SPAM: Dealing with unwanted e-mail

by

Steven E. Fitch MBA

www.stevenefitch.com

This document is for information purposes only!

Spam email is taking over the Internet as spammers exploit its openness and free-sharing nature. Spammers make a living through the harvesting of millions of email addresses, and then using them as targets for fraudulent, misleading, or pornographic messages. Spam has become the Internet's version of the Black Plague, and it will cost businesses in excess of $10 billion in 2003 in lost bandwidth and compromised IT resources. In short, I view the battle against spam as one that must include anti-spam technology, federal legislation, and user education on how to avoid becoming a target of spam.

The Origins of Spam

Spam e-mails are as old as the Internet, and began 25 years ago on the Internet's predecessor, ARPNET. With the Internet boom in the mid-1990's technologists dubbed the e-mails, "*spam*," not to be confused with Hormel Foods' canned meat product, SPAM. It is said that the inspiration for using the word "*spam*" is said to have originated from a Monty Python skit about a restaurant that served only processed meat, and a group of Vikings singing, "*spam, spam, spam, spam*," drowning out other conversations in the restaurant, much as spam e-mail tends to drown out legitimate e-mail communication.

Key Points

- Spam is fast becoming a #1 pain point for enterprises worldwide.
- As of March 2003, 45% of all email is now spam.
- Anti-spam technology has become a crucial component of Internet security.

The anti-spam market is fragmented with over 50 vendors, and is expected to see continued consolidation as the larger Internet security companies look to include anti-spam technology into their repertoire.

Network Associates and Symantec are the *two* companies in our Internet security coverage universe with recently introduced anti-spam solutions for businesses and home users. According to the NPD Group, Network Associates has captured 90% off the consumer market with its SpamKiller product, while Symantec offers an anti-spam solution within its Norton Internet Security software suite.

It is expected that the anti-spam market to boast the strongest growth within the Internet security arena over the next five years. According to estimates and those of Ferris Research, they believe the market could reach $850 million by 2008.

Like other types of corporations, ISPs must deal with lost storage space, diminishing bandwidth, and special software and IT staff requirements that together result in higher costs. Gartner Research estimates that an ISP with one million users spends approximately $7 million per year defending against spam.

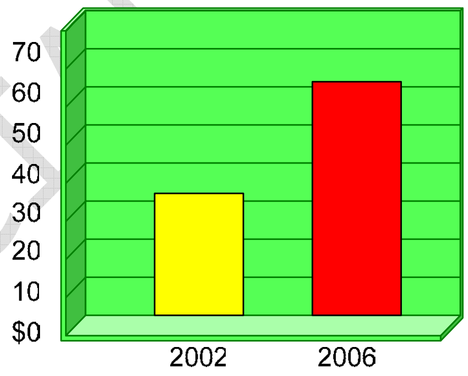
Early this year (2003), *three* of the major ISPs – ***AOL, MSN, and Yahoo!*** have made a push to stopping the proliferation of spam. In addition to enhancing their anti-spam technology, these *three* will continue to lobby for tougher federal legislation that seeks to punish spammers.

On April 28, 2003, ***AOL, Microsoft, and Yahoo!*** entered into a partnership to combat spam. Together they will press for legal action against spammers. Each of these ISPs has recently instituted some defenses against spam.

- **AOL** – As the largest ISP, AOL claims to block approximately 1 billion spam emails per day. The company employs numerous types of email filters and mail controls. It also responds to spam complaints and attempts to uncover the original senders.
- **MSN** – Microsoft uses the Brightmail anti-spam product for its Hotmail accounts. Microsoft has also introduced an anti-spam platform for its Exchange Server 2003, which will accommodate various third party anti-spam technologies. Microsoft asserts that the eventual solution to spam goes well beyond current technology developments and must incorporate effective legislation and user awareness.
- **Yahoo**! – For its email users, Yahoo! Has developed its own anti-spam technology called *SpamGuard*. Bulk email and junk email can be automatically isolated in a dedicated folder that the user can simply empty without opening.

