Manager's Guide to

Coping with Spam

From The Open Group Messaging Forum
Prepared by Leslie Ogonowski of Johnson Consulting

THE Open GROUP

The views expressed in this Guide are not necessarily those of any particular member of The Open Group.

Manager's Guide to Coping with Spam

ISBN: 1-931624-37-2

Document No.: G034

Published by The Open Group, March 2004.

Any comments relating to the material contained in this document may be submitted to:

ogpubs@opengroup.org

contents

Introduction 1
What is spam? 3
Why should we be concerned with spam? 12
What can we do to minimize spam? 18
Best Practices 32
Summary 35
Glossary 37

About The Open Group

The Open Group is a vendor-neutral and technology-neutral consortium, whose vision of Boundaryless Information Flow will enable access to integrated information within and between enterprises based on open standards and global interoperability. The Open Group works with customers, suppliers, consortia, and other standards bodies. Its role is to capture, understand, and address current and emerging requirements, establish policies, and share best practices; to facilitate interoperability, develop consensus, and evolve and integrate specifications and Open Source technologies; to offer a comprehensive set of services to enhance the operational efficiency of consortia; and to operate the industry's premier certification service.

Further information on The Open Group is available at www.opengroup.org.

The Open Group has over 15 years' experience in developing and operating certification programs and has extensive experience developing and facilitating industry adoption of test suites used to validate conformance to an open standard or specification.

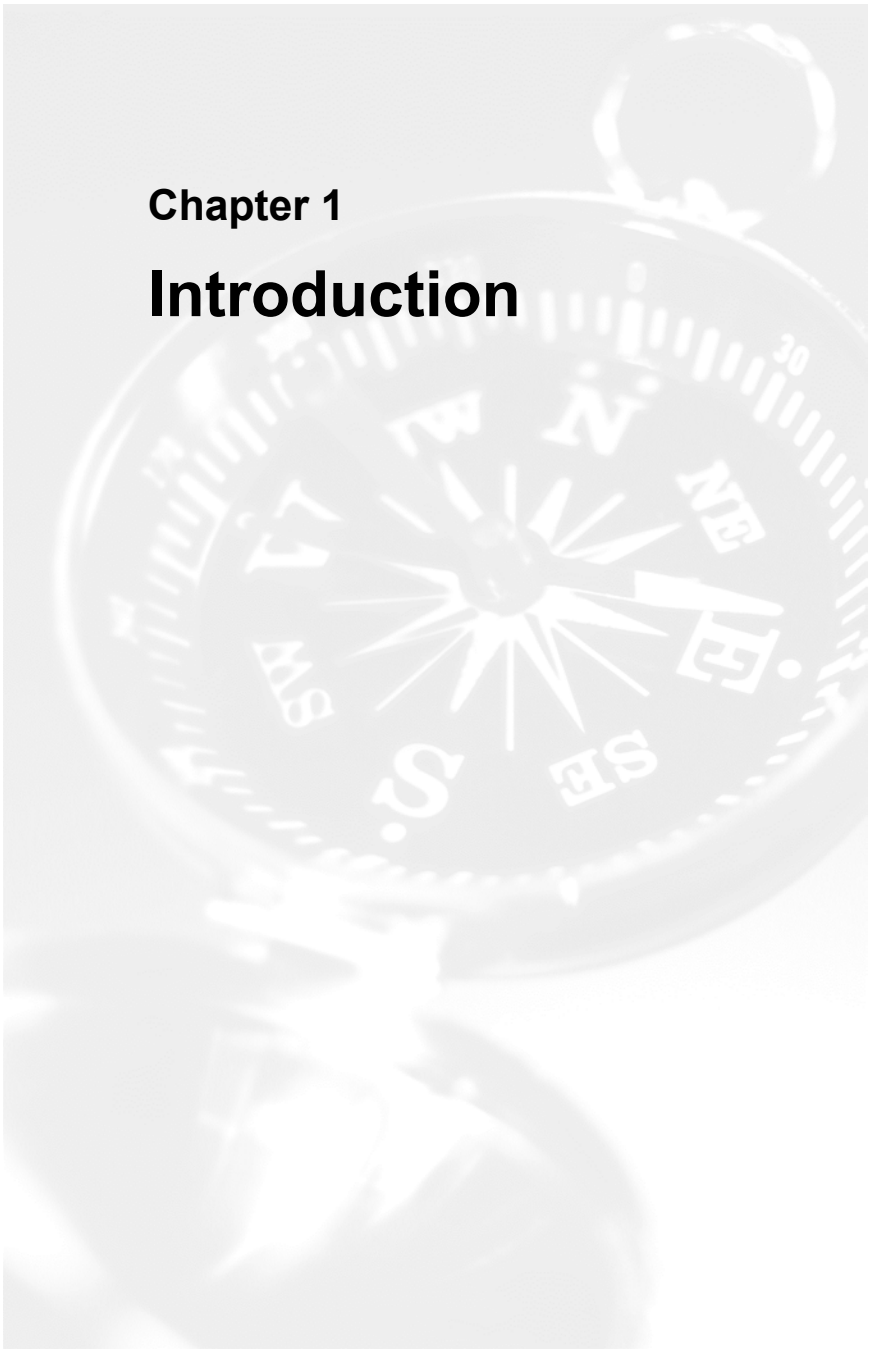
The Open Group publishes a wide range of technical documentation, the main part of which is focused on development of Technical and Product Standards and Guides, but which also includes White Papers, Technical Studies, and Business Titles.

A catalog is available at www.opengroup.org/publications.

Trademarks

Acknowledgements

The Open Group gratefully acknowledges the members of The Open Group Messaging Forum for their contribution of raw material and for reviewing and refining this Guide.



Chapter 1

Introduction

Email has grown from a novel communications tool to a mission-critical aspect of corporate life. Not only does it provide a simple and efficient means for delivering personal and business messages, it is also a backbone service for electronic transactions. Marketers find email a very inexpensive way to reach potential customers. Many customers find email an effective way to learn about new products and promotions.

We have worked for many years to make email messaging systems reliable as well as easy-to-use and administer, but unfortunately we seem to have made it a little too easy.

Unscrupulous marketers have latched onto email as an effective way to reach a huge number of people at a minimal cost. These marketing emails, and other unsolicited messages, commonly known as *spam*, are threatening to overwhelm the Internet and make email unusable.

To keep email as a viable communication option, we need to find ways to minimize the effects of spam, both for the spam received in our corporate messaging systems, and the spam that can be sent using insecure email relays. There is no one technology that will completely block all spam messages while allowing every legitimate message through, and legislation alone will not be able to eradicate spam.

In this Guide we discuss ways to bring the amount of spam we receive in our corporate email systems down to a level where we can cope with it.

Chapter 2

What is spam?

A simple definition of spam is unsolicited email, generally commercial or promotional in nature, usually sent in bulk. The key word here is *unsolicited*.

Spam comes in many forms. Aside from commercial advertisements for products or services, many non-commercial messages can also be considered spam, such as:

- Chain letters
- Charitable solicitations
- Games
- Hate mail (including racist and sexist messages)
- Jokes and funny stories
- Malicious code (email borne viruses)
- Petitions
- Political messages
- Pornography
- Prayers
- Press releases
- Promotional messages
- Scams and "get rich quick" schemes (such as pyramid schemes)

Another factor often used to define spam is whether the message is unwanted by the mailbox owner; either by the individual user or the corporation that owns the messaging system.

Depending on corporate messaging policies, personal messages may be treated as spam by system administrators. On the other hand, bulk commercial email of interest to the recipient may technically qualify as spam. Everyone has a different opinion of what is and isn't unwanted email; corporate messaging policies can serve here as a useful guideline.

Instant messages and wireless messages can also be defined as spam, if they are unsolicited, but that discussion is outside the scope of this document.

There is some debate about the source of the term, but the generally accepted version is that "spam" comes from a Monty Python's Flying Circus skit. A group of Vikings in a restaurant sings a chorus of "spam, spam, spam . . ." in an increasing crescendo, drowning out other conversation. This is comparable to unsolicited email drowning out normal Internet communications.[1]

What is not spam?

