

## Spam

**Disclaimer!** Spam is a complicated and difficult issue. What is spam anyway? Unfortunately, there is no silver bullet that can stop all YOUR spam, yet never stop an email you want to receive.

This seminar tries to give you an overview of the development of spam, some guidelines to help reduce spammers getting your email address, and finally various methods of reducing the spam in your Inbox.

### Definition of Spam

#### Wikipedia:

*E-mail spam is a subset of spam that involves sending nearly identical messages to numerous recipients by e-mail.*

*Most definitions of spam are based on the e-mail being Unsolicited Bulk E-mail (UBE). That is, spam is e-mail that is both unsolicited by the recipients and there are many substantively similar e-mails being sent. Spam is usually also unwanted, commercial and sent by automated means and some definitions include those aspects.*

#### Webster:

*Pronunciation: 'spam*

*Function: noun*

*Etymology: from a skit on the British television series Monty Python's Flying Circus : unsolicited usually commercial e-mail sent to a large number of addresses.*

### So, where did "spam" get it's name?

In the Monty Python skit -called spam, a couple are attempting to order breakfast at a restaurant where all the menu items feature spam. The woman doesn't like spam, and asks the waitress if they offer any spam free dishes. Alas - there is no escape from spam.

Meanwhile a chorus of vikings at another table express their appreciation for spam by chanting "spam, spam, spam, lovely spam!"

Eventually any hope of meaningful communication is lost in the drone of "spam."

As the recipient directly bears the cost of delivery, storage, and processing, one could regard spam as the electronic equivalent of "postage-due" junk mail. However, this does not mean that all commercial email is spam; for example, some recipients may have opted in (i.e., willingly chosen) to receive the marketer's email. This is one of the key problems - what each of us regards as spam - can be different.

Spam is sent by organizations of varying sizes and motivations. Some are large, well-known companies; spam from these sources is sometimes called mainsleaze. A

widely-known instance of mainsleaze was Kraft Foods' marketing of its Gevalia coffee brand. "Advance fee fraud" spam such as the Nigerian "419" scam may be sent by a single individual from a cyber cafe in a developing country. Organized "spam gangs" operating from Russia or eastern Europe share many features in common with other forms of organized crime, such as turf battles and revenge killings.

## **Statistics**

### **Spam:**

- 1978 - An e-mail spam is sent to 600 addresses.
- 1994 - First large-scale spam sent to 6000 newsgroups, reaching millions of people.
- 2005 - (June) 30 billion per day
- 2006 - (June) 55 billion per day
- 2006 - (December) 85 billion per day
- 2007 - (February) 90 billion per day (or about 150 per PC)
- 600M PC's and 1.2B email addresses. Each day, 50 spam emails are sent per email address!
- 90% of all email is spam
- Annual cost (2004) to US businesses - \$10B. A 1000 employee company gets 3M spam emails a year (Bill Gates gets 4M spam emails per year - mostly get rich schemes!)
- Annual cost to non-corporate - \$500M
- 28% reply to Spam email
- 8% purchase from Spam email

### **Type of Spam Categories (% of total Spam)**

- Products 25%
- Financial 20%
- Adult 19%
- Scams 9%
- Health 7%
- Internet 7%
- Leisure 6%
- Spiritual 4%
- Other 3%

### **Most Annoying**

- Pornography 91%
- Mortgage and Loans 78%
- Investments 68%
- Real Estate 61%
- Software 41%
- Internet 7%
- Leisure 6%
- Spiritual 4%
- Other 3%

The major sources of spam in 2006 were the United States (23%), China (20%), Russia (10%) and South Korea (6%).

