Global Information Assurance Certification Paper

SPAM - How to survive in 2003

GSEC Practical Assignment (v.1.4b)

Ingo Quilling
April 17, 2003

TABLE OF CONTENTS

ABSTRACT

Today spam is a threat to the survival of the internet. It floods ISPs companies and end users and causes huge costs. Spam has to be fought.

This paper wants to help end users and companies in reducing the pain spam causes.

INTRODUCTION

Definition

Spam: “To mass-mail unrequested identical or nearly-identical email messages, particularly those containing advertising. Especially used when the mail addresses have been culled from network traffic or databases without the consent of the recipients.”[1]
The term „spam“ is commonly used for **U**nsolicited **B**ulk **E**mail (UBE) or **U**nsolicited **C**ommercial **E**mail (UCE).

Origin

Some people say that the word spam is originally adopted from a sketch from Monty Python’s Flying Circus.[2] Others say it is from an event happened around 1985 of typing “SPAM SPAM SPAM SPAM SPAM SPAM…” by a keyboard macro in the MUD user groups.[3]
The first spam email was send by a DEC-company sales representative in 1978.[4]

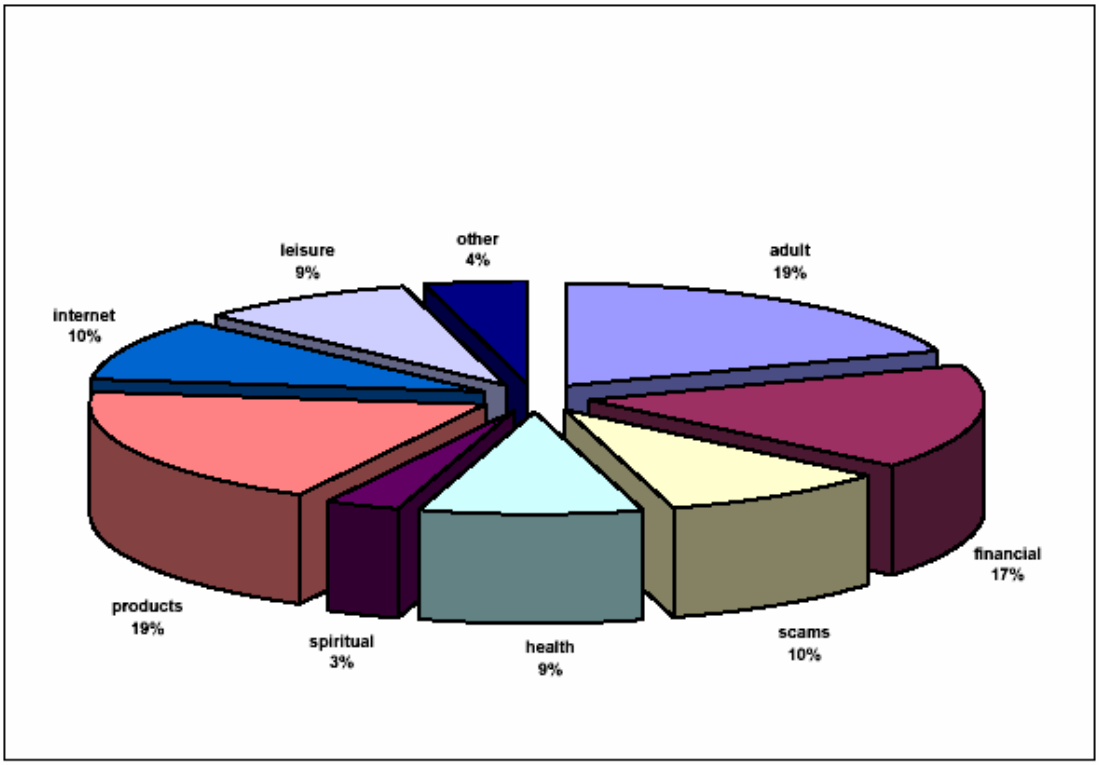
Characteristics

The nature of spam is that it causes very little costs for the sender - compared to traditional postage. Spammer use other people’s resources and shifting costs away from the sender.

According to CAUCE[5] spam usually contents one or more of the following themes:

- Chain letters
- Pyramid schemes (including Multilevel Marketing, or MLM)
- Other "Get Rich Quick" or "Make Money Fast" (MMF) schemes
- Offers of phone sex lines and ads for pornographic web sites
- Offers of software for collecting e-mail addresses and sending UCE
- Offers of bulk e-mailing services for sending UCE
- Stock offerings for unknown start-up corporations
- Quack health products and remedies
- Illegally pirated software ("Warez")

See the tables on the next page to realize the content of spam and the frequency of categories.