Spam should not be confused with permission-based email marketing messages. There are legitimate marketing services that send messages to customers who have specifically requested to be notified of promotions, contests, sales, etc. Some marketers (for example, MyPoints and Yahoo) provide services in exchange for a user agreeing to accept promotional email. Although these messages may have the exact same content as a spam message, they are not spam, because the user has agreed to receive them.

Note that Yahoo has reset its members' marketing preferences to accept all promotional email. We would consider those messages unsolicited and unwanted, and therefore spam.[2]

Types of spam

There are, of course, different types of spam.

There is nuisance spam – for example, nonsense chain letters – which is mostly harmless except for the storage space and time wasted–. System administrators may have to deal with other types of inappropriate personal email that is outside the limits of the corporate messaging policy.

Also mostly harmless are commercial or promotional messages from legitimate marketers or companies with whom the recipient may have a relationship. However, if these messages are outside corporate

---

[1] *Spam and the Internet*, Hormel Foods; refer to www.spam.com/ci/ci_in.htm.
[2] *Yahoo Resets Member Spam Preferences*; refer to www.pcworld.com/news/article/0%2Caid%2C91984%2C00.asp.

messaging policy, although they technically are not spam, system administrators may still treat them like spam as they take up bandwidth and storage space.

Then we have abusive spam, usually in the form of unsolicited commercial email sent to random recipients who have no prior relationship with the sender. Many of these messages are for dubious claims or outright fraud. Others may be obscene, violating corporate anti-pornography policies.

Abusive practices of spammers

Additional qualifiers that cause email to be identified as spam come from the practices of unethical marketers, called spammers by the online community.

These spammers abuse email by transferring the cost of sending unsolicited messages to Internet Service Providers (ISPs), open relay email servers, corporate messaging systems, and individual recipients.

Illegal, unethical, and immoral content

Many spam messages contain blatantly illegal offers, scams playing on the recipients' gullibility, and even pornographic advertisements. Some spammers insist they won't market anything illegal or anything they don't believe will work – though few of the products advertised seem to be able to live up to their claims.

To whom it doesn't concern

Spammers frequently send their messages to people who have expressed no prior interest in the products or services they are advertising. Spammers often work on a commission basis. A 1% response can be enough to turn a profit for a spammer, so the more messages they send, the greater the potential to make money.[3]

---

[3] Interview with Ron Scelson and Intellireach, 06/25/03; refer to www.intellireach.com/events/0625w.html.

Spammers can purchase lists of email addresses from a marketing service or advertising broker at very little cost (for example, one site offers 60 million email addresses for $150[4]). It is not difficult for them to then find other marketers to trade address lists with, and to harvest addresses on their own.

It is easier for a spammer to send a message to 1,000,000 email addresses than it is to clean up a list or target specific recipients who have expressed interest in a product or service.

Address harvesting

Spammers send "spiders" or agents to scour web sites, mailing lists, and newsgroups for addresses. Even details from instant messaging programs such as ICQ are not safe. Spammers can be a little more selective using this method, however. As an example, they may harvest names from web sites associated with golf to target a mailing about a marvellous new golf ball that guarantees at least 400 yards for every stroke.

Directory harvest attacks

Directory harvest attacks overwhelm email servers. Spammers send variations of email addresses to a domain email server. Any addresses that do not fail are considered valid and are added to their mailing lists.

These attacks can be easily recognized. The email server generates an unusually large number of failed messages in a short period of time.

There is not much that can be done to prevent directory harvest attacks, however.

Opting out brings more spam

Recipients who click on a spammer's "remove me from this mailing list"

---

[4] *60 Million Fresh Email Addresses + Opt-Ins*; refer to http://sources.redhat.com/ml/crossgcc/2001-09/msg00015.html.

link are often only confirming that they have a monitored email address, thus inviting more spam.

Setting web beacons

Web beacons are another way to verify active email addresses – if an HTML message is downloaded and rendered, even in preview mode, a web beacon tells the sender that the email address is being used.

Using open email relays

Spammers often use insecure email servers to redistribute their spam, without the owner's permission. In effect, they "steal" server resources. This practice can cause delays or disruptions to an email service, also known as *denial of service* attacks – the server is too busy processing spam to deal with regular email.

Often spammers choose backup servers pointed to by secondary email exchange (MX) records. Backup servers usually accept and relay all email to the primary MX host without checking it.

"Drive-by" spamming

An interesting new tactic is called *drive-by spamming*.[5] Spammers drive around and find unsecured wireless networks. Sitting in a van with a laptop, spammers can send email from "inside" the network to the Simple Mail Transfer Protocol (SMTP) server. Any messages sent by the spammer would appear to come from within the organization's network.

Open proxy servers

Another way spammers exploit open relays is through insecure or misconfigured proxy servers. Proxies are normally set up for users in a

[5] "*Drive-by spam hits wireless LANs*", Graeme Wearden, CNET; refer to http://news.com.com/2100-1033-956911.html.

network for control over Internet access or to cache data for frequently used web sites – properly configured, they route data from a Local Area Network (LAN) to the Internet. However, if they are misconfigured, they may be able to route data from the Internet into a LAN, or perhaps to another part of the Internet. Using an open proxy, spammers can find internal email servers and use them to route email, or they can anonymously abuse SMTP servers elsewhere on the Internet.

Proxy servers can be installed without the system administrator's knowledge. In January 2003, the virus Sobig.a was released. This virus downloads a Trojan executable that, as part of its payload, installs a specially modified proxy server that is hidden, runs on non-standard ports, and does not generate a log. It is not known whether this virus was developed by a spammer for practical reasons or simply by a hacker for malicious reasons.

In June 2003, up to 70% of spam was sent from hijacked machines, according to Mark Sumner, Chief Technology Officer of MessageLabs, Inc.[6]

Forged headers and "Joe-jobs"

Spammers often forge or tamper with the headers of email messages to conceal their identity and location. This could lead to a system administrator receiving emails complaining about the spam it appears he sent.

The practice of using a fake return address is known as a *Joe-job*. A spammer may do this maliciously, in order to damage another organization's reputation and possibly trick their provider into revoking their Internet access, or they may do it simply to hide their own identity. Joe-jobs are named after Joes.com, which was victimized in this way by a spammer some years ago.

[6] "*SoBig spam-virus still spreading*", Bob Sullivan, MSNBC; refer to www.msnbc.com/news/931205.asp.

The "Big Murkowski"

Spammers sometimes try to give their messages an air of legitimacy with text such as the following:

```
This message is sent in compliance with the new email
bill section 301. Per Section 301, Paragraph (a)(2)(C)
of S.1618, further transmissions to you by the sender
of this email will be stopped at no cost to you. This
message is not intended for residents in the State of
WA, NV, CA, and VA. Screening of addresses has been
done to the best of our technical ability. If you are
a Washington, Virginia, or California resident please
remove yourself. We respect all removal requests.
```

Or, to make themselves less likely to receive negative responses, spammers may even sound a little threatening:

```
Under these provisions, this letter cannot be dealt
with as spam, and no further action can be taken by
the reader against this company/person. Any report of
this letter as spam to any independent agency or site
is a violation of U.S. Bill S.1618 TITLE III of the
U.S. Congress and will be dealt with promptly.[7]
```

At this point in time there is no such law in the U.S. If this disclaimer is actually found in an email, it is obviously spam, and the spammer is trying to avoid complaints.

This practice is so prevalent it actually has a name: "Murk", as defined in the online spam glossary:[8]

(n.) A disclaimer at the end of an email spam assuring you that the spam complies with Bill S.1618 which makes the spam legal. Also known as a "Murkogram".

(v.) The act of sending spam containing a Murkogram.

The term comes from Frank Murkowski (R-AK), the Senator who wrote Bill S.1618. This would have made certain types of spam illegal, unless the message included full contact information at the start and made no attempt at hiding its origin.

---

[7] Refer to www.internet-tips.net/Email/SPAM_1618.htm.
[8] Refer to www.rahul.net/falk/glossary.html.