Spam Statistics

Currently, approximately 31 billion emails are sent each day around the world; if spam continues to multiply as it has over the last two years, this number could double by 2006. If this happens, the efficiency of the Internet as we know it will be severely compromised.



The Problems Caused by Spam

Spam is wreaking havoc on corporate IT infrastructures around the world. Below are several of the major problems corporations face with the flooding of spam emails.

- **Sapped IT Resources** – Spam eats up an organization's IT resources by taking valuable storage space and occupying the time of an IT staff that should be concerned with the core competencies of the network.
- **Lost Bandwidth** – An average employee deals with a least 1500 spam emails each year. Often, these messages contain video clips and other multi media that spa the bandwidth of an organization. Bandwidth can become expensive, if a corporation must continually increase its supply.
- **Virus Potential** – Spam emails are the primary avenue for email borne viruses to infiltrate a corporation's network. Viruses can shut down an entire network within minutes, leading to a myriad of problems from lost customers to corrupted files.
- **Legal Liability** – With the majority of spam containing pornographic and other potentially insulting material, an organization faces the responsibility of dealing with an employee or customer who resents receiving an email containing this type of content.
- **Employee Productivity** – The process of opening and reading spam emails takes away from an employee's ability to work. With the flood of junk email coming through every day, the time wasted is significant.

Spamming Techniques

Like other Internet vandals and thieves, spammers have become extremely adept at finding ways to cultivate immense lists of legitimate email address. These lists are continually updated and sold within the spamming community. Spammers use such techniques as "*dictionary attacks*," which involve special software that generates millions of emails by combining simple names with integers such as JDOE@anywhere.com.

"*Directory harvest attacks*" are another popular technique of cultivating addresses. Spammers have developed software that scans a company's list of IP addresses and "*harvest*" the valid ones. Once the spammer has built a list of addresses, he utilizes numerous mail servers connected to the Internet to send fraudulent requests to the targeted corporation's mail server. This method is normally a very efficient and quick way of netting thousands of email addresses in a few minutes. Many anti-spam solutions, as well as firewalls, are unable to detect harvest attacks because these attacks simply exploit the free-sharing nature of email.

Another common technique involves the use of a "*beacons*." A beacon is used by a spammer to identify who is actually opening the spam messages. A spam message often contains a graphic image and if the recipient opens the message, a request is sent to a web server that contains the image to be downloaded into the spam message. This request contains the recipient's email address, which is then saved into a database that the spammer can use for future mailings.

In many cases, spammers are able to hide or obscure the originating address for a spam mailing through the use of special software. Spammers can also hide their URLs in graphics or even within the body of a message.

Spammers have also been known to take over a company's mail server that has been left open with no security. The spammer will then use the server to send the bulk emails.

The Cost of Spam

Businesses are suffering more and more from the outbreak of spam on a daily basis. Loss of costly bandwidth, high virus potential, legal liability and crowded email servers are just a few of the problems encountered by virtually every business with an email system. Enterprises are frustrated with the technology currently available, because while it may block a decent amount of spam, it also blocks legitimate emails crucial to the enterprise's business. In response, anti-spam vendors are developing solutions that are flexible and customizable, so that a customer can select the messages they wish to receive.

Don't believe that spam and other junk e-mail in your employees' e-mail inboxes impact your corporate bottom line? ***Think again***

Number of employees with e-mail	1,000
Number of workdays per year per employee	230
Average hourly salary per employee	$25
Average number of junk e-mails per day per employee	50
Average number of seconds wasted per junk e-mail	5

Total Corporate cost of Spam		**Spam cost per employee**	
Yearly	$399,305	***Yearly***	$399
Daily	$1,736	***Daily***	$1.74
Time	1,056 days	***Time***	25 hours per employee per year

Legal Action Against Spam

What is the best course of action? Anti-spam legislation has proved to be a difficult path to fight spam. First, by its nature, the Internet is a tough place to enforce laws, as users can operate anonymously and easily hide the origin of their correspondence. A hacker can easily move to a jurisdiction that down not enforce laws put in place to control Internet abuse. A second problem is that the definition of spam is open to interpretation.