March 2003 Spam Category Data[6]
(as measured by Brightmail's Probe Network)

Category	Description	Types (e.g.)	Subject Lines (e.g)
ADULT	Email attacks containing or referring to products or services intended for persons above the age of 18, often offensive or inappropriate.	Porn, Personal Ads, Relationship advice	"Enjoy the fantasy or regret the reality" "Instantaneously Attract Women" "Watch Me Free On Webcam." "Looking for Love, Romance or Friendship?"
FINANCIAL	Email attacks that contain references or offers related to money, the stock market or other financial "opportunities."	Investments Credit reports Real estate Loans	"Have you checked your personal credit reports recently" "Get Cash, No Equity Required! Apply Today?" "We WILL Help You To Succeed With Wealth Builders"
PRODUCTS	Email attacks offering or advertising general goods and services	Devices Investigation services Clothing Makeup	"Flat Rate Long Distance" "Winter Clearance Leather Blowout Sale" "Buy 1, get 2 FREE Inkjet Cartridges" "Get your Free satellite TV system now"
INTERNET	Email attacks specifically offering or advertising Internet or computer-related goods and services	Webhosting Web design Spamware	"High Speed Internet Access is Here!" "Get 80,000 Direct Email Leads A Month!!"
SPIRITUAL	Email attacks with information pertaining to religious or spiritual evangelization and/or services	Psychics Astrology Organized religion outreach	"Free Psychic Readings" "Get 72 hours of Unlimited LIVE Psychic Readings" "The Truth About Your Power"
SCAMS	Email attacks recognized as fraudulent, intentionally misguiding, or known to result in fraudulent activity on the part of the sender	Nigerian investment Pyramid Schemes Chain letters	"Urgent Business Letter" (Nigerian investment scam) "Money in the Mail" (Be a millionaire in 1 year scam) "University Diplomas (Fake diplomas from non-accredited universities)
LEISURE	Email attacks offering or advertising prizes, awards, or discounted leisure activities	Vacation offers Online casinos Games	"Update Now for Your Florida Vacation!" "Collect $150 Free Chips!" "Free Cruise Certificate#243694" "Free CASINO hotel rooms & Money to Play with!"
HEALTH	Email attacks offering or advertising health-related products and services	Pharmaceuticals Medical treatments Herbal rememdies	"Online Consultation and Prescription Ordering" "Turn back your body's biological clock with ABC Oral Spray" "Breakthrough Skin Resurfacing Facial..."
OTHER	Emails attacks not pertaining to any other category.		

Brightmail spam categories (March 2003)[7]

SPAM TECHNIQUES

Spammer usually use tools called spamware. There are two categories of spamware: *pull* tools, which search for e-mail addresses, and *push* tools, which send bulk mailings.
Pull tools operate automatically by navigating websites and public spaces on Usenet. The software collects the email addresses found on defined web pages and newsgroups.
A push tool is software that sends bulk e-mails without going through a specific mail server or a particular ISP. It manipulates message headers to break through the mail servers' anti-spam filters. Often these tools generate randomly email addresses by composing words or letters and adding an existing domain.
Other ways to push spam to the recipients is to use open mail relays.

IDENTIFICATION

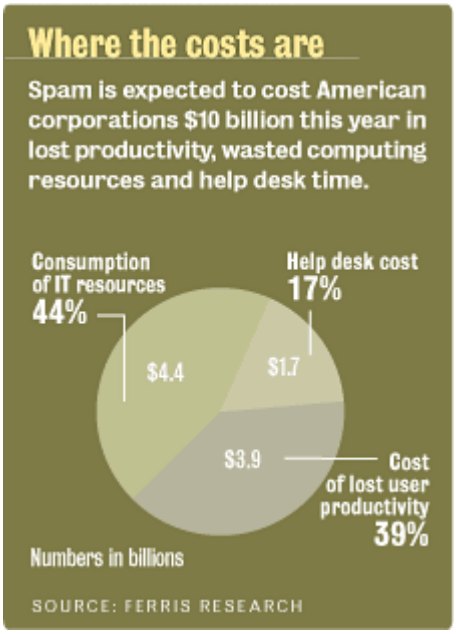
Spam has turned form an annoyance to a threat. Today we are confronted with a flood of spam. Spam causes tremendous costs, concerning the following parties:

Concerned party	**Resulting effects**
ISPs	• Support costs • Traffic costs • CPU time costs • Bandwidth consumption
ESPs	• Support / cleanup costs • Traffic costs • CPU time costs • Storage costs • Bandwidth consumption
Companies	• Loss of worker productivity • Support costs • Traffic costs • CPU time costs • Storage costs • Bandwidth consumption
End users	• Loss of productivity • Traffic costs • Storage costs • Bandwidth consumption

The content can also be a threat to the recipient. Spam is send by email and emails can content malicious code like undesirable dialers and of course viruses and worms. Additionally spam mostly is advertisement and leads to internet resources for porn, drugs, weapons and other content, which children or young adults should be guarded against.

Regarding the sheer quantity, spam can often be considered as a denial-of-service attack. It takes time to clean the mailboxes from spam to find the real

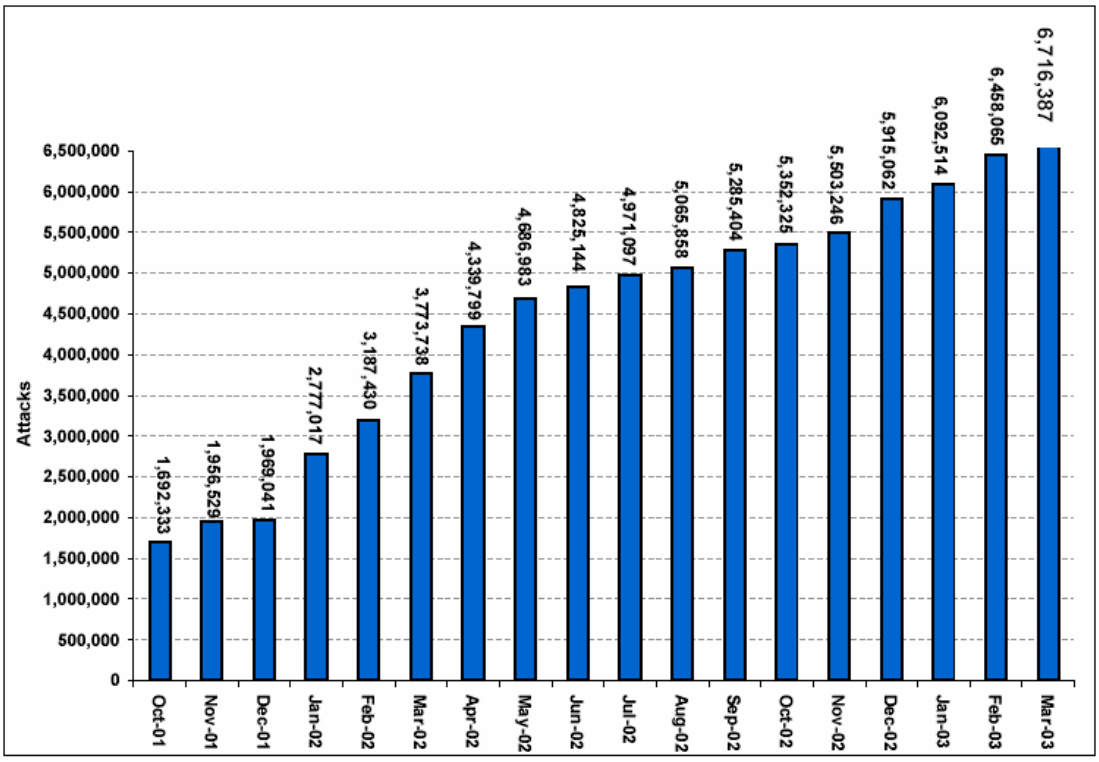
important messages. To separate spam from desirable content it is often necessary to read all the spam emails.



The Cost of Spam[8]

The cost of spam is measurable. "Spam costs corporate US over $8.9 billion in 2002 and will rise to over $10 billion in 2003" says Ferris Research.[9]

In the last 18 months the volume of spam has increased by the factor of 4 form 1.7 Million to 6.7 Million "attacks".



Unique Spam Attacks October 2001 to March 2003[10]
(as measured by Brightmail's probe Network)

A European Commission study suggests that an estimated 10 billion Euro per year will be wasted by Internet users around the world in dealing with junk emails or spam.[11] The study comes to a future scenario, where 20 Billion emails are transmitted by email marketer daily. If this scenario comes true, spam will threaten the survival of the internet.

