

# Response to the OFTA Consultation Paper, "Proposals to Contain the Problem of Unsolicited Electronic Messages"

*Prepared by Hong Kong Computer Society*

## 1 Contents

1	Contents .....	1
1.1	Figures .....	3
2	Introduction.....	4
3	Extent of the Problem .....	4
3.1	Technical.....	4
3.2	User.....	5
3.3	Blocking.....	5
3.4	Communications .....	5
4	Industry Co-Operation .....	5
4.1	Codes of Practice .....	5
4.2	Practical Guidelines and Tips .....	6
4.3	Common Industry Blacklist .....	6
5	Information Campaign.....	6
6	Technical Solutions.....	6
6.1	Combining Techniques .....	7
6.2	Hiding Addresses .....	7
6.3	Blacklists.....	8
6.3.1	Known Abusive Server Blacklists .....	8
6.3.1.1	SpamCop.....	8
6.3.1.2	Spamhaus .....	9
6.3.1.3	Osirusoft.....	9
6.3.1.4	MAPS RBL.....	9
6.3.2	Open Relay Blacklists.....	9

6.3.2.1	MAPS RSS.....	9
6.3.3	Dynamic User Blacklists .....	10
6.3.3.1	MAPS DUL .....	10
6.3.4	Blanket Blacklists .....	10
6.3.4.1	Blackholes.US.....	10
6.3.5	Summary .....	11
6.4	Pattern Matching.....	11
6.4.1	The "Scunthorpe Problem" .....	11
6.4.2	Advanced Pattern Matching.....	12
6.4.3	Summary .....	12
6.5	Bayesian Filtering .....	12
6.6	Computational and Human Interaction Challenges.....	12
6.6.1	Computational Challenges.....	12
6.6.2	Human Interaction Challenges.....	13
6.7	SPF.....	13
6.8	DomainKeys .....	15
7	Legislative Solutions.....	15
7.1	Existing Legislation .....	15
7.2	Principles .....	16
7.2.1	Solicited .....	16
7.2.2	Appropriate .....	17
7.2.3	From and Easily Identifiable Source .....	18
7.2.4	Legal and Ethical .....	18
7.2.5	The recipient subscribed to the list .....	18
7.2.6	The List Should be Well-Managed.....	18
7.3	Network Control .....	19
7.4	Blocking.....	19
7.5	Prevention of Address Harvesting .....	21
7.6	Penalties and Prosecutions.....	21

7.7	Exemptions .....	22
8	Other Problems .....	23
9	Annex: Case Studies .....	23
9.1	Case 1: Information Security and Trading SME .....	23
9.1.1	Mail Server Log Analysis .....	23
9.1.1.1	Events.....	24
9.1.1.2	Derived Statistics .....	26
9.1.1.3	Sources of Error .....	27
9.1.1.4	Results and Analysis.....	28
9.1.2	Port 25 Probes .....	32
9.1.2.1	Sources of Error .....	34
9.1.2.2	Discussion.....	34
9.1.3	Faxes .....	34
9.1.3.1	Sources of Error .....	34
9.1.3.2	Results and Analysis.....	34
9.1.4	Mailing List.....	35
9.1.4.1	Cost Comparison.....	35
9.1.4.2	Discussion .....	36
9.1.5	Conclusions.....	36
9.2	Case 2: User at a Large non-Profit Organisation.....	36
9.3	Case 3: User at a Small Consultant Company .....	37
9.4	Other Spam Studies.....	38
10	Annex - Glossary and Abbreviations.....	38

## 1.1 Figures

Figure 9.1.1-i: Mail Statistics .....	27
Figure 9.1.1-ii: Estimated Cost of Not Blocking Spam.....	29
Figure 9.1.1-iii: Anti-Spam Techniques .....	30
Figure 9.1.1-iv: Anti-Spam Simple Violations.....	31