Legitimate businesses that are selling products through mass email campaigns claim that to include their actions in the definition of spam that would impede on the right to free speech. These businesses would argue that spam is bulk email that contains things like potentially harmful or insulting content, fraudulent claims, or unwanted advertising. Another important part of this definition is that any email that is sent deceptively without a proper return address or is sent by hacking into another computer or server constitutes spam.

Others would say that any bulk email that does not allow the user to "*opt out*," or choose whether to receive it is spam. The latter definition is far more broad and encompassing regarding the bulk email that floats through the Internet.

On the state level, at least *28* have passed anti-spam laws. While effectiveness of these laws is questionable, anti-spam legislation is starting to gain some momentum. Many of these laws have been hard to enforce, and the penalties to date have been quite soft.

The most recent and toughest spam law to be passed resides in Virginia. With several of the largest ISPs, including AOL, located in Virginia, it becomes an important state in the fight against spam. The law, enacted on April 29, 2003, prohibits the sending of fraudulent emails that contain deceptive subjects or false addresses or emails that have been sent illegally by hacking into another computer. If a spammer is found guilty of sending at least 10,000 of these emails in one day, he or she would be sentenced to jail time of one to five years, and a forfeiture of any profits or assets associated with his / her spamming business.

While action at the state level is positive, many experts maintain that for legislation to stand a chance of success there must be a unified federal system of laws governing spam. There are several important bills currently in the Senate and the House that focus on the reduction of spam:

- **CAN-SPAM ACT** – Sponsored by Senator Conrad Burns (R-Mt.) and Senator Ron Wyden (D-Or.), the CAN-SPAM Act probably stands the best chance of passage into Federal law. The Act strives to separate fraudulent emails from legitimate ones, by requiring the emails contain subject lines revealing their contents and real return addresses. These emails must also have an "*opt out*" option, which users can click if they do not want to receive further emails from the given source. The penalty for spammers would be up to one year in jail and a fine of $10 per email capped at $500,000.
- **REDUCE Spam Act** – Sponsored by Representative Zoe Lofgren (D-Ca.), the REDUCE Spam Act seeks to offer a bounty to people who track down illegitimate spammers. The penalty for those caught is up to one year in jail and a fine of $10 per email.
- **Ban on Unsolicited Bulk Electronic Mail Act** – Senator Bill Nelson (D-Fl.) sponsors this bill, which seeks to outlaw emails that contain false subject lines or "*from*" addresses. It also serves to prohibit spammers from probing the Internet for addresses to use in a mass email campaign. The penalty could result in up to five years in jail.
- **Reduce in Distribution of Spam Act** – Representative Richard Burr (R-Nc.), Rep. Billy Tauzin (R-La.), and Rep. James Sensen Brenner (R-Wi.) sponsor this bill, which requires an "*opt-out*" option for consumers to remove themselves from a bulk email list. The bill also bans type of fraudulent email and requires valid return addresses. Emails that advertise something must include "*ADV*" in the subject line as well. The bill allows ISPs to sue for up to $1.5 million in damages, and states can sue for up to $3 million.

States with AntiSpam Solicitation Legislation (28)

Arkansas	*Iowa*
Rhode Island	*Kentucky*
Louisiana	*California*
Colorado	*South Dakota*
Tennessee	*Maryland*
Minnesota	*Connecticut*
Delaware	*Utah*
Virginia	*Kansas*
Nevada	*Missouri*
Illinois	*Wisconsin*
Oregon	*Pennsylvania*
Idaho	*Florida*
West Virginia	*Washington*
Oklahoma	*North Carolina*

Leading Anti-Spam Vendors and their Technology

There are well over 50 public and private anti-spam vendors that are releasing numerous new products each month, as spam continues to multiply at a feverish rate. We have outlined below some of the larger vendors and the type of technology they offer.