## Spam Made Up 94% Of All E-Mail In December

By Thomas Claburn InformationWeek - Jan 29, 2007 02:00 PM

*The Postini report says the rise of botnets and image spam is causing e-mail systems at some companies to melt down.*

*Legitimate e-mail now constitutes a rounding error when compared with spam, thanks to a standing army of more than a million zombie PCs waging war on in-boxes worldwide on any given day.*

*Some 94% of all e-mail last December was spam, according to Postini's annual communications intelligence report, which the managed e-mail security company released today.*

*In 2006, the volume of spam rose 147% by Postini's measure. The company attributes the surge in spam to PCs that have been commandeered by cybercriminals without the knowledge of their owners.*

*"There were two fundamental changes in the world of business communications in 2006 that are going to get even bigger in 2007," says Daniel Druker, executive VP of worldwide marketing for Postini. "The major event in communications security is the emergence of botnets. This has changed the game, the dynamics, and economics of the Internet security marketplace."*

### Why it's hard to detect

First of course - "One man's meat is another man's poison" - what one person considers spam, another may not.

1. In the early days of spam, the content could be checked for certain words, e.g. viagra, porn, stock price, etc. But this was soon overcome by use of special characters - v!agr@ which we all know what is meant - but is hard for a computer.
2. Most spam was sent from certain email addresses - which again could be easily blocked - known as black lists or blocked senders. In response to this spammers started to change the email address on every email sent out.
3. Although the "From" email addresses were constantly changing, the domain used or IP address where the email originated, often stayed the same. So databases (DNSBL) were created that kept a note of the IP address of anyone reported sending spam and made this information available to ISP's so that any email originating from these IP addresses could be dumped.
4. Spammers again responded by frequently changing domains. Initially they used open mail relays and vulnerabilities in mail servers. The response to this was requiring users to authenticate, and improved security on mail servers. So then spammers started buying domains, but needed to change them frequently. In June last year, 35 million new domains were registered, of which 32 million

were never paid for (used stolen credit cards, etc.) - probably used by spammers. Today, a spamming domain is usually only active for 4 hours!

5. The big growth in the last year has been in the use of "zombie" machines. It's estimated that today, 1/4 of all PC's - 150 million, have been quietly infected (mainly thru drive-by) and are acting as mail servers for the spammers. This means that many of us here, could well have a "zombie" and be part of a "botnet", sending out spam emails!
6. The use of hijacking a mail server, exploiting a poorly written mail script on a website, or using a zombie machine has caused problems for legitimate users. Their "stolen" email server IP address gets listed in a DNSBL, and any ISP using this service to block email (now days, only usually the smaller/less sophisticated ISP's), block all the email from a legitimate user of that server. This happened to our domain - gcclc.org
7. Image spam. The latest wave of spam is image based. To help avoid detection, the message (the spammers want you to read) is made as an image, which is hard for a computer to "read". However, the email also contains a lot of text, that is "good", to help convince the anti-spam program that the email is not spam. Over the last year, this has gone from 1% to 25% of all spam.

### **How do spammers get your email address?**

1. Dictionary attack
2. Spam bots searching websites
3. From infected computers.
4. From online vendors (not reputable ones)
5. From subscriptions (magazines, etc)
6. Buy and sell (amongst themselves)
7. Voter registration,
8. On-line "causes"
9. etc.,
10. etc.

### **Spam Safety Tips**

Spam can be extremely frustrating to some individuals while others don't seem to be bothered by it. If you are one of the many trying to reduce the amount of Spam in your inbox, the following guidelines can help.

#### **Encrypt email addresses**

When you first create your email address, come up with a combination of letters and numbers that are cryptic in nature—something you couldn't find in a dictionary. For example, instead of using sally, or sally1, or sallysmith, choose: s18all56y. This number/letter combination is inconvenient for humans to remember but it provides

more of a challenge for the spammer's programs to randomly send Spam to your email address.

**Use Fake email addresses**

On many websites, you are required to enter an email address into a standard form before you can proceed through the website. If you don't feel comfortable giving out your email address to the particular website, leave a fake email address.

**Guard your email addresses**

Treat your email address the same way you do most of your personal information. Don't give it out to anyone you don't trust. If you are not sure you can trust a particular website, read their privacy policy to see what they will do with your email address.