CONSEQUENCES

We have to fight against spam.
Companies: Close your open relays. With open relays companies are in real danger. They lose their reputation. They will be charged for the traffic and possibly will be disconnected from the internet by their ISPs.

PREVENTION

Fundamentals

It is a pleasant behavior to use the Blind Carbon Copy (BCC) function in mail clients, if an email is sent to multiple recipients which do not know each other. This guarantees the privacy of each addressee.

If spammed, it is useless to reply to spam emails. On the contrary it gives spammers the certainty that the email address is really existing and valid. Also if the spam can be identified by the subject line or the senders address, it is important to not open it. The reason is that spammers often use HTML emails with possibly hidden pictures which have a unique address on a web server. By checking the web server's protocol the spammer is able to verify the email address.

Avoid online surveys, contests and price competitions. They are an important source to collect email addresses.

End users

End users may avoid to be spammed by carefully using their email accounts. It is recommended to give one's email address only to people one knows and trusts. But this reduces the value of email near to zero, because if nobody knows the address, no one can use it.
Also a long non conformal email address will help to avoid being found by robots. Robots compose email addresses in combining letters and / or words. For example with an address like JimSmith@email.com one can be sure to be a spam victim. Especially if this email account is provided at a provider like aol.com, hotmail.com, etc. These providers are a common target for spammer. Using an address like Jim_Smith_from_New_York_98716165@email.com will probably never be spammed automatically. Do not use your valuable email address in Chat rooms and Newsgroups. These internet resources are scanned for email addresses persistently and one can be sure to get a lot of UBE. Instead use an address like Jim_Smith_at_email_dot_com@Read_this.No_Spam. People (hopefully) can read this, but robots will have difficulties.
It is a good idea to use multiple email addresses, i.e. hosted by a free mailer. If this mailbox is flooded with spam, it can be replaced by a new one.

Companies

Companies are in another position. They have email addresses that cannot be changed frequently. They may have to publish their email addresses to get in contact with customers and partners. Here is another approach necessary:

Train the personnel.
It is important to train the employees to make them aware of the risks of using email addresses carelessly. Important rules are not to use a company email address for private emails, never to use a company email address for news-groups and chats.

Use tools like scripting to compose the addresses and write the addresses in form of pictures like JPEG or GIF. Another possibility is to use web forms to transmit messages.

Use spam filters.
All these steps above cannot avoid being spammed. Companies have to use spam filtering software

LEGAL CONSIDERATIONS

Criminalization of spam

Some countries and several U.S. States made the attempt to and are still developing laws, which criminalize spam.

A directive of the European Parliament says:
Safeguards should be provided for subscribers against intrusion of their privacy by unsolicited communications for direct marketing purposes in particular by means of … e-mails... These forms of unsolicited commercial communications may on the one hand be relatively easy and cheap to send and on the other may impose a burden and/or cost on the recipient. Moreover, in some cases their volume may also cause difficulties for electronic communications networks and terminal equipment. For such forms of unsolicited communications for direct marketing, it is justified to require that prior explicit consent of the recipients is obtained before such communications are addressed to them. The single market requires a harmonized approach to ensure simple, Community-wide rules for businesses and users.[12]
This directive has to be transformed in local laws in every country of the EU.

8 U.S. States, including Connecticut, Delaware, Illinois, Oklahoma, Rhode Island, Virginia, and West Virginia, have passed legislation that has made the sale or distribution of stealth spamware applications illegal.[13]

And more often we can see press announcements like “AOL Wins $7 Million From Spammers”.[14]

But, as far as there are no world wide agreements about the criminalization of spam it is impossible to catch and punish spammer. They will switch to other countries where spamming is not prosecuted and continue their bad business.

ANTI SPAM SOFTWARE

It is important to be aware that software will never find and filter 100 per cent spam. It is like natural evolution. Spammer and their programmers will always find ways to fool filtering software.

ISPs often offer email filtering. Email filtering at the ISP level prevents from heavy traffic form the ISP to the user / company's mail server caused by spam mail. On the other hand it will lead to the danger of loosing solicited and important emails.

In house anti spam software is easier to tune, because the complete control over all emails lies inside. The drawback of this solution is that it does not save any traffic.