Figure 9.1.1-v: Anti-Spam Blacklists .....	31
Figure 9.1.1-vi: Honeypot Addresses .....	32
Figure 9.1.2-i: Port 25 Probes .....	33
Figure 9.1.2-ii: Port 25 Sources and Probes per Destination.....	33
Figure 9.1.3-i: Faxes, 1 - 9 August 2004 .....	35
Figure 9.1.5-i: User B's Email .....	37
Figure 9.1.5-i: User C's Email .....	38

## 2 Introduction

This paper is a response to the OFTA Consultation Paper, "Proposals to Contain the Problem of Unsolicited Electronic Messages", and has been prepared by the Hong Kong Computer Society (HKCS). The Information Security Specialist Group of HKCS previously submitted a paper, "Anti-Spam Recommendations: Appropriate Legislation" ([http://www.hkcs.org.hk/en\\_hk/doc\\_general/as-recommend-final.pdf](http://www.hkcs.org.hk/en_hk/doc_general/as-recommend-final.pdf)) that was referred to on page 19 of OFTA's consultation paper; this paper is intended as a supplement to, not a replacement for, the earlier paper. It seeks to address the specific questions raised in OFTA's consultation paper, and to add information where there are new developments. The sections are named according to the questions they address.

In short, spam is a significant and growing problem that must be addressed. Improvements to industry cooperation are possible, but even the best possible cooperation will not solve the spam problem. Likewise, user education and technical control methods are only partial solutions. Legislation will be an important component in a multi-faceted anti-spam solution, and the Government has the opportunity to learn from Overseas experience in drafting world-leading laws worthy of Asia's IT Hub.

## 3 Extent of the Problem

Some members of HKCS have prepared reports of the spam problem, based on the data they have collected either as an individual, or from their organisation. These have been included in the Annex, section 9 .

Calculating the impact of the problem can be difficult but in summary can be divided into four categories:

### 3.1 Technical

Bandwidth to transfer the spam, processing power and disk space. For ordinary organisations, these are generally negligible, but they become significant for the very small and very large cases. Dial-up users have very limited bandwidth, and pay per-minute, so transferring mail can become quite tedious and costly -

spam hits at those struggling to overcome the Digital Divide. At the very high end, major ISPs may need to install more bandwidth, mail servers and storage - their cost per user will be very low, but their margins will also be tight.

## **3.2 User**

These are generally the highest category of costs. Even reading the subject and deleting a message takes time. Companies will also be concerned about the time wasted by staff who choose to open messages and read inappropriate content during office hours. The worst-case costs are the losses of users who are taken in by the fraudulent get-rich-quick schemes.

## **3.3 Blocking**

Organisations may use various techniques in an effort to control the costs and effects of spam. There are costs of purchasing, installing and maintaining these.

## **3.4 Communications**

Whether done by a user inadvertently, or by an automated anti-spam solution, blocking a message that is not spam disrupts our communications. In the worst case, a major deal worth millions of dollars could be jeopardised by the misclassification of a single message.

Calculating the human costs of spam is even more difficult. Users have no effective methods for dealing with spam - replying with an 'unsubscribe' message results in yet more spam arriving, and deleting it is tedious and error-prone. A person facing a problem they cannot influence becomes frustrated and anxious. This contributes to job dissatisfaction, disempowerment, absenteeism and maybe even psychological problems.

# **4 Industry Co-Operation**

## **4.1 Codes of Practice**

The voluntary codes of practice have been ineffective at controlling the problem of spam. This is particularly evident in the case of email and fax, as can be seen from the case studies in section 9 . The Codes also have deficiencies, the code on SMS only addresses inter-operator messages, not intra-operator messages, and certainly not addressing operator-originated messages: current mobile operators do not seem to understand that customers have not bought their service to receive promotional and other annoying messages from the operators themselves.

