making the entire phishing profit model appear much more appealing.

6 Conclusion

We have witnessed the maturing of anti-spam technology over the last few years as well as the increase in deployment rates. Consequently, a significant amount of mailboxes today are protected from spam messages. This has caused spammers to resort to more extreme and desperate measures. We are seeing the final mass efforts of spammers to find ways to maintain profitability in the face of the new world of email. As the technology is focusing more on identifying legitimate email and reputation of senders, it is becoming more difficult for spammers to hide. As CAN-SPAM was enacted at the beginning of 2004 and is starting to be enforced, spammers face a very real risk of retribution. Many have given up and decided that it is no longer worth it to attempt to make a living based on these activities. For example, recently we had conversations with a gentleman that was formerly one of the top ten spammers in the world. He made it clear that even though he used to make millions of dollars a month based on a spamming business, the new pressures of technology and legislation created an environment that does not allow him to maintain such a business. He has now walked away from the world of spamming. Certainly this is only one data point, but it is an indication of the turning point in the battle against spam.

References

[1] Aaron Emigh, Radix Partners Anti-Phishing Technology, Report in conjunction with United States Secret Service San Francisco Electonic Crimes Task Force, 2004.

[2] CipherTrust, Inc, Spam Statistics. http://www.ciphertrust.com/resources/statistics/index.php

[3] Controlling the Assault of Non-Solicited Pornography and Marketing Act of 2003, Public Law, 108th US Congress

[4] Email Address Harvesting: How Spammers Reap What You Know, FTC Consumer Alert, November 2002.

[5] IDC. http://www.idc.com

[6] Jack Ferreri, Entrepreneur Magazine's Knock-Out Marketing: Powerful Strategies to Punch Up Your Sales, *Entrepreneur Press*, June 1999

[7] Jeremy Jaynes Trial Documents

[8] Kraut, Robert E., Sunder, Shyam NMI, Telang, Rahul and Morris, James H., *Pricing Electronic Mail to Solve the Problem of Spam*

[9] Matthew Prince, Project Honeypot, *The Third Spam Conference*, MIT Jan 2005.

[10] Michael Pastore, Phishing is Up and It Has Consumers Down, *Inside ID*, September 2004

[11] National Technology Readiness Survey 2004. R.H. Smith School of Business, University of Maryland

[12] Phishing Activity Trends Report, Anti-Phishing Working Group, January 2005

[13] Avivah Litan, *Phishing Attack Victims Likely Targets for Identity Theft*, Gartner FirstTake FT-22-8873, May 2004.

[14] Serge Gauthronet and Etienne Drouard, Unsolicited Commercial Communications and Data Protection, January 2001

[15] Templeton, Brad, Reflections on the 25th Anniversary of Spam, March 2004