As an example I will introduce Symantec AntiVirus for SMTP Gateways 3.1, software with anti spam functionality. I will show that a combination of techniques will help to manage spam. But it is not a "click and forget" solution. Rules and filters have to be adjusted permanently.

Example: Symantec AntiVirus for SMTP Gateways 3.1

Spam Detection:

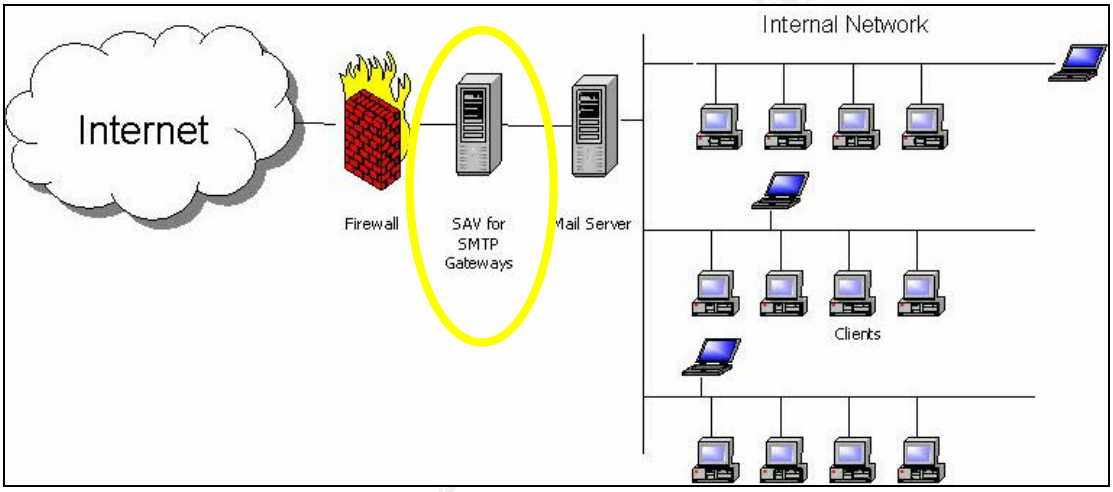
- Support for Multiple Real-time Blacklists
- Heuristic Antispam Engine
- Custom Blacklists (v3.0)
- Blocking by relay syntax

False Positive Prevention:

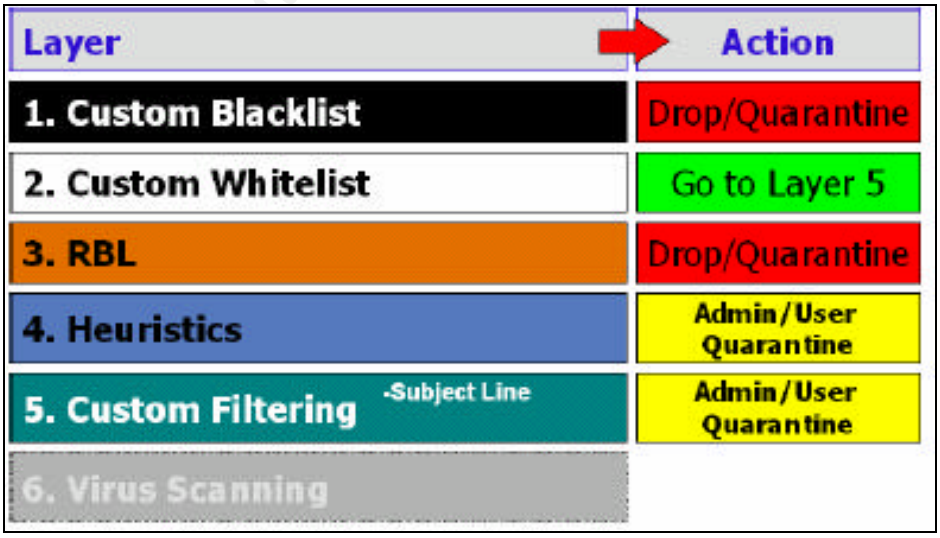
- Custom Whitelist

False Positive Management:

- Subject Line tagging – allowing administrator to filter at the gateway or client levels.
- "Quarantine" means "forwarding" to special administrative email account



Spam mail filtering



Symantec's spam rule order

Detecting Spam

An effective first layer against spam is **DNSBL-based** (**Real-time) Blacklists**. An administrator can use up to three different blacklist services (there are roughly 125 services available on the Internet, most are free, with the exception of MAPS). Since various services categorize different source addresses, using a combination of lists will minimize the total number of spam messages received from various sources. The databases or lists are DNS based which means that they can be queried from SAV SMTP via simple DNS query through the product. The list can either be accessed locally on the network or remotely across the Internet. The list can be replicated (via Zone Transfer) on a scheduled basis, through DNS (This process is external to the product itself). To avoid resource expensive remote lookups, replicating the database locally (Zone Transfer) is highly recommended.