The major contribution that operators can make, and can be encouraged to make through codes of practice, is in preventing their users, and themselves, sending spam. One improvement that could be made is in feedback to spam reporters: the codes describe mechanisms for reporting spam, but not any mechanism for informing reporters of the result of an investigation, or action taken. Typically, sending a report to the "abuse" address of an ISP generates an automated reply along the lines of, "Thank you, we will investigate.", and nothing else. It is easy to believe that no investigation took place, and no action was taken, especially when the overall level of

spam remains high. Therefore, people do not bother to report. Perhaps reporters could be directed to a list of investigations and results, to look up their case-number.

To be at all useful, the codes of practice must be improved, and made mandatory.

## **4.2 Practical Guidelines and Tips**

Any competent Operator could produce practical guidelines or tips for their users, there should be no need for a Government-backed, Industry-wide initiative to implement something so simple. Making such an initiative a major component, or the only component, in an anti-spam strategy would imply that the Government sees spam as primarily the responsibility of the users and therefore does not consider the source of spam to be a problem that requires any form of remediation action.

## **4.3 Common Industry Blacklist**

Creating a reliable blacklist requires a large, and continuing, investment of resources. Reports must be verified, evidence collected, appeals considered. Before the list can be established, the policy for inclusion must be debated and decided. Enforcing such a blacklist across all Hong Kong operators is tantamount to making and implementing laws; therefore it should be the Government that decides the policy, and an independent body that implements it.

However, the intention of the blacklist could also be questioned: if it identifies spam sources in Hong Kong, why not simply shut down the source? If it identifies external sources, it might be misconstrued as an intentional barrier to free communication and trade.

Additionally, commercial blacklists are already available, and organisations can choose one that has an inclusion policy that matches their requirements. A Government-funded, compulsory list would, inevitably, compete with these. There might be good reason to regulate commercial blacklists: require them to publish their policy for inclusion, and have a defined and effective appeals procedure. Consumer choice cannot be effective unless consumers are provided with accurate and relevant information.

## **5 Information Campaign**

An information campaign would be a useful part of an overall anti-spam strategy. However, even if such a campaign is, say 99% effective at educating consumers, the economics of spam mean that the remaining 1% will still support the spammers activities.

## **6 Technical Solutions**

There is discussion of some technical solutions in our previous paper, "Anti-Spam Recommendations: Appropriate Legislation" ([http://www.hkcs.org.hk/en\\_hk/doc\\_general/as-recommend-final.pdf](http://www.hkcs.org.hk/en_hk/doc_general/as-recommend-final.pdf)), in particular section 5 (Ineffective Solutions) discusses email charging, whitelists, and digital signatures. The discussion here references some data from the case studies (section 9).

It should also be noted that none of these techniques are solutions - they do not resolve the problem, in most cases they reduce the impact by preventing the spam reaching the user. They are also not permanent: we are facing intelligent adversaries, when spammers note that they are not getting through, they change their approach.

The discussion concentrates on email techniques: fax does not lend itself to automatic classification of messages, and mobile devices are currently generally too limited to support the sophisticated techniques used on email, though this could be expected to change.

One significant difference between email and mobile messaging is that, because mobile operators generally charge per message, they should always be able to identify the sending device, unlike email, where forging of the sender is currently a major contribution to confusion. However, it should be noted that more powerful, programmable mobile devices are vulnerable to viruses, worms and Trojans, which could be utilised to send spam from the compromised devices. Indeed, the first worm that spreads from mobile phone to mobile phone, SymbOS/Cabir, was reported in June 2004.

## **6.1 Combining Techniques**

As noted in the consultation paper, most anti-spam solutions combine several techniques. Least sophisticated is serial testing, as used by Company A in section 9.1 : the tests are applied in turn, and the message is rejected as soon as it is identified as spam. This provides the lowest false negative rate, at the expense of a false positive rate that is the sum of all the components' false positive rates.