**Use Bcc to send to a lot of people**

Jokes or other "informative" emails are often forwarded to many friends and/or family. Frequently, all the recipients (either individually or using a group/distribution list) are placed in the To: or Cc: address lines. All the recipients now see the all these email addresses. Only one of these recipients needs to have a worm or spyware, to have the complete list available to spammers.

**Don't open Spam**

If the Spam is HTML (one of those attractive graphic emails) and you open it, the graphic is pulled from the spammer's server. Your computer informs the spammer that your email address is in use.

**Don't Reply**

Remember those pesky telemarketers or the unrelenting door-to-door salesman? Once you answer the telephone or door, they know you are home and are a challenge to get rid of. The same is true of spammers. Once you reply to a Spam email, you have just confirmed for the spammer the legitimacy of your email address.

**Don't post your email address**

Once your email address has been placed on a website (personal or corporate) or entered into an online guest book, newsgroup, contact list, ezine, chat room, or a host of other online activities, you have just invited a spammer to take your email address. Spammers "harvest" your emails through programs called spiders, crawlers, and bots. These programs scour the web for email addresses to be used in the spammers future email campaigns.

**Opt -out**

When you are purchasing something online or signing up for a service or promotion, be sure to opt-out on any additional services or promotions you don't want cluttering your inbox.

**Don't unsubscribe (?)**

Honorable marketers will unsubscribe your email address if you request it, but distinguishing between legitimate companies and those who are not is a challenge. Check their privacy policy and complaint procedures. Submitting an unsubscribe request can be used against you—your email address may be confirmed by or sold to spammers. When this happens, your Spam will increase when you thought you'd submitted an unsubscribe request. Generally, reputable companies are fine (Home Depot, Lands End, HGTV, etc.)

**Use a Spam filter**

No matter how thorough your Spam prevention measures, you will still receive Spam—accept this reality. Even though the perfect Spam filter doesn't exist, there are many good Spam filters in the marketplace

**Spam filter techniques.**

As with any major and growing problem, there are a 1,001 companies out there, offering to sell you a solution that will eliminate all your spam. There is no such product - unless you stop all email! However, there are a number of different techniques that can be used to significantly reduce the amount of spam you might receive.

With all these techniques, you do need to continually "train" them in one form or another, and to make sure you receive ALL the emails you want to see, you will have to go into your spam/bulk folder and scan for emails wrongly classified as spam. No way around this. Sorry.

**Rules based.**

As the name implies, this uses "rules" to look for certain words in any of the headers. The headers are primarily To:, Cc:, From:, Subject:, Body:, Attachment. So you can set a rule to look for "viagra" in the subject or body. If found, take some action (move to a Junk folder or delete it., etc.). The problem is that these rule based filters look for an exact match. You also need to be careful, because setting a rule to look for "stock" in financial spam, will find stockcar Woodstock, stocking, etc. Can be very effective used for managing a "White List".

**Bayes filter.**

Bayes' Theorem is based on probability theory. Using collected data points it statistically attempts to determine the probability that a message is spam and whether it should be filtered or not. In other words, the Bayes filter looks at all the emails you identify as spam, and tries to statistically look for words and phrases that constitute spam.

**Fuzzy logic**

Fuzzy logic anti spam filters make observations about a message to determine to what extent it's a member of the "spam set" and to what extent it's a member of a the "non-spam set." If it crosses a certain threshold it's considered to be spam and is filtered. A fuzzy logic filter uses both machine collected data in addition to human selected data points to make its spam filtering decisions. By combining both machine and human intelligence, it's claimed that fuzzy logic has superior spam fighting performance.

*If these spam filtering techniques sound similar - it's because they are. In practice the differences between Bayesian filters and a fuzzy logic filters spam stopping ability comes down to the data set.*

**Black list**

Black lists are basically rules that block particular email addresses or domains. These are not very effective against modern spam – where the sending domain and email address are constantly changing. Can be useful for deleting email newsletters or those forwarded emails from your friends and family!