Email identified as "blacklisted" can be dropped at this level, as a response can be returned to the sending server, which would identify the reason for the "refusal". In other words, email would not be delivered, but it wouldn't be lost. Should legitimate email be rejected, the sender's mail administrator can either address the issue for the blacklisting (i.e. improper mail server configuration) or they can work with the intended recipient's mail administrator to have their domain added to a whitelist, which would the bypass an blacklist check in future (see False Positive Management below).

Another layer of eliminating unwanted email is the **custom blacklist** (blocking by sender). This list allows custom entries for fully qualified addresses, as well as top (e.g. .com) and 2nd level (e.g. mydomain.com) domains. Blocking by top-level domain will be an increasingly important tool as spammers use resources and top-level domains from countries where spam is unregulated. Second level domains can be used to stop mail coming from domains that have not been blacklisted, but are common sources of spam or otherwise unwanted email. Similarly, using fully qualified email addresses is helpful if the domain cannot be effectively blacklisted, without losing wanted email. An example might be unwanted email emanating from a specific Hotmail account. Spammer@hotmail.com could be used to block emails specific to that one account, without blocking all Hotmail emails. At this blacklisting level, email can be dropped, without review.

A third spam detection layer is the **heuristic antispam engine**, which tags messages if a certain threshold is reached by the neural networks based scanning engine. The "sensitivity" of the heuristics engine can be adjusted to maximize detections and minimize false positives. The sensitivity threshold can be set from 1 (low) to 5 (high); where 1 will minimize false positives (and detections), and 5 will maximize detections (and false positives).

When tagging messages, SAV SMTP 3.1 places custom text in front of the subject line of the message. It's important to note that the heuristics engine does not delete or drop messages based on the results of the analysis. This decision is made at the subject line blocking level. The default text is "Spam:", but this can easily be changed to something more subtle, such as "Bulk:" or "UBE:"

A simple telltale sign of spam is the use of special strings in the recipient field that use characters like “!” and “%” as relaying instructions (these are respectively referred to as the “bang path” and “percent hack”). These strings are used in the recipient line to force a closed relay to openly relay to an external mail server, providing it can interpret the syntax. **Blocking by characters** in the email address allows the administrator to reject emails that use these relay strings.

Finally, **subject line filtering** can also be used to eliminate spam using common keywords or phrases. This can be used to address any missed detections (of spam) that may arise. In many cases (depending on the line of business), common words like “Mortgage” can be clear indicators of spam content. To more flexibly match subject lines containing keywords, wildcards are supported.

Preventing False Positives

The **custom whitelist** feature can be used to both prevent unnecessary false positives, as well as to remediate false positives. For example, since spam content and email newsletters often share characteristics not typical to standard personal or business email, i.e. multiple links, text advertisements etc., newsletters are some of the most common false positive detections. To resolve these misdetections, administrators can add domains for email newsletters (which, unlike spam, typically remain constant over time); to the whitelist. Adding a domain to a whitelist will exempt it from being scanned in future by either the RBL or the heuristic antispam layers.

The whitelist can also be populated proactively with customer and partner domains to ensure that standard business email communications are delivered without unnecessary delay.

Finally, in the case of RBLs, because legitimate mail servers can be improperly configured as open relays and consequently blacklisted, the whitelist can serve as remediation in these cases. As the sending customer’s mail administrator works with the blacklisting service to have their mail server removed from the blacklist (for example, after an “open relay” configuration is changed), adding the domain to the whitelist will allow the email to be delivered in the interim.

Since some blacklist services will blacklist entire ISP domains because they feel the ISP is “spammer friendly”, whitelists allows the administrator to exempt specific ISP domains, but still leverage the other entries on the blacklist. Some blacklist services like Spews.org will list the most common originating domains for spammers on their website, which will also give the administrator insight on which domains to exempt using the whitelist.

Managing False Positives (and tagged spam Messages)

Since the heuristics engine only tags messages as being spam, a decision needs to be made on how tagged messages are handled, i.e. either locally in the user's email client or centrally in an administrative account (holding area).