A little more sophisticated is including tests that will give an automatic pass - for example, a message from an address on a whitelist will get through, even though the sending server is on a blacklist.

Most sophisticated are solutions that use a complex scoring technique. These might have hundreds of tests that each produces a score. The scores are combined to give a final "spaminess" that is used to decide whether to reject the message or not. These solutions can be delicately tuned to provide low false positives and low false negatives, but understanding the interaction of all the rules and scores is a specialised skill. These will probably be tuned by the vendor's support personnel, and will come with a support contract.

## **6.2 Hiding Addresses**

The only really effective form of this technique is to never give out the email address, and never use it for sending email, that is, give up using email. However, spammers have various methods for collecting and verifying addresses, and, by restricting your activities it may be possible to lengthen the time before the spammers find your address:

- i. Do not put email addresses on websites (section 9.1.1.1.4.3 demonstrates that spammers harvest addresses from websites).
- ii. Do not post to newsgroups (see section 9.1.1.1.4.3 again).

- iii. Do not reply to spam.
- iv. If email addresses must be put on websites, conceal them from simple address harvesters, e.g. use HTML entity encoding, or a word trick (e.g. "name-trousers@domain.com - remove clothes to send email" - this type of trick is probably confusing to people with a different native language).
- v. Choose email addresses that are difficult to guess - some spammers use lists of names or job functions. Unfortunately, it also makes the addresses difficult to remember.

Spammers can bypass all of these controls. Decoding entity encoding is simple, if they became common, spammers would automate solving word tricks too. Brute force guessing of addresses in a domain (a, b, c, ...z, aa, ab... - this type of attack is easier for phone and fax numbers) will eventually find any address, but using an address to communicate with anyone puts it at risk of being discovered by spammers: a virus or Trojan on the recipient's machine could find the address and disclose it.

## 6.3 Blacklists

There are many publicly available IP-based blacklists, and some commercial ones, usually implemented through special DNS domains. Common mail server software can be configured to lookup the IP address of systems attempting to send messages, and reject the attempt if the system is found on the list.

The most important limitation of such lists is that listing affects all users, and all domains, that send mail from an IP address. If the system is the mail server for a major ISP, the ability to send email of many organisations, and thousands of users, can be disrupted because a tiny number of abusive users.

The statistics in Figure 9.1.1-iii and Figure 9.1.1-v show that even the most successful of these blacklists let through significant amounts of spam, and false positives were sometimes a problem.

There are different listing policies; this is by no means an exhaustive list:

### 6.3.1 Known Abusive Server Blacklists

The list administrators add to the list based on reports of spam-sending systems. The following partial policy descriptions are taken from the websites of the list administrators:

#### 6.3.1.1 SpamCop

The well-known SpamCop blocking list describes its policy. From <http://www.spamcop.net/fom-serve/cache/297.html>:

"The goal: To provide a system which will flag the most spam with the least 'collateral damage' (flagging of wanted email). This is a never-ending and difficult task. This system is driven by people (some would say "mob-rule"). As such, it is intelligent but also faulty. People are instructed to report only email which is UBE (Unsolicited Bulk Email). Users being what they are, they may accidentally or maliciously report other email. So the results are fuzzy, but surprisingly accurate and agile (try it yourself - the results speak for themselves)."

SpamCop clearly warns that this policy can result in false positives, from <http://www.spamcop.net/bl.shtml>:

"However, it should be noted that SpamCop is aggressive and often errs on the side of blocking mail - users should be warned and given information about how their mail is filtered. Ideally they should have a choice of filtering options. Many mailservers can operate with blacklists in a "tag only" mode, which is preferable in many situations."

### **6.3.1.2 Spamhaus**

Spamhaus places more emphasis on the evidence and its quality. From <http://www.spamhaus.org/SBL/sbl-rationale.html>

"The Spamhaus Block List ("SBL") is a database of IP addresses of direct spam sources; spammers, spam gangs and spam support services (but not open proxies or open relays), queryable in realtime by mail systems throughout the Internet for the purpose of refusing mail from known spam senders.