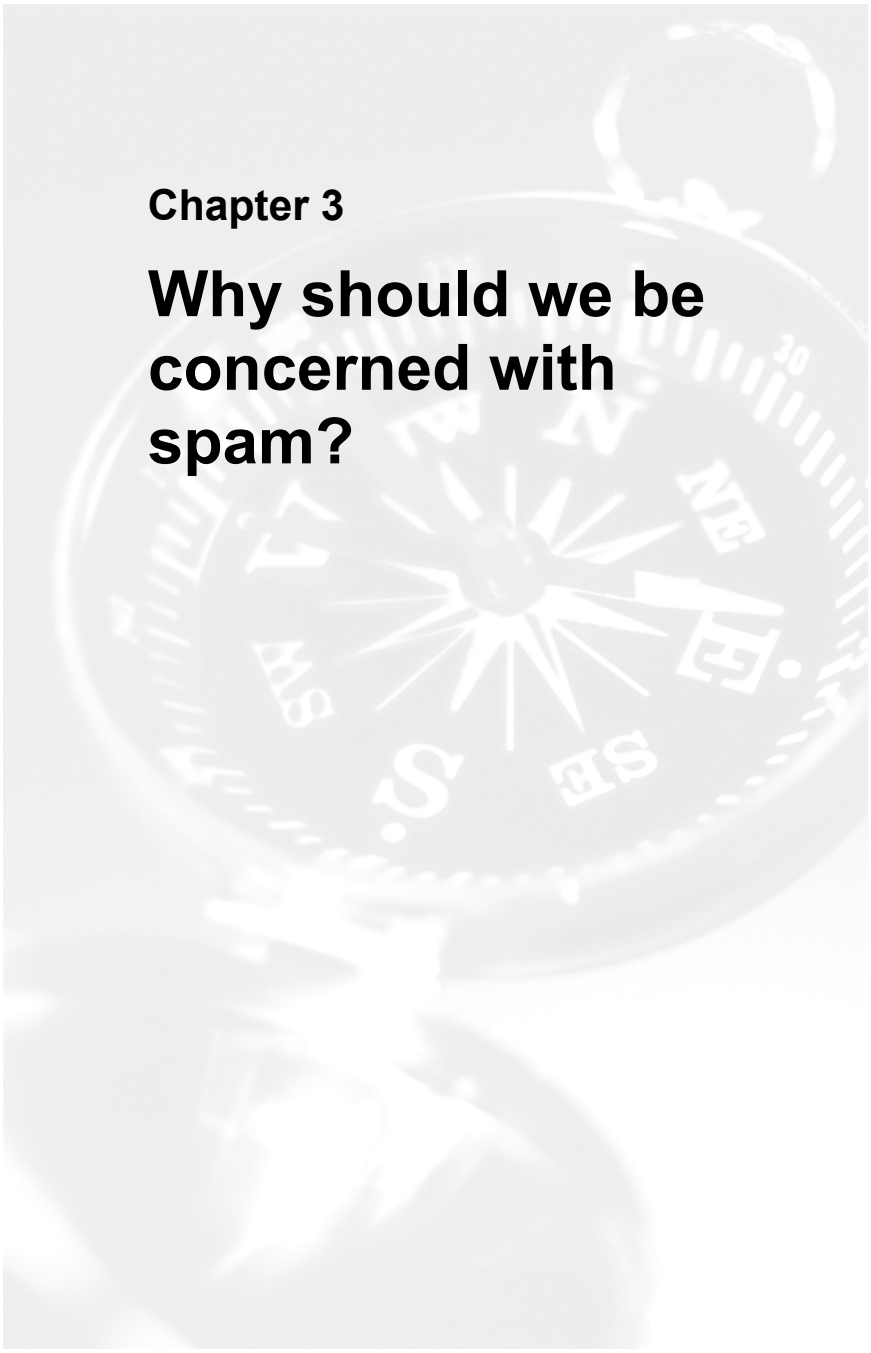
Filter busters

Spammers also add *filter busters* (strings of nonsense characters) to the subjects and bodies of their messages or send HTML graphics, in the hope of confusing filters that look for known spam messages.

Many spammers actually buy filtering software to test and fine-tune their messages to get around the filters.

Chapter 3

Why should we be concerned with spam?



To some companies, spam is just a fact-of-life, accepted as an annoyance but basically ignored. Users simply delete their spam on a regular basis and work goes on. However, companies with this attitude do leave themselves open to wasted time, money, and resources, as well as legal risks.

A Radicati Group study concluded that 94% of companies consider spam to be a very serious problem, but 43% still do not have a formal anti-spam policy in place.[9]

Reasons to stop spam

Spam attacks disrupt electronic messages and transactions, impacting an organization's ability to do business.

Spam wastes system resources, including bandwidth, email server processing cycles, and storage capacity. Spam can overwhelm email servers that are not secured against relaying.

Spam wastes human resources; not just the time of employees who read and respond to spam, but that of system administrators, help-desk staff, and human resources personnel as well.

Spam can violate corporate policies regarding non-business use of corporate messaging systems. Offensive or pornographic spam can also violate corporate anti-harassment policies. Content filtering can reduce the risk of legal exposure due to a "hostile environment" when the contents of messages are offensive to employees.

Business-related messages and solicited advertising from legitimate marketers can get lost in the proliferation of unsolicited commercial advertisements and other spam.

Risks of blocking spam

However spam is filtered, false positives are a risk. A false positive is a

[9] *Spam and Virus "Blended Threats" Top Email Dangers in 2002*, Advisor; refer to http://marketingadvisor.net/doc/11611.

message that has the characteristics of spam, but is actually a legitimate message. Business deals and important information can be lost because of a false positive.

Other than false positives, the only major risk of blocking spam is to be sued by the message sender or recipient on the basis of censorship or violation of personal privacy. It can be argued that a corporate mailbox is owned by the company; therefore, all messages are subject to corporate policies.

The costs of spam

Fighting spam does have its costs, but allowing spam to continue to grow unchecked could be disastrous. Ferris Research stated that spam cost U.S. companies $8.9 billion in 2002, and $2.5 billion for European businesses. These costs are roughly divided between soft costs such as lost productivity *versus* hard costs such as equipment, system administration, and help desk personnel.[10]

Is it worth the cost to fight spam versus simply accepting it as a side-effect of using email?

Prices for anti-spam solutions start at approximately $15 to $20 per user.[11]

Vendors and researchers provide greatly varying results on the cost of spam per employee per year, from Ferris Research's $168[12] to Nucleus Research's $874.[13]

The Radicati Group's study – *Anti-Spam Market Trends 2003-2007* – suggests that deploying an anti-spam solution is often more than 50%

---

[10] *The Spam Police*, Network World; refer to www.nwfusion.com/research/2001/0910feat.html.
[11] *Spam by the Numbers*, ePrivacy Group; refer to http://cobb.com/spam/numbers.html.
[12] *The Next Step in the Spam Control War: Greylisting*, Evan Harris; refer to http://projects.puremagic.com/greylisting/.
[13] *New laws will make it "legal" to spam your work address*, Silicon News; refer to www.silicon.com/news/165/1/4970.html.

less costly than living without one.[14]

Is it worth the cost of additional software and hardware to filter spam versus the cost of additional software and hardware to process and store spam?

The *Anti-Spam Market Trends 2003-2007* study reports: "A 10,000-user company, running Microsoft Exchange 2000, is deploying an average of five messaging servers just to process spam in 2003, out of a total of 21 messaging servers. By 2007, if nothing is done to stop spam, this will spiral to 25 servers processing spam, out of a total of 50 messaging servers.[15]"

In the US, the Radicati Group estimates that an organization of 10,000 users with no anti-spam protection will spend an average of $49 per mailbox per year processing spam messages in 2003.[16]

Is it worth the cost of a system administrator's time to install and manage spam filters versus the time it takes users to sort out legitimate email from spam?

Osterman Research reports that spam costs for a system administrator run at about $15 per user per year.[17]

MessageLabs estimates that spam costs over $750 per user per year in wasted time alone.[18]

Is it worth the risk of losing business due to false positives versus the cost of receiving obscene material?