For administrator side management, a filtering rule can be created in the Subject Line blocking feature based on the spam tag, e.g. "Spam:*" (note the wildcard). The option "forward messages" to a specified email address, e.g. bulk@mydomain.com, should be chosen as well.

For user side management, a simple filtering rule can be created in the mail client to place all tagged messages in a special "bulk" folder for review. The filtering rule would be based on the tag selected by the administrator, e.g. "Spam:*" (note the use of wildcard). A process should be put in place to collect any false positives reported by users. This could be as simple as setting up a special administrative email account such as notspam@mydomain.com, where users could forward false detections. From this account, each case can be analyzed and domains can be added to a whitelist to prevent false positives from these sources in future. Again, the majority of false positives will typically revolve around newsletters or "solicited" mass e-mailings, which may look like spam, but which are not. The task of whitelisting these domains should decrease over time as the system is "taught" more about your email environment.

CONCLUSION

Until world wide legal regulations are accepted and enforced, the only way to fight against spam is to do it yourself or to use a service. Several technical tools and solutions are available. But technical measures alone will never solve the problem. Training and adequate behavior is an important element in the battle against the threat of spam.

ABBREVATIONS

CAUCE	Coalition Against Unsolicited Commercial Email
DEC	Digital Equipment Corp
DSN	Domain Name Service
DSNBL	Domain Name Service Blackhole List, see RBL
ESP	Email Service Provider
EU	European Union
ISP	Internet Service Provider
MAPS	Mail Abuse Prevention System
MUD	Multi User Dungeons
RBL	Realtime Blackhole List
SAV	Symantec AntiVirus
SMTP	Simple Mail Transfer Protocol
UBE	Unsolicited Bulk
UCE	Unsolicited Commercial Email

[1] The Jargon Dictionary. Spam, 5 URL: http://info.astrian.net/jargon/terms/s/spam.html (19 Apr. 2003)
[2] "The Infamous Monty Python Spam Skit!" URL: http://www.detritus.org/spam/skit.html (20 Apr. 2003)
[3] Southwick, Scott and Falk, J.D. "The Net Abuse FAQ", 23 Dec. 1998, URL: http://www.cybernothing.org/faqs/net-abuse-faq.html#2.4 (20 Apr. 2003)
[4] Templeton, Brad. "Reaction to the DEC Spam of 1978". URL: http://www.templetons.com/brad/spamreact.html (20 Apr. 2003)
[5] CAUCE, The Problem. URL: http://www.cauce.org/about/problem.shtml (20 Apr. 2003)
[6] Brightmail. "Spam Categories", Page 1. URL http://www.brightmail.com/pdfs/0303_spam_definitions.pdf (20 Apr. 2003)
[7] Brightmail. "Spam Categories", Page 2. URL http://www.brightmail.com/pdfs/0303_spam_definitions.pdf (20 Apr. 2003)
[8] Fontana, John, „Spam: New year, same old story". 27 Jan. 2003. URL: http://www.nwfusion.com/news/2003/0127spam.html (20 Apr. 2003)
[9] Morrissey, Brian. "Spam Cost Corporate America $9B in 2002" The Big Picture. URL: http://cyberatlas.internet.com/big_picture/applications/article/0,,1301_1565721,00.html (19 Apr. 2003
[10] Brightmail, Unique Spam Attacks URL: http://www.brightmail.com/pdfs/0303_spam_attacks.pdf (20 Apr. 2003))
[11] COMMISSION OF THE EUROPEAN COMMUNITIES. Unsolicited Commercial Communications and Data Protection - Summary of Study findings - January 2001 URL: http://europa.eu.int/comm/internal_market/privacy/docs/studies/spamsum_en.pdf (20 Apr. 2003)
[12] Official Journal of the European Communities. "DIRECTIVE 2002/58/EC OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 12 July 2002, concerning the processing of personal data and the protection of privacy in the electronic communications sector (Directive on privacy and electronic communications), Article 40" 31 Jul. 2002 URL: http://mineco.fgov.be/internet_observatory/pdf/legislation/directive_2002_58_en.pdf (20 Apr. 2003)
[13] URL: http://www.spamhaus.org/rationale.html (20 Apr. 2003)
[14] Weiss, Todd R. PcWorld.com. URL: http://www.pcworld.com/news/article/0,aid,108007,00.asp (20 Apr. 2003)