All SBL entries are backed up with evidence which has fully satisfied the Spamhaus Project team that the IP is under the control of a spammer, spam operation or a spam support service and that the IP or netblock represents an unwanted nuisance or threat to mail systems using the SBL."

### **6.3.1.3 Osirusoft**

Osirusoft was a large anti-spam blacklist, but it was forced to close down by a sustained Distributed Denial of Service attack in August 2003. An article about its demise (<http://www.zdnet.com.au/news/communications/0,2000061791,20277794,00.htm>) described its policy:

"The service used a variety of techniques to maintain a real-time dynamic list of IP addresses that are known to have spam originate from them."

### **6.3.1.4 MAPS RBL**

The MAPS RBL takes a very cautious approach. From [http://www.mail-abuse.com/services/mds\\_rbl.html](http://www.mail-abuse.com/services/mds_rbl.html):

"The MAPS RBL is a carefully maintained list of IP addresses that have been shown to send spam as well as those who support the sending of spam (i.e. offering services to spammers, or allow their resources to be used by those who send spam). All listings have been carefully investigated and confirmed to be a source of spam. Before any IP address is added to this list, effort is made to contact the owner of the address and correct the issue of spam being sent. Addition to this list is only as a last resort after an agreeable resolution cannot be made."

## **6.3.2 Open Relay Blacklists**

They list mail servers that can be used as open relays, usually whether or not they have ever been used for sending spam.

### **6.3.2.1 MAPS RSS**

MAPS also has an open list it calls the "Relay Spam Stopper", from [http://www.mail-abuse.com/services/mds\\_rss.html](http://www.mail-abuse.com/services/mds_rss.html):

"The MAPS RSS (Relay Spam Stopper) is an extensive list of IP addresses of known insecure ("open relay") mail servers. A third-party relay, also known as an "open relay" or "insecure relay", is a mail server that will route mail for any third party to any other third party, no

questions asked. Spammers often hunt for and abuse open relays in an effort to cover their tracks because they know their spam is unwelcome and unwanted."

### 6.3.3 Dynamic User Blacklists

These list dynamic addresses that should not, normally, be used as outgoing mail servers.

#### 6.3.3.1 MAPS DUL

Again, MAPS provides an example of this category, from [http://www.mail-abuse.com/services/mds\\_dul.html](http://www.mail-abuse.com/services/mds_dul.html):

"The MAPS DUL (Dynamic User List) is a comprehensive list of IP addresses that should not be running mail servers, usually dynamically assigned IP addresses. These are addresses that should not be a direct source of email. When they are, it is often due to a virus or other type of illegal or unknown use of the end-users resources. Mass emailers often use compromised machines, often home computers, to directly connect to the mail servers of their targets, bypassing the usual ISP gateway. The victim of this compromise is often unaware that this illegal activity is taking place on their computer. It is for this reason that the DUL was created, to block those addresses where spam should not be sourced and to help prevent the end user from unknowingly supporting spam."

### 6.3.4 Blanket Blacklists

The most aggressive type of blacklist, these contain the IP address ranges for entire organisations, or even countries.

#### 6.3.4.1 Blackholes.US

This has separate lists for many major ISPs and countries. From <http://www.blackholes.us/>:

"Blackholes.us does not list spammers, spam supporters or vulnerable hosts at the present time. These lists are meant to contain all known networks assigned or allocated to the respective provider or organizations within the respective country. Lists are created for research purposes, primarily, and are made public for any use others see fit."

Blackholes.US is not alone in using this type of disclaimer to avoid responsibility for the use of the data provided. The next paragraph clearly indicates that they are aware that their lists are being used to block email:

"If you are listed, your only recourse is to contact the administrator of the host using the data to reject mail and ask them to either "whitelist" you or to stop using the data to reject mail."