12% of spam received by Brightmail for June 2003 was "adult" in nature.[19] Users can lose their jobs for viewing obscene material at work,

---

[14] *FTC finally realizes that spammers lie*, ZD Net UK News; refer to http://news.zdnet.co.uk/story/0,,t269-s2134105,00.html.
[15] *Virginia Seeks Jail for Spammers*, Internet Advertising Report; refer to www.internetnews.com/IAR/article.php/2199001.
[16] *"Buffalo Spammer" Arrested*, InternetNews.com; refer to www.internetnews.com/xSP/article.php/2206311.
[17] Soumushou Research; refer to www.soumu.go.jp/joho_tsusin/top/m_mail.html (Japanese only).
[18] Interview with the Spammer, Intellireach Webcast; refer to www.intellireach.com/events.
[19] Refer to www.brightmail.com/spamstats.html - spam_categories.

or be sued for sexual harassment. Companies can also be sued for allowing the messages to be delivered to the desktop.

These are questions each organization must answer in the best interests of their business, though statistics indicate that the costs of spam are growing. Radicati suggests that the $49 per user per year figure for 2003 will grow to $257 per user per year by 2007.[20]

How much spam is out there?

What makes spam such a problem is the amount of it. Even with all the developments in anti-spam technology, spam is increasing at an alarming rate. Personal and business email inboxes are flooded with spam messages, and ISPs and corporate email servers have to cope with steadily increasing traffic.

Statistics do vary, but all agree that spam will only continue to increase.

In June 2003, almost 7.7 million different spam messages were caught by Brightmail's Probe Network, which has a statistical reach of 250 million mailboxes.[21] For 2002, Brightmail had blocked more than 50 million spam attacks.[22]

Brightmail projects that, by September 2003, over half of the messages sent via the Internet will be spam.[23]

At any given time, 5% to 30% of the email messages received at AOL are spam.[24] The AOL Time Warner Internet unit is blocking almost a billion unsolicited bulk email messages every day.[25]

---

[20] Refer to www.radicati.com/cgi-local/brochure.pl?pub_id=202&subscr=&back_link=/single_report/.
[21] Refer to www.brightmail.com/spamstats.html.
[22] *Brightmail Reveals Annual Top 10 Spam Messages for 2002*; refer to www.brightmail.com/pressreleases/121202_top_spam.html.
[23] *Spam on course to be over half of all email this summer*; refer to www.brightmail.com/press-releases.html.
[24] *AOL spam dispute escalates*, CNN Interactive; refer to www.cnn.com/TECH/9712/31/aol.addresses/.
[25] *AOL: Spam problem is getting worse*, Internet Advertising Report; refer to www.internetnews.com/IAR/article.php/2091641.

Research firm Jupiter Media estimates that consumers will be inundated with 206 billion junk emails in 2006, which is double the number received in 2002.[26]

In a December 2002 study, the Gartner Group found that the junk mail rate for companies is approaching 50% and continuing to rise.[27]

The Radicati Group finds that one in three corporate email messages are spam; they expect this to rise to 39% by 2006.[28]

Medium-sized companies routinely get 20,000 spam messages per day, according to the Meta Group.[29]

Brightmail estimates that the average spammer sends 250,000 messages per day.[30]

---

[26] *The Anti-Spam Cookbook*, Network Computing; refer to www.networkcomputing.com/1320/1319f3.html.
[27] *Study suggests spam-stopping tricks*, CNET News.com; refer to http://news.com.com/2100-1024-993333.html.
[28] *The Anti-Spam Cookbook*, Network Computing; refer to www.networkcomputing.com/1320/1319f3.html.
[29] *The Anti-Spam Cookbook*, Network Computing; refer to www.networkcomputing.com/1320/1319f3.html.
[30] *Living in "Spammed" Times*, Internet.com; refer to http://asia.internet.com/news/article.php/1436501.

Chapter 4

What can we do to minimize spam?

There is no completely effective solution to rid the world of spam, short of getting rid of email altogether. However, we do have some suggestions to reduce the amount of spam received. Educating users is one part of the solution, and using technology to secure messaging systems and filter spam is another. Legislation will help provide guidelines, but it won't have a big effect without heavy penalties and enforcement.

User education

Teach users what is and what is not spam. Any unsolicited email they receive from an organization with which they have no prior relationship is spam. Any email from a list to which they have subscribed, or promotional email they have agreed to accept in return for a service, is not spam.

Educate your users about what they can do to minimize the possibility of spammers getting their email addresses, both corporate and personal. Teach them not to post email addresses on web sites or newsgroups, or if they must, suggest they make their addresses "human readable"; for example, *leslie at jconsult dot com*.

When users sign up for an online service, remind them to examine privacy policies, terms, and conditions. They should also look for check boxes that indicate that the user agrees to allow the vendor to share their email address with other marketing partners.

Make users understand that it is useless to respond to spam; that in many cases "unsubscribe" links usually just confirm a live, monitored address. If the sender is a well-known organization with a good reputation, then unsubscribing should be a valid option.

If users need to share an email address with a marketer, suggest they create a "throwaway" account, and share their legitimate business and personal email addresses with those who will not abuse them.

Make users understand that if they never respond to spam, it makes it an ineffective way to advertise. If nobody ever bought anything a spammer marketed, they would all be out of business. Unfortunately, there will always be people out there who think that the "miracle product" will be

the one that finally works.

If users really want to order a product that was advertised using spam, suggest they go directly to the web site by typing the web address into the browser; i.e., not to click on a link contained in the message. In this way, the vendor does not know that the spam message worked. It is worth reminding them, however, that they are supporting organizations that hire spammers.

Do also inform them that nobody in Nigeria is really going to share $240,000,000 just for the use of a bank account. (The Nigerian "advance fee" scam is expected to gross $2 billion in 2003, according to MessageLabs.[31])

Corporate messaging policies

Develop a corporate messaging policy that includes a definition of spam. Enforce it.

A "no personal use of the corporate messaging system" policy is one way to prevent spammers from getting addresses that are shared carelessly, and personal messages can open up legal liability depending on content. Also, employees are not wasting company time and resources for personal business. However, this is extreme; email is a valuable perk that some companies want to provide to their employees.

Technology solutions

The first technological step to help prevent spam is to secure all wireless networks, proxy servers, and email relay servers, so that spammers must use their own resources to send email. This makes it easier to trace spammers back to their sources and complain to their ISPs or have them prosecuted in areas where the law allows. Spammers' accounts are more often closed because they contravened their ISP's Terms of Service than by any laws.

---

[31] *Spam and Virus "Blended Threats" Top Email Dangers in 2002*, Advisor; refer to http://marketingadvisor.net/doc/11611.

There are many different types of software that can detect and filter spam. These programs can be layered to make an effective system of spam filtering to maximize the number of spam messages blocked, while minimizing the number of false positives.

Spam can be filtered at the Mail Transport Agent (MTA) or gateway level, when it is delivered to the server, or when it reaches the desktop. The closer to the perimeter, the less work the corporate messaging system has to perform.

Ideal solutions examine messages before they are actually brought into the messaging system to save processing and storage requirements.

Server solutions can quarantine messages at the incoming SMTP server level, which prevents further processing by other servers.

By the time a spam message reaches the desktop, it invalidates most of the reasons why spam should be filtered – the message has been processed and stored by the messaging system already, and user action may still be required.

Even with a gateway or server solution, the ability to tune filters on an individual level can be useful, especially since spammers often purchase off-the-shelf spam filters to develop messages that will get through the filters.

Ben Littauer, an independent consultant from Boston, recommends the "silver shotgun" approach: "If enough people are using enough different spam filters, then the number of spam messages that get through is greatly diminished. If fewer messages get through, the response rate is reduced and thereby the total revenues for a fixed-size mailing are also reduced. If the response rate drops low enough, the mailing cost (though small) will eventually exceed the return and the spammer will stop. If everyone uses the same filtering approach, however, a spammer can "tune" the mailing to that filter and still get through. Thus, I recommend that organizations and individuals use more than a single spam filter, and consider standardization on a single filtering approach as helpful to the spammer."