- **Network Associates** – Network Associates has developed McAfee SpamKiller, which was first introduced in May 2002. SpamKiller for Microsoft Exchange Small Business quarantines the spam messages into a dedicated folder on the email server. The messages can then be disposed of quickly by simply emptying the folder. Equipped with open-source technology from its Deersoft acquisition in January, Network Associates expects to release SpamKiller for its WebShield appliance in the second half of 2003. By placing SpamKiller on this gateway appliance, spam can be stopped before it reaches the email server. The NPD maintains that McAfee's consumer SpamKiller product has captured 90% of the consumer anti-spam market as of the first quarter of 2003.
- **Symantec** – Symantec is developing its own anti-spam technology, and it recently released its Antivirus for SMTP Gateway Version 3.1, which includes the company's latest anti-spam technology. The technology employs a heuristic pattern matching approach, which examines the behavior and content of an IP packet to determine if it is spam or not. The technology also uses blacklists, whitelists, and subject-line blocking. Symantec also supports 125 third party lists of known spamming addresses. Symantec will soon offer a stand-alone anti-spam product for consumers as well as bundling more anti-spam capability into its Norton Internet Security Suite for consumers.
- **Postini** (*private*) – Postini is the leading vendor in the hosted anti-spam space. As opposed to the installed solutions Network Associates and Symantec, Postini takes on the management and prevention of spam for a customer. Postini's spam filtering servers utilize a heuristic-based approach to analyze the behavior of spam emails and stop them before they reach a customer's network. The heuristic engine is constantly updated and revised based on the 100 million emails that Postini processes daily. Postini allows customers to configure their anti-spam protection with web-based controls that can set up

blacklists, for example. Trend Micro has recently licensed the Postini heuristic technology for its own enterprise anti-spam solution.

- **Brightmail** (*private*) – The company has emerged as the leader of the ISP anti-spam market, protecting many of the large service providers including MSN, AT&T, Earthlink, and Verizon Online. Brightmail offers anti-spam technology that uses several techniques, including "*honey pots*," to fight spam and stop it at the gateway of the network before it hits the email server. Real-time detection, 24-hour update service, and protection from spam attacks are crucial components of the solutions. Brightmail also offers an enterprise edition of its software.
- **SurfControl** – SurfControl offers a complete set of web, email, and instant message filtering products. Its Anti-Spam Agent compares incoming email messages to a large database of spam emails and their electronic signatures. The customer retains a large degree of administrative control and can decide how the spam is dealt with once it has been stopped; it can be quarantined, deleted, or allowed to remain. Customers can subscribe to daily signature updates.
- **CipherTrust** (*private*) – The IronMail anti-spam solution from CipherTrust garnered the top results for accuracy in a recent PC Magazine test of anti-spam products. IronMail employs multiple detection techniques that include the FirstAct update service, header analysis, content filtering, heuristic scanning, whitelists, and blacklists.
- **Tumbleweed** – With its Email Firewall Appliance for Anti-Spam, Tumbleweed has asserted itself in the anti-spam market. With a long list of large enterprise customers facing overwhelming amounts of spam, Tumbleweed has become a leader in the enterprise anti-spam arena. In April 2003, the company introduced the Dynamic Anti-Spam Service, which is an Internet-based subscription service with a heuristic engine at its core. The engine is continually updated by the Tumbleweed Message Protection Lab, which constantly analyzes spam and the new spamming techniques. The Dynamic Anti-spam service can be integrated into the Email Firewall Appliance for Anti-Spam. The Email Firewall, also available in software form, employs numerous anti-spam techniques that include "*exact match*" lexical analysis, HTML tag filtering, relay and server protection heuristics, support for blacklists and content-based signature lists, whitelists, and pornographic image detection. The Email Firewall also provides anti-virus protection through the use of McAfee anti-virus solutions.