Some people see these blanket lists as a viable method for controlling spam, and advocate others to do so too. This extract is taken from responses to an article about the closedown of Osirusoft (see 6.3.1.3), on the page <http://slashdot.org/articles/03/08/27/0214238.shtml?tid=111&tid=126>:

"sad news, but there are alternatives (Score:5, Informative)  
by Indy1 (99447) <...> on Tuesday August 26, @11:26PM (#6801630)  
(http://.../) For mail admins around the world try these alternatives.  
... (list of blacklists omitted) ...  
hongkong.blackholes.us  
another spammy asian country"

The information provided by blackholes.us is available from other sources, in fact, the Internet could not work unless it was available in some form. The IANA is responsible for allocating the IP address space to Regional Internet Registries, who allocate the space to Local Internet Registries, so where an IP address is can be deduced from the allocation information. Blackholes.us merely presents it in a form that is easily searchable. There are even good reasons to have the information searchable, knowing the countries of visitors to a website is important marketing information. It is the choice of individual mail system administrators whether to use this information for blocking.

It is a sad fact that while Hong Kong is regarded as "*another spammy asian country*", legitimate users and businesses in Hong Kong, who do not send spam, will suffer from having their email blocked.

### 6.3.5 Summary

Blacklisting by IP address is inherently *grainy* - it affects whole hosts, not the users of hosts individually. There are a wide variety of blacklists available, from cautious, that will carefully avoid listing mixed sources, to aggressive, that regard affecting non-spammers on blacklisted hosts as a means of increasing pressure on the spammers. Some mail system administrators are willing to blacklist entire countries or even continents that are regarded as prolific spam sources.

## 6.4 Pattern Matching

At its simplest, this looks for unacceptable keywords (often sex related) in the subject or body of the message. Spammers look for ways to bypass the filters by misspelling words that will be easily recognised by a human recipient, but will fool the automatic scanner. The following short excerpt from an anti-spam ruleset will give some idea of the problem:

```
# Check for deliberate misspellings of common "trigger" words -
the
# misspelling is a dead giveaway that the message is spam.
if body contains "p0rn" or subject contains "p0rn" weight 51
if body has "p@rn" or subject has "p@rn" weight 51
if body contains "penls" or subject contains "penls" weight 51
if body contains "penís" or subject contains "penís" weight 51
if body contains "tenage" or subject contains "tenage" weight 30
```

This becomes a competition between spammers and anti-spammers to think up new ways of spelling these words.

### 6.4.1 The "Scunthorpe Problem"

Scunthorpe is an unremarkable English town whose name has become the byword for the major problem with simple pattern matching. The "unacceptable" keywords can occur in perfectly acceptable contexts: the name Scunthorpe contains a rude word, so people innocently discussing the town found their email was blocked. Other terms causing false positives by simple pattern matching systems include:

Essex, Middlesex      English Counties. Some companies lost contact with regional branches when they introduced simple content filtering.

Fanny	A female name.
Dick	A short form of Richard.
prick	to make a hole or a mark with a sharp point.
Horniman	Surname, a museum in South London.

### 6.4.2 Advanced Pattern Matching

Spammers constantly advance their techniques for hiding suspicious words from content scanners while leaving them visible to human recipients. Advanced pattern matching targets the techniques: if a message contains words spelt with punctuation, or broken up by HTML tags, or "sliced and diced" by HTML tables, it is highly likely to be spam.

### 6.4.3 Summary

Pattern matching is an arms race against spammers that try to make their messages unmatchable. Using end-user filtering rules is asking all users to solve the same problem independently. Specialists can produce good results.

## 6.5 Bayesian Filtering

Bayesian filtering uses statistical techniques to identify spam and non-spam. Starting with a corpus of known spam and non-spam, the messages are split into tokens. Then the probability that a token indicates that a message containing it is spam is calculated for each token. When a message arrives, it is split into tokens and the probabilities linked to them are combined to give an overall probability that the message is spam.