Content Filtering

Content filtering examines messages to find phrases or patterns common to spam, or messages that match a spam signature database.

Filtering may take place at the email gateway, mailbox server, and/or at the email client.

Heuristic filtering is a statistical process that weights common spam phrases and characteristics and assigns a token or value to each one found in a message. If the combined weight of the tokens exceeds a set limit, the message is tagged as spam. Heuristics understand spam, even if it's a first-time spam message not in a spam database.

Messages tagged as spam through heuristics may automatically generate rules to identify future spam. These products basically learn as they are used, and can be very effective at catching most spam.

Many content filtering products use databases of spam definitions provided by the manufacturer. They may scan messages by subject lines or specific phrases. These signature files are highly useful for catching known spam.

One way spam databases are created is using "honey-pot" systems, set up using decoy email addresses posted on web sites or newsgroups. Any email delivered to one of these addresses is spam. Filters are created based on these messages.

These databases are delivered as updates to the clients. Some anti-spam packages make updates available every ten minutes, which is close to real-time.

Content filtering can also be useful to block viruses before anti-virus vendors release pattern files. It reduces the risk of data loss and data exposure, as well as lost business.

It also can eliminate pornography and hate mail, which reduces legal liability, minimizes workplace harassment, and diminishes inappropriate email usage.

Of course content filtering isn't that useful for foreign-language spam.

Content filtering has one major drawback – it cannot differentiate between solicited and unsolicited email based simply on the content of a message. Many content filtering packages include an option to allow email from specified senders to be added to a whitelist, allowing the messages to be delivered irrespective of the content.

Blocked messages should be examined, especially when a spam filter is first installed. Some programs use a quarantine folder, giving access to either a system administrator or the users. This means a human must spend time scanning the filtered email to make sure nothing legitimate has been misfiled.

Pattern Recognition and Header Analysis

Pattern recognition and header analysis can be used to filter messages with inconsistencies in headers – such as forged information – and envelope characteristics or patterns common to spam – such as delivery paths using multiple servers or a large number of recipients.

This type of filter is useful for poorly forged messages; however, it can cause delivery failures of legitimate messages if a company uses a circuitous route to send SMTP email, adding to the hop count and perhaps going over the limit of the spam filter.

Reverse DNS and Address Verification

Reverse Domain Name Service (DNS) lookups and address verification can catch inconsistencies in messages, but they use a large amount of resources and also run a high risk for false positives.

Not all companies have their outgoing MX server listed in their DNS, which could cause legitimate email to be filtered.

Real-Time Black-Hole Lists (RBLs)

Some spam filtering applications use real-time black-hole lists (RBLs) or blacklists, which are a fairly clumsy form of filtering, based on IP addresses. RBLs don't let spam servers connect to your email servers.

RBLs are maintained by various system administrators who hate spam.

RBLs use different criteria to add IP addresses to their lists. Some go by whether a server is an open relay or open proxy, others by whether the list administrator has received reports of spam being sent by a particular server or ISP, and others just by whomever the administrator is not happy with that day. Email system administrators who are considering using RBLs should fully understand what qualifies an entry to be added to the list before subscribing.

This type of filter runs a huge risk for false positives, as an ISP may play host to legitimate users as well as spammers. David Nelson, a senior industry analyst at Giga Information Group, says a recent study found that MAPS blocked 24% of spam with 34% false positives. This hurts all legitimate Internet email users.[32]

If you have open relays in your environment and they are discovered by spammers, you may be added to a blacklist. It is not always easy to be removed from a blacklist.

Domain Name Service-delivered Blocking List (or DNS-delivered Blocking List) is a type of black-hole list, using DNS rather than IP addresses.

Spammers sometimes use lists of servers with open relays to find more servers from which to relay their email.

Whitelists and Certification

Whitelists work on a similar principle to blacklists, except that servers on the list are sending messages that are certified not to be spam. The lists can contain individual email addresses, or lists of domains or ISPs with "no spam" policies.

Several consumer and marketing groups are developing whitelists based on their certification standards. These lists are made available, usually for a fee.

According to field tests of 40,000 consumers by a large consumer company, there was a 52% improvement in click-through rate per

[32] *The Spam Police*, Network World; refer to www.nwfusion.com/research/2001/0910feat.html.

delivered email for messages containing a Trusted Sender trust stamp, *versus* the same message without a stamp.[33]

Habeas is one company providing certification. Their Habeas Sender Warranted Email program inserts a haiku and special text headers in outbound email that has been certified by Habeas as "not spam". Anti-spam filters can be programmed to recognize the headers and allow the message to be delivered.

```
X-Habeas-SWE-1: winter into spring
X-Habeas-SWE-2: brightly anticipated
X-Habeas-SWE-3: like Habeas SWE(tm)
X-Habeas-SWE-4: Copyright 2002 Habeas(tm)
X-Habeas-SWE-5: Sender Warranted Email (SWE). The sender of
X-Habeas-SWE-6: this email in exchange for a license for this
X-Habeas-SWE-7: Habeas warrant mark warrants that this is a
X-Habeas-SWE-8: Habeas-compliant Message (HCM) and not spam.
X-Habeas-SWE-9: Please report use of this mark in spam to
                <http://www.habeas.com/report/>.
```

Habeas also maintains a DNS-based whitelist of IP addresses of Habeas bulk email licensees and other enterprise and individual licensees who use confirmed opt-in and meet all other standards for a Habeas-Compliant Message (HCM).

Greylists

Greylisting is a recently proposed method that filters spam at the Sendmail MTA level, so there is no network traffic other than the connection. Greylisting looks at the IP address of the host attempting delivery, the envelope sender address, and the envelope recipient address. If none of these three qualifiers is in the greylist database, the message is refused with a temporary failure. A properly configured SMTP server will retry after a specified time period – spamming software generally does not. Greylisting is not intended to replace other methods of spam filtering, but it is an addition that can reduce spam processing. In testing, this method blocked 97.4% of spam with no false positives.[34]

---

[33] *Spam by the Numbers*, ePrivacy Group; refer to http://cobb.com/spam/numbers.html.
[34] *The Next Step in the Spam Control War: Greylisting*, by Evan Harris; refer to http://projects.puremagic.com/greylisting/.

Do-Not-Spam Lists

The Federal Trade Commission (FTC) is developing a do-not-spam list for marketers. However, unethical spammers may not bother to take the time to clean up their lists, or even worse, use those lists as new addresses to spam. The FTC will most likely put honey-pot addresses on the list, and aggressively go after those who do spam them.

Some marketing firms such as the Direct Marketing Association (DMA) have similar lists.

Peer-to-Peer Reporting

In peer-to-peer reporting, or collaborative filtering efforts, users tag messages as spam, and the anti-spam program forwards a copy to a system administrator. These messages are added to the program's database and pushed to the other users of the package. A risk with peer-to-peer programs is that users may tag email as spam when the message was something they once subscribed to. This can be a final step in the message screening process, though it is more of a community service than any type of resource or time-saving measure for the organization.

Challenge/Response Systems

Some desktop users favor challenge/response systems, which give them much more control over their email. Any message received from a sender not listed on an "approved" list gets an automated reply requiring the sender to type a code shown in an attached graphic. This requires a human response. If the code is correct, the message is delivered.

This does require more work on the part of the legitimate sender to get their message delivered, and it will also result in delayed email if the sender doesn't see the challenge right away. Spammers don't generally even see these challenges, and for the small number of messages that will be blocked this way, they have another 999,999 addresses on their mailing list.

Technology Summary

There is no one technology solution, but a layered approach using several of these methods can be very effective at weeding out spam.

Legal options

The legal situation is very fluid. At this point many countries are developing enforceable financial and legal penalties for spammers, though a global standard will most likely be necessary to have a real effect.