**On line data base of known spammers (DNSBL)**

Many anti-spam programs include this technique. As I said earlier, many spammers are "hijacking" legitimate mail servers, and so information from these databases has to be used with care. However, these databases are kept very current and can be helpful in identifying large scale spammers.

**White list (Safe Senders list)**

This is a list of email addresses from which you **will** accept email. All others are rejected. This method is the most effective, but needs to be constantly updated.

An additional feature that some programs use, is called a "challenge/response". This is where an automatic "challenge" is sent as a response to any email received from someone not on your White List. Spammers never receive replies to their emails, so the challenge goes unanswered. However, a "real" person receiving the challenge can respond and will be put into your White List. However, legitimate emails from say your Credit Card company will also go unanswered.

**Reducing Spam in real life**

Ok, so what do we do in real life? Well, this is different if you're using web mail or POP3 mail. Web mail is where you use your Internet browser to read and send your email. POP3 email is where you use an email client such as Outlook Express, Thunderbird, Endura, etc.

## Web mail

When you use web mail, your emails all stay on your ISP or web mail provider's computer. As such, you have very little control of which anti spam program is used. The large providers such as Yahoo (provide email for SBC, AT&T, Pacbell, etc.) Hotmail, Google, etc. provide excellent spam filtering techniques. They will use a combination of the above techniques. So if the sender of an email is in the DNSBL, they will look at other factors before putting it in the Spam, Junk or Bulk folder.

**However, if a legitimate email ends up in the Bulk folder, its up to you to see this and identify it as "Not Spam"** When you change a "Spam" email to "Not Spam", it is automatically added to your White List – so any emails from that sender in the future, will be automatically put in your Inbox. So you still need to check your spam or bulk folder.

Yahoo has another great feature, called AddressGuard – to create an individual email address for each person, or class of person. First you create a Base Name. I created rogerspace. Now, I can create an email address [rogerspace-utilities@yahoo.com](mailto:rogerspace-utilities@yahoo.com) or [rogerspace-creditcards@yahoo.com](mailto:rogerspace-creditcards@yahoo.com) or something very specific, such as [rogerspace-bankofthewest@yahoo.com](mailto:rogerspace-bankofthewest@yahoo.com) In this way, you can give a very restricted and specific number of people a particular email address. If it ever gets "spammed" it can be deleted.

Many of the smaller ISP's do not supply the level of sophistication detailed above.

## POP3

This is where you can install a program, that "sits" between your ISP mail server and your email client (e.g. Outlook Express). There is a huge selection from freeware to expensive solutions. Again, many of these programs use a number of methods to try to identify spam, without trashing good emails. Most email providers allow you to manage your email by either web mail or POP3. Remember however, if you use Yahoo mail (e.g. –sbcglobal.net) and you connect by POP3 and have the Yahoo spam filter turned on, **you must still go into the Bulk folder** (use web mail to view) to identify any good emails misidentified as spam.

One of the differences is usability. Some of these programs

### Outlook Express – Rules

Create and demonstrate a simple rule as well as using rules to create a White list. Creating a White list or Safe senders list rule, is actually quite easy, as you can import your entire address book with one mouse click!

### Outlook Express – Bayes filter – K9 (freeware)

The freeware program K9, uses a Bayesian filter, and therefore needs some "training". This program also has a White list, a blacklist, Regex filters for advanced users. The program has an easy interface, but is not "elegant".

### **Outlook Express – White list (Safe Senders list) – Computer Associates**

Demonstrate the use of an integrated White list anti-spam program. Being integrated to the Outlook Express program, it is very easy to use.