Bayesian Filtering is usually applied on a per user basis: each user has their own corpus of known spam and non-spam. However, the results in section 9.1.1.4.2 show that it can be successfully applied on an organisation basis, for small organisations. Bayesian Filtering is the most effective of the tested methods, and is low-maintenance.

## 6.6 Computational and Human Interaction Challenges

These techniques are similar to email charging schemes: they attempt to make bulk sending of email expensive or difficult. They also share the same problems: spammers can still fool them, and they are an additional barrier to those who are already disadvantaged: those trying to cross the "digital divide", and the disabled.

### 6.6.1 Computational Challenges

When a message is received, the mail server asks the sending host to solve a computational puzzle, thus using CPU resources. The argument is that this would be a minor burden for normal email senders, but, because spammers send huge volumes of mail, their computers would be unable to cope with the load. It increases the cost of email without requiring a payment.

The problems with this are:

1. Charities that have legitimate large mailing lists will face an extra burden.
2. Computational power is increasing rapidly, a challenge that is only moderately difficult for a new computer would take many times longer on an older machine. The computational challenge further disadvantages those who are already disadvantaged: using old, second-hand computers.
3. Spammers can avoid the problem by using compromised computers. It would be the victim's computer that is burdened by the challenge, and the spammer will be unaffected.

## 6.6.2 Human Interaction Challenges

When a message is received, the mail server asks the sender to solve a puzzle that a human can do easily, but which is difficult or impossible for a computer. It is claimed that this would prevent spam by ensuring a real person is sending the message.

One example of a puzzle is displaying a picture of a number, and asking the user to enter the number. This technique is already used by the Hong Kong Domain Name Registry to prevent harvesting of addresses from their Whois database (see <http://www.hkdnr.net.hk/hkdnr/Whois.do>), or asking the user to click a certain point in a picture.

The problems are:

1. Legitimate mailing lists will have difficulties.
2. The challenges may be an insurmountable barrier to people with disabilities. People with impaired vision would not be able to complete puzzles based on viewing images, and people with impaired coordination or fine motor skills would be unable to click on a point accurately.
3. The claims that such puzzles cannot be solved by computers may be exaggerations: intelligent character recognition may be able to recognise even disguised numbers, and simple pattern recognition can defeat some of the "click the point" puzzles.
4. Spammers can defeat the control by passing on the challenge: for example, the spammers can take the image displayed, and use it for access control to a website they control, "To view our free pictures, enter the number you see above". By making the supposed website content sound attractive, they can effectively recruit a large number of humans to solve the puzzles.

## 6.7 SPF

The Sender Policy Framework (previously known as "Sender Permitted From", see <http://spf.pobox.com/>) is a draft Internet standard (<http://spf.pobox.com/draft-ietf-marid-protocol-00.txt>) that aims at preventing email address forgery and making it easier to identify spam, worms and viruses.

SPF allows domain administrators to declare which hosts are allowed to send email for their domain. This is done through an extra record in their DNS. When a mail server receives a message, it can look up the record for the claimed source domain, and check whether the host is allowed to send the mail. If it is not, then the message is some sort of forgery and can be rejected.

SPF does not prevent spam being sent, but it does allow forged emails to be identified. It has an impact on various types of email abuse:

1. It prevents the "Joe Job", where mail from a particular address is forged so that the owner of the address is deluged by return messages, whether they are bounces (because the recipients' address does not exist), or complaints.
2. Many recent self-emailing viruses forge the address of the sender. If a gateway anti-virus system detects and blocks the virus, it may try to send a warning to the "sender", using the forged address. This is currently causing a lot of unnecessary mail traffic that SPF can prevent.
3. Phishing is attempting to gather authentication details (e.g. for online banking) by sending forged emails. SPF would force the senders to use a real sender's address, making it more likely that the recipients will realise that the message is suspicious.