European Legislation

The European Union's Privacy and Electronic Communications Directive came into force on 31 October 2003, and makes sending of unsolicited email to an individual illegal unless they specifically request it. This ban does not apply to existing customer relationships, so retailers may continue to send marketing messages to consumers, as long as they provide an opt-out feature. The definition of email is broad enough to also cover text-messaging systems such as mobile telephones. The EU member states must pass this regulation individually as part of their own national laws, which could take years.

The UK government is considering exempting business email accounts from this directive, which would make it legal for spammers to continue to send spam to business addresses. "Many people feel strongly that anti-spam measures could hamper business-to-business (B2B) commerce. Others feel equally strongly that unsolicited email is just as big a problem for businesses", Timms explained at the Spam Summit at the House of Commons.[35]

United States Legislation

In the US, there is no Federal legislation, although there are nine bills pending as of July 2003.

---

[35] *New laws will make it "legal" to spam your work address*, Silicon News; refer to www.silicon.com/news/165/1/4970.html.

At the present time, the only US Federal Government agency that is taking any action against spam is the Federal Trade Commission (FTC). However, the FTC is only taking action when email messages contain evidence of fraud or illegal activity. The FTC does not take action against ordinary unsolicited commercial email. The FTC receives about 130,000 forwarded spam messages a day at uce@ftc.gov, the agency's unsolicited commercial email mailbox.[36]

Many individual states have enacted anti-spam laws, though enforcement has been sporadic at best.

US laws will not have any effect on spammers who operate from other countries, and even those who operate in the US won't necessarily heed the laws as long as they feel they won't be punished seriously.

However, a recent Virginia law may make them reconsider. Certain types of spam can be punishable with up to five years in prison, and the forfeit of all profits earned from the deceptive solicitations, as well as all computer equipment, computer software, and all personal property used in connection with the illegal act.[37] This law doesn't apply only to email originating in Virginia, but any message that passes through any server in Virginia.

Most spammers that have been prosecuted in the US have not found themselves in trouble for sending spam, but rather for the methods they use to send spam. Howard Carmack, *aka* the "Buffalo Spammer", was arrested and arraigned in New York for credit card and identity theft, not for the 825 million unsolicited emails he allegedly sent.[38]

Spammers often use overseas servers in countries with no anti-spam legislation, such as China and Korea, so they don't have to worry about US laws, even though it raises their costs significantly.

---

[36] *FTC finally realizes that spammers lie*, ZD Net UK News; refer to http://news.zdnet.co.uk/story/0,,t269-s2134105,00.html.
[37] *Virginia Seeks Jail for Spammers*, Internet Advertising Report; refer to www.internetnews.com/IAR/article.php/2199001.
[38] *"Buffalo Spammer" Arrested*, InternetNews.com; refer to www.internetnews.com/xSP/article.php/2206311.

Legislation in Japan

The Japanese Government acted to cope with spam by passing the Specific Electronic Mail Law, 2002 #26, effective from 17 April 2002. This law was authorized at the 154th Assembly.[39]

Proposed Rules for Fair Spamming

Common suggestions for legislation include:

- Making it illegal to falsify the routing information of the email message
- Failing to honor "opt-out" requests, or failing to include opt-out instructions
- Failing to identify a message as an advertisement through use of the ADV, ADV-ADULT, or ADVERTISEMENT labels at the beginning of the subject line

Spammers' Rights Under the Law

Spammers raise the issue of censorship, claiming they have the right to send their messages under the First Amendment.

A group of spammers in Florida, Emarketers America, is suing anti-spam organizations SPEWS, The Spamhaus Project, and Joker.com, claiming they are destroying spammers' right to market via the Internet.

Self-Regulation of Legitimate Marketers

Marketing companies and organizations are working to develop their own guidelines and policies regarding what is and is not spam, perhaps hoping that if they regulate themselves, there will not be a need for Federal legislation.

The Email Service Provider Coalition (ESPC), formed by the Network Advertising Initiative, is a coalition of email service providers that plans

[39] Soumushou Research; refer to www.soumu.go.jp/joho_tsusin/top/m_mail.html (Japanese only).

to develop registries to certify legitimate email marketers from spammers. Marketers' performance would be rated as a way to remain on the registry, which the ESPC hopes companies will adopt as a whitelist. Marketers would be evaluated and given a score somewhat like a credit rating based on customer complaints, how many times people have to unsubscribe, and other factors.

There are disagreements between marketers and anti-spam activists about where the lines are to be drawn.

The Direct Marketing Association (DMA) is the oldest and largest marketing trade association in the US. The DMA's Commercial Solicitations Online Guidelines approved in January 2002 state that acceptable commercial solicitations are those sent to a marketer's own customers, or to individuals who have consented to receive solicitations online or have not opted out when offered the chance. Each solicitation should include a link to request removal from the marketer's mailing list, and a link to request that the email address not be shared with other marketing organizations for online solicitation if the marketer provides such a service. The DMA does provide an Email Preference Service suppression file for members to use to filter email lists.

While some anti-spam activists feel these guidelines are fair, many others do not agree that a prior business relationship is enough to justify sending marketing messages; they believe that if a customer hasn't specifically agreed to accept promotional email from an organization, it is spam.

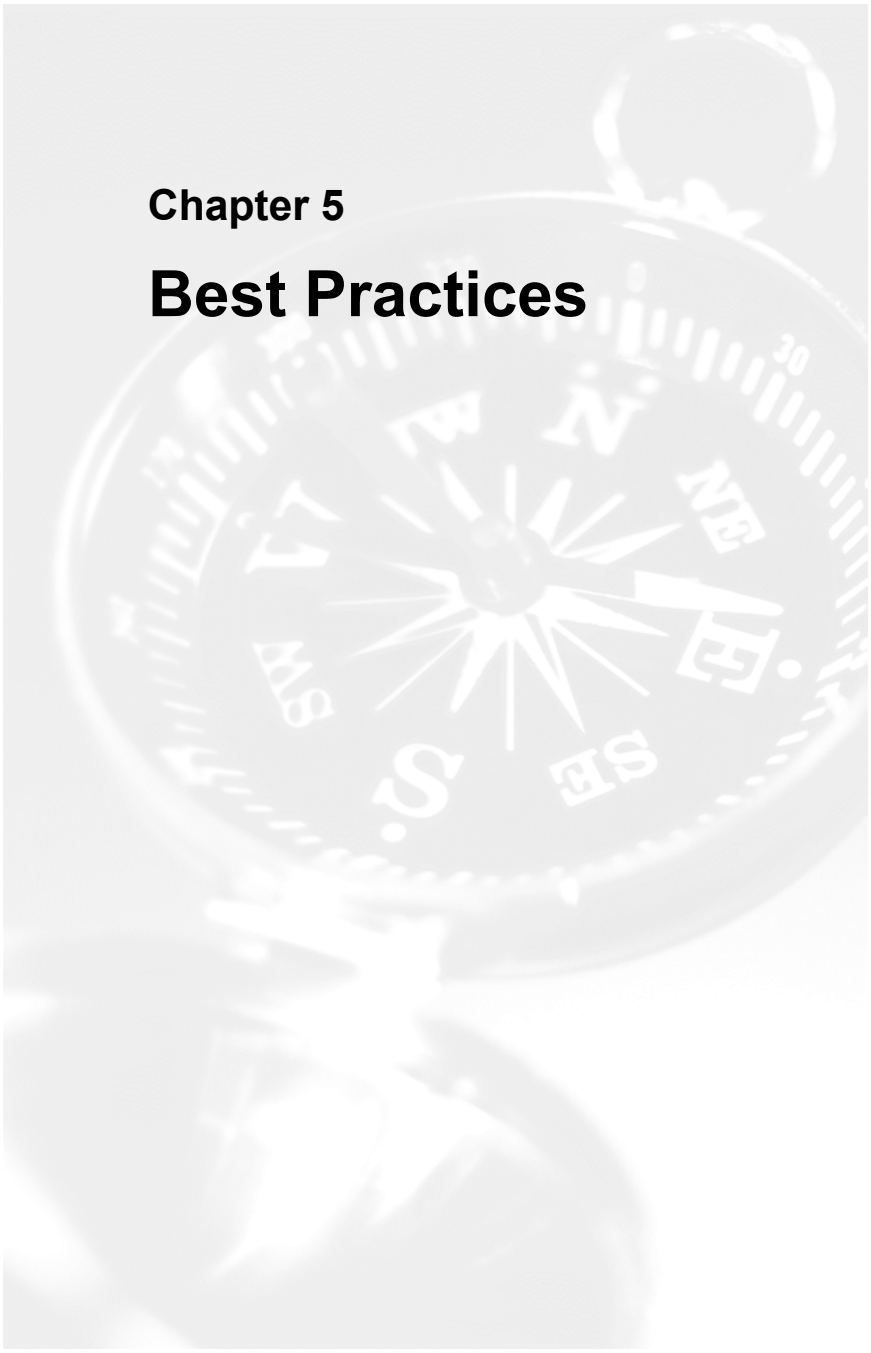
Many activists believe legitimate marketers should adopt an "opt-in" policy, where users should not be contacted unless they have requested to be added to a mailing list via an organization's web site, a contest entry blank, or some other medium. They believe it is not appropriate to put the onus on the recipient to opt out. Lack of response does not indicate consent. Additionally, an "opt-out" procedure confirms to the spammer that the recipient's email address is valid, thus opening the door to additional spam.