### **Outlook Express – Server based – MailWasher**

One of the disadvantages of the above programs, is that the spam e-mail needs to be downloaded to your computer, before the anti-spam program can identify which is spam. For those on a dial-up connection, this means a lot of wasted time, just downloading spam. The MailWasher program looks at your e-mail whilst it is still on the server, allows you to review the headers and make a decision about whether the e-mail is spam. All e-mail marked as spam, will then be deleted at the server and not downloaded to your PC. This program neatly integrates a White list, Black List, on line DNSBL, Bayesian filter as well as user created rules and filters. A very complete and easy to use program.

### **Phishing**

Online phishing (pronounced like the word fishing) is a way to trick computer users into revealing personal or financial information through an e-mail message or website. A common online phishing scam starts with an e-mail message that looks like an official notice from a trusted source, such as a bank, credit card company, or reputable online merchant. In the e-mail message, recipients are directed to a fraudulent website where they are asked to provide personal information, such as an account number or password. This information is then usually used for identity theft.

These fishing expeditions - identity theft attempts using for phony e-mails sent to gain Social Security, credit card and bank account information - have doubled in the last two years. Many of these thieves are using new "lures" and targeting consumers with incomes of \$100,000 or more.

	<b>2004</b>	<b>2006</b>
<b>Losses from phishing attacks</b>	<b>\$137 million</b>	<b>\$2.8 billion</b>
<b>Number of US adults who received at least one phishing e-mail</b>	<b>57 million</b>	<b>109 million</b>
<b>Number of victims</b>	<b>53 thousand</b>	<b>2.25 million</b>
<b>Per-victim loss</b>	<b>\$257</b>	<b>\$1,244</b>
<b>Money recovered by consumers</b>	<b>80%</b>	<b>54%</b>

(Source: Gartner)

To avoid getting caught, don't open e-mails from anyone you don't know even from recognized businesses you don't deal with (and call your bank or credit card company if you receive e-mails asking you to update your account information on line).

## Summary

### Web mail

With web mail you are entirely dependent upon the Web mail provider's capability and offerings. Many of the large Web mail provider's, like Yahoo, Google, Hotmail, etc. provide very sophisticated and often flexible and configurable options to help you reduce your spam e-mail.

All of these systems however, do need to be trained. **It is therefore important that you go into your Junk/Bulk/Spam folder and identify any e-mail that is not meant to be junk.**

For those willing to take the time, the Yahoo AddressGuard is a great system for protecting your e-mail address.

### POP3 mail

Here there are a multitude of products that use either a single or often a combination of methods to reduce your spam mail. At the end of the day, the most effective method is to use a White list, or approved senders list. This does demand a small amount of ongoing effort to keep your white list or approved senders list up to date. Using the built-in rules capability of Outlook express works well, but isn't as convenient as a purchased product like the one from CA (Computer Associates).

Above all, protect your e-mail address, as it is a valuable piece of your personal information.

### Here's some Links to useful resources.

- [http://en.wikipedia.org/wiki/Spam\\_e-mail](http://en.wikipedia.org/wiki/Spam_e-mail) Wikipedia Reference
- <http://spam.abuse.net/userhelp/> Links to resources and anti-spam filters
- <http://spamlinks.net/> More links to many anti spam resources
- <http://spamlinks.net/filter-client-win.htm> (More spam filters)
- <http://spam-filter-review.toptenreviews.com/> Review of some spam filters
- <http://keir.net/k9.html> Freeware Bayes filter (plus White and Black list)
- [http://shop.ca.com/STContent/landingpages/Antispam/ASPM001/index.aspx?sc\\_lang=en-US](http://shop.ca.com/STContent/landingpages/Antispam/ASPM001/index.aspx?sc_lang=en-US) Computer Associates anti spam program (Or Google ca spam). Paid, easy to use White list – integrates to Outlook Express.
- <http://www.mailwasher.net/> Mail Washer free and paid anti-spam versions. Uses multiple methods for detection
- <http://www.spambutcher.com/> Spam Butcher – fuzzy logic anti-spam