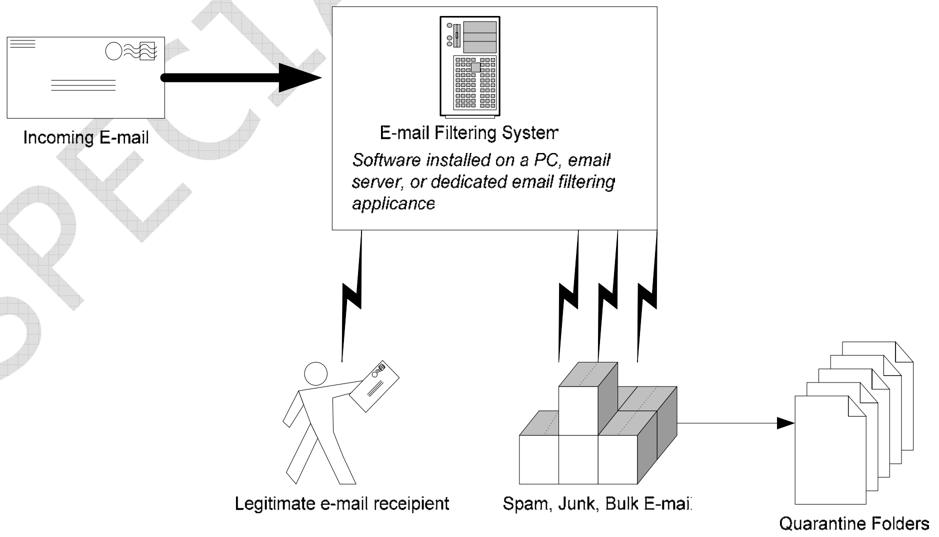
Methods of Spam Prevention

Some experts assert that to truly fight the problem of spam, it would take a complete overhaul of the Internet infrastructure. Email servers and systems would have to be reconstructed and shaped into a less open and free-sharing system, violating the original intent and spirit of the Internet.

On a less extreme front, there are numerous anti-spam technologies and methodologies introduced each month. Some methods work better than others, and it depends largely on the needs of an organization or individual user as to which method to employ. Unfortunately, most of these anti-spam solutions do not work well enough individually, and many times a "*layered*" approach is the best way to go. A layered approach involves the use of several anti-spam technologies at once.

Anti-spam tools come in the form of software and dedicated hardware appliances. For the home user, anti-spam protection is offered in the form of software that can be installed on the users desktop. A home user may also utilize anti-spam protection offered by an ISP that has the software running on its own servers. In a large network environment of an enterprise, anti-spam protection is either installed as software on the company's email servers, or the company may choose to outsource its protection to an anti-spam managed service provider. These service providers use their own anti-spam software installed on servers that filter their clients' email, wedding out spam. A managed anti-spam service will allow customers to configure the anti-spam solution to their needs, despite the fact that the anti-spam protection is hosted remotely.

Anti-spam software may also be installed on an appliance that normally sits at the Internet gateway of a network. A dedicated appliance can preserve the speed of the network, and it can also be installed easily with minimal configuration needed.


Incoming E-mail
E-mail Filtering System
Software installed on a PC, email server, or dedicated email filtering applicance
Legitimate e-mail receipient
Spam, Junk, Bulk E-mail
Quarantine Folders

Source: ***JMS Research***

List below are several anti-spam solutions which are used in one form or another by most anti-spam vendors. The most effective solution is most often a combination of the techniques listed below:

- **Whitelists and Blacklists** – These "*lists*" are simple email filtering systems. Whitelists will accept emails only from domains and IP addresses that are contained on a list, which is stored on a server. A blacklist works the same way; the list contains the domains and addresses of unwanted emails. There are many ways around this system for spammers, such as changing the address of the originating email repeatedly through the use of special software. Vendors will populate blacklists through the use of things like "*honey pots*," which are decoy email address set up to attract spam. When a spam email comes in the origin is traced, recorded, and added to the blacklist for future reference.
- **Basic Filtering** – Much of anti-spam software contains email filters used to detect messages with spam content. If a message contains enough elements of spam, a pre-set threshold is exceeded and the filter identifies the message as spam. This method can be effective when used in concert with other solutions, but when used alone it is not very scalable, nor can it evolve quickly enough to keep up with the latest spamming techniques.
- **Heuristic Engines** – These software programs examine the behavior of email traffic down to the packet level to determine if it is spam or not. Heuristic engines, through non-linear mathematics, identify a message as spam by correlating the results of various spam filters that are triggered by an incoming message. Only if sets of related filters specific to spam are triggered at the same time is the message classified as spam. If a group of filters is triggered, but there is no correlation that points to spam, the message is left alone. A true value of the Heuristic engine is that it becomes more refined and efficient as it processes more and more email. Heuristic engines can also detect directory harvest attacks.
- **Challenges / Response** – This technique requires the sender of an email to verify his authenticity before the email is accepted.
- **Bayesian Filtering** – This type of spam filter assigns a probability to an email as to whether it is spam or not. Bayesian filters scan an entire email message for key words or characters that are most often found in spam messages. Each word is assigned a "*spam*" probability and if the total probability of the email exceeds a sets threshold, it is classified as spam. Words that do not have high assigned probabilities help to decrease the overall chance that the message is spam, while those with high probabilities increase the chance of spam. Bayesian filters evolve as spammers alter their methods of sending messages.

The Problem with False Positives

A major gripe with anti-spam technology is the fact that it often blocks email that does not fall into the spam category, eliminating "good" email from user's inbox before it can be read. To reduce the number of false positives, anti-spam technology must become more customizable, so that the software can be tweaked depending on the type and origin of an organization's email. Most vendors allow a customer to configure the email filtering system in such a way that preserves legitimate messages, but no technology completely avoids the problem of false positives.

Conclusion

With approximately 31 billion emails are sent each day around the world; if spam continues to multiply as it has over the last two years, this number could double by 2006. Spam has become the Internet's version of the Black Plague and has cost associated with it – tangible, as well as intangible: It is estimated that businesses will spend in excess of $10 billion in 2003 in lost bandwidth, time spent filtering email vs. productive work, high virus potential, legal liability, and compromised IT resources.

What is the best course of action? Unfortunately, there isn't a straight one-line remedy or magic potion – that is if you still plan on using e-mail within you lifetime.

All of the products I have tested and read about seem to use a combination of techniques. Some combine heuristics with blacklists, or Bayesian analysis with source IP checking, and/or creating whitelist, or a combination of all accessible method.

At this time, a good starting point would be as follows:

- *Provide a clear definition (interpretation) of spam*
- *Enforceable laws governing certain events / action via the Internet*
- *Up-to-date and affordable anti-spam tools in the form of software (applications) and/or dedicated hardware appliances.*

Remember, no solution is perfect

Spam Glossary

Acquaintance Spam: These mails are sent based on some previous relationship between the sender and recipient, such as purchasing products online and providing your e-mail address for order information

Open Relay: A mail server that permits relaying by anyone. Spammers often abuse such systems.

Spam: Generally defined as e-mail that you did not request or expect to receive.

Spambot: A robot that specializes in gathering e-mail addresses for a spammer to use.

Spamhause: Used to refer to a site (*company*) that is actively producing spam.

Spamnest: A place (*company*) that produces spam.

Unsolicited Bulk E-mail (*UBE*): Same definition as Spam.

Unsolicited Commercial E-mall (*UCE*): Typically advertising some commercial service, product or company, these e-mails serve as an inexpensive marketing tool to sender to reach potential customers.