While self-regulation is one possible solution to reduce spam since legitimate marketing firms will abide by the guidelines, spammers will continue to abuse this medium until it becomes unprofitable. The best

way to get rid of spam is to discourage spammers from sending it. The bottom line here is money – if they don't earn enough from sending spam, or if they have to pay for their abuses, they'll be forced to find other ways to make money. For this reason, a "private right of action" would be a critical component of any new legislation.

Chapter 5

Best Practices



We recommend the following steps to help eliminate possible delivery routes for spam from corporate messaging systems:

- Secure SMTP servers by either disallowing any external relay, or requiring authentication.
- Verify that proxy servers are configured to only route traffic from the LAN to the Internet.

We recommend the following steps to minimize the amount of spam received by individual organizations:

- Use a content filter that subscribes to a regularly updated database of spam. Quarantine all filtered messages at first, and examine them to fine-tune the system to avoid false positives.
- To block directory harvests, do not post member or employee databases on corporate web sites.

The ideal spam filter

The ideal spam filter will have the following features:

- Easy installation and administration
- High level of flexibility and customization
- High filtering rate – 90% or higher
- Extremely low false positive rate
- Filtering at the perimeter of the messaging system
- Little impact on email delivery times or server performance
- Automatically updated database of confirmed spam
- Rules and heuristic scoring system for new spam
- Ability to teach itself what is and is not spam
- Customized whitelists
- Access to quarantine folder for either system administrator or users
- Integration with anti-virus filtering

Make sure you understand the filtering criteria used, and tune it to an appropriate level for your organization.

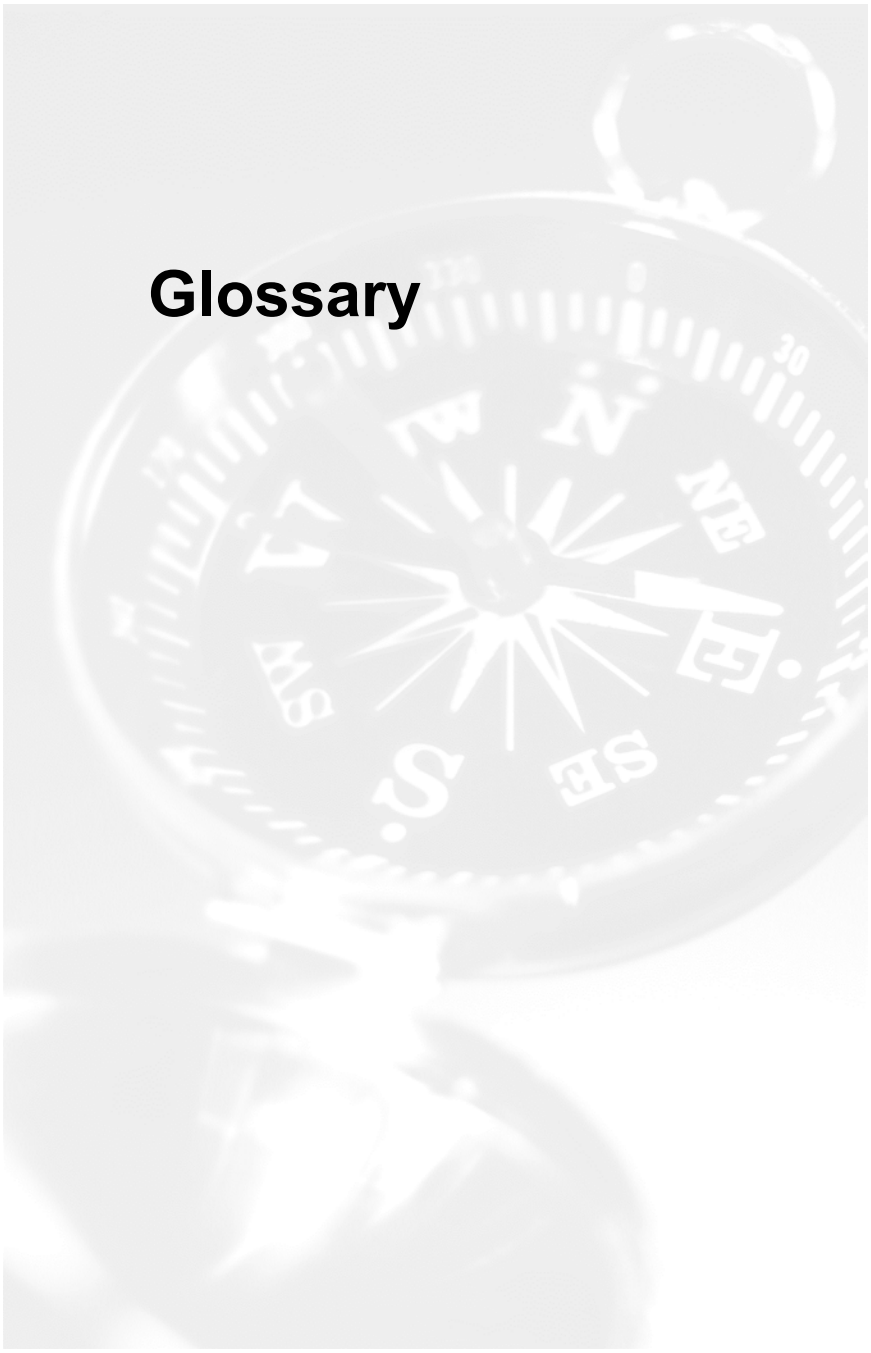
Chapter 6

Summary

To continue to leverage our investment in email, we need to minimize the amount of spam that gets transmitted and received through messaging systems. If less spam is delivered, the spammer's potential for profit is reduced. Spammers will not stop sending spam until they stop making money.

The measures required to fight spam can be quite frustrating – there is no simple solution, but there is some relief in sight from better filtering technologies, certification systems under development, and potential legal remedies applied on a global scale.

Glossary



Address Harvesting: Spammers send "spiders" or agents to scour websites, mailing lists, and newsgroups for addresses.

Bayesian Filtering: A statistical process that weights each word or phrase in a message and assigns a token. If the combined weight of the tokens exceeds a set limit, the message is tagged as spam.

Blacklists or Blocklists: Lists of IP addresses and domain names of confirmed spammers (or the ISPs who host them). Messages from senders on these lists are blocked.

Bulk Email: Messages sent to multiple recipients, usually commercial in nature.

CAN-SPAM: Controlling the Assault of Non-Solicited Pornography and Marketing Act (CAN-SPAM) became United States law on December 15, 2003. The bill requires that spam emails be labelled as such, and include the sender's physical address and instructions for how to opt-out. It also bans deceptive subject lines and false headers.

Chain Letters: Usually nonsense emails encouraging the recipient to forward the message to all users in their address book.

Challenge/Response Systems: Available for desktop users. Any message received from a sender not listed on an "approved" list gets an automated reply requiring the sender to type a code shown in an attached graphic.

Collaborative Filtering or Peer-to-Peer Reporting: Users tag messages as spam, and the anti-spam program forwards a copy to a system administrator.

Content Filtering: Searches for phrases common to spam, and identifies spam by content.

Denial of Service (DOS) Attack: When spammers use unsecured email servers to relay their messages, they can cause delays or disruptions in email service – the server is too busy processing spam to deal with regular email. Alternatively, a spammer may flood an email server to maliciously disrupt service.

Directory Harvest Attacks: Email servers are overwhelmed, as spammers send variations of email addresses to a domain email server. Addresses that don not fail are considered valid and added to their mailing lists.

Do-Not-Spam Registry: The Federal Trade Commission (FTC) is studying the feasibility of setting up a do-not-spam registry modeled after the national do-not-call list of people who do not want to receive telephone solicitations.

Drive-By Spamming: Spammers drive around and find unsecured wireless networks. Sitting in a van with a laptop, spammers can send email from "inside" the network to the SMTP server. Any messages sent by the spammer would appear to come from within the organization's network.

Email: Electronic messages sent from one computerized device to another.

False Positive: A message that has the characteristics of spam, but is actually a legitimate, solicited, or welcome message.

Filter Buster: A string of nonsense characters added to the subjects and bodies of their messages in the hope of confusing filters that look for known spam messages. Many spammers actually buy filtering software to test and fine-tune their messages to get around the filters.

Forged Headers: Spammers often forge or tamper with the headers of email messages to conceal their identity and location.

Greylisting: A method that filters spam at the Sendmail MTA level, so there is no network traffic other than the connection. A properly configured SMTP server will retry failed messages after a specified time period – spamming software generally does not. Greylisting is not intended to replace other methods of spam filtering, but it is an addition that can reduce spam processing.

Header: Part of an email message that contains the subject and routing information. Headers are useful in tracking the true source of a message, which is why spammers may forge them to hide their identity and location.

Header Analysis or Pattern Recognition: Can be used to filter messages with inconsistencies in headers (such as forged information) and envelope characteristics or patterns common to spam (such as delivery paths using multiple servers or a large number of recipients).

Heuristic Filtering: A statistical process that weights common spam phrases and characteristics and assigns a token or value to each one found in a message. If the combined weight of the tokens exceeds a set limit, the message is tagged as spam.

Honey Pot: A decoy email address set up to capture and identify spam. Filters are created based on these messages.

Munging: Editing an email address so that it is invalid to automated processes but understandable to humans; for example, leslieNOSPAM@jconsult.com or leslie at jconsult dot com.

Joe-joe: A maliciously forged header, making another person or domain appear responsible for a spam message.

Murk (aka Murkogram): A disclaimer in spam attempting to give the message an air of legitimacy, stating it complies with Bill S.1618. There is no such bill, so messages referencing it can be automatically classified as spam. The term comes from Frank Murkowski (R-AK), the senator who wrote Bill S.1618. This would have made certain types of spam illegal, unless the message included full contact information at the start and made no attempt at hiding its origin.

Nigerian 419 Scam: An advance fee scam where the recipient is promised a share of unclaimed wealth if they provide their bank account information and perhaps cover a "transfer tax". These scams originated in Nigeria, and were actually expected to gross $2 billion in 2003, according to MessageLabs.

Open Proxy Servers: Spammers exploit internal SMTP relays through insecure or misconfigured proxy servers, normally set up for users within a network to control Internet access or to cache data for frequently used web sites.

Open Relay: An email server that permits relaying of email by anyone, allowing spammers and others to send messages.

Opting-In: Requesting or otherwise agreeing to receive email from a list or other email source.

Opting-Out: The process of requesting to be removed from a mailing list. However, if it's a spammer's mailing list, an opt-out request is likely to be taken as confirmation of a live email address.

Peer-to-Peer Reporting: See Collaborative Filtering.

Realtime Blackhole Lists (RBLs): See Blacklists.

Reverse DNS Lookups: Match the sending IP address with the domain of the sender before a message will be accepted. It uses a lot of resources and also runs a risk for false positives if MX records are not configured correctly.

Signature Files: Many content filtering products use databases of spam definitions provided by the manufacturer. They may scan messages by subject lines or specific phrases.

Spam: Marketing emails and other unsolicited messages. A simple definition of spam is unsolicited email messages, generally commercial or promotional in nature, usually sent in bulk.

Spambots: Robots that scan webpages to extract email addresses, which are then used as targets for spam.

Spammer: A marketer or mass mailer who generally does not act ethically; i.e., sends messages to people who have not expressed interest, uses others' unsecured email relay servers, or sends fraudulent offers.

Spamware: Software used by spammers to harvest or deliver spam messages.

Spoofing: Using a false return address.

Throwaway Account: An email account created by a user for short-term use to avoid spammers. Alternatively, an account created by a spammer for short-term use to send spam messages.

Unsolicited Commercial Email (UCE): See Spam.

Web Beacons: A way to verify active email addresses. If an HTML message is downloaded and rendered, even in preview mode, a web beacon tells the sender that the email address is being used.

Whitelists: Similar principle to Blacklists, except that only servers on the list are permitted to send messages to your server. The lists can contain individual email addresses or domains of business contacts, or lists of domains or ISPs with “no-spam” policies.

About the Authors

Leslie Ogonowski

Leslie Ogonowski is a messaging and anti-virus consultant with Johnson Consulting. She received her first spam in 1998 and has been seeking revenge ever since.

The Open Group Messaging Forum

The Open Group Messaging Forum is a leading association for the ebusiness and messaging industries. The Forum comprises customers, suppliers, and consultants. Its diverse membership focuses on providing interoperable solutions for business leaders through education, fulfilling customer-driven requirements, promotion and endorsement of standards-based solutions, and influencing public policy.

index

abusive practices 6
abusive spam 6
address harvesting 7
address verification 23
anti-harassment policies 13
anti-spam laws.....28
anti-spam policy 13
anti-spam solutions.....14
blacklists.....23
blocking spam 13
Buffalo Spammer 28
censorship.....29
certification 24
challenge/response systems.26
commission 6
content filtering 13, 22
corporate messaging policy.20
denial of service 8
directory harvest.....7
DMA 26, 30
DNS.....24
DNS-delivered blocking list24
domain email server 7
do-not-spam lists 26
drive-by spamming.....8
education 19
email.....2
email address lists
  purchase of 7
Emarketers.....29
ESPC 29
fair spamming.....29
fake return address 9
false positive.....13
filter busters.....10
filter software.....21
First Amendment 29
foreign-language spam.....22
forged headers.....9
fraud.....6
FTC 28
greylists.....25
Habeas.....25
harvesting addresses.....7
HCM 25
header analysis.....23
heuristic filtering.....22
honey-pot systems.....22
illegal content.....6
immoral content 6
insecure email 2
insecure email servers 8
instant messages.....4
IP address.....23
ISP.....20
Joe-jobs 9
Joker.com.....29
legal options.....27
legislation.....2, 19
  EU 27
  Japan.....29
  US.....27
mail transport agent.....21
MAPS.....24
marketing 2
marketing services 5
messaging policy.....4
MTA.....21
Murkowski.....9
newsgroups 19

nuisance spam 5
open email relays.....8
open proxy servers 8
opt-in policy 30
opting out 7
pattern recognition 23
peer-to-peer reporting.....26
personal privacy 14
promotional messages 5
proxy servers 8
RBL.....23
real-time black-hole lists 23
recommendations 33
reverse DNS 23
self-regulation 29
signature files 22
Sobig.a.....9
spam 2
  cost of.....14
  examples.....4
  increasing traffic.....16
  types of.....5
spam database 22
spam filter 33
Spamhaus.....29
spammers 6
spammers' legal rights 29
spiders 7
technology solutions 20
text-messaging 27
The Open Group iv
Trojan executable.....9
unethical content.....6
unsecured wireless networks.8
unsolicited email 4
unsolicited email 2
unsubscribing.....19
virus 9
web beacons.....8
whitelist.....23, 24
wireless messages 4