Despite still being an Internet Draft, SPF has been widely adopted (though adoption in Hong Kong is currently low).

Other protocols with similar objectives are RMX, DMP and Microsoft's CalledID for Email (renamed SenderID). SPF has now incorporated the best ideas from these other schemes, but there is controversy over licensing terms that Microsoft wants to have included to protect its intellectual property. Major open source development groups, including the Debian Project and The Apache Software Foundation have stated that the licensing terms Microsoft wants are incompatible with their licensing, and they will therefore be unable to support the proposed standard if the terms are not changed. It would be a pity if the best emerging standard for identification of email senders fragmented or failed because of licensing inflexibility.

More information:

<http://spf.pobox.com/>

<http://www.ietf.org/html.charters/marid-charter.html>

<http://www.microsoft.com/senderid>

<http://news.bbc.co.uk/1/hi/technology/3624798.stm>

<http://www.apache.org/foundation/docs/sender-id-position.html>

<http://www.imc.org/ietf-mxcomp/mail-archive/msg04260.html>

## 6.8 DomainKeys

Yahoo! Inc. published an Internet Draft of another scheme, called DomainKeys, to authenticate sending servers in May 2004. The abstract describes the scheme, "'DomainKeys' creates a domain-level authentication framework for email by using public-key technology and the DNS to prove the provenance and contents of an email.'

The draft is available at:

<http://antispam.yahoo.com/domainkeys/draft-delany-domainkeys-base-00.txt>

No working implementations of the scheme are known at the time of writing, and it is unclear whether the extra complexity of public-key cryptography would provide significant advantages over SPF.

## 7 Legislative Solutions

Legislation on its own cannot stop spam. A lot of current spam is apparently already breaking existing laws: in particular, the most prolific spammers are making use of relays installed on poorly-protected computers, without the knowledge or approval of the system owners. There is no reason to think that such spammers would stop, just because they were also breaking new laws.

However, anti-spam legislation would have a lot of beneficial effects:

- i. Some messages that are regarded as legitimate by the sender are classified as spam by the receiver. There is an obvious mismatch in expectations of reasonable behaviour. Legislation would define Society's expectations, and responsible people, companies and organisations would respect that and follow the law.
- ii. International co-operation
- iii. Protection of Hong Kong business interests. As noted in section 6.3.4.1 , businesses will suffer while Hong Kong is regarded as "another spammy Asian country". Effective legislation can improve the reputation of Hong Kong.

We are not experts at drafting legislation, so this section aims to give a direction, not to be a draft Ordinance. Hong Kong is not the first jurisdiction to legislate, so some references are made to the experience of other jurisdictions. Section 7.1 looks at related existing legislation that should be amended. Section 7.2 sets out principles for the legislation to follow, and discusses their implications.

### 7.1 Existing Legislation

As we discussed in our earlier paper, ([http://www.hkcs.org.hk/en\\_hk/doc\\_general/as-recommend-final.pdf](http://www.hkcs.org.hk/en_hk/doc_general/as-recommend-final.pdf)) ISPs in Hong Kong work under the Telecommunications Ordinance (TA), which was designed for telephone networks. It is therefore ill-equipped to deal with the realities of Internet communications.

Particular consideration should be given to:

- i. Delaying messages: Sections 24 and 25 of the TA prohibit delaying messages, but SMTP email is store-and-forward - delays are technically unavoidable. It would be useful to clarify the meaning of delay in different technical environments.
- ii. Blocking messages: Sections 24 and 25 of the TA prohibit blocking of messages. The complex issues of when blocking is required, desirable or damaging are discussed in section 7.4 .
- iii. Forging messages. Sections 24 and 25 of the TA prohibit the "uttering" of a message that is known to be forged. What standard of proof is required? If a message is suspect there is a danger of ISPs being trapped between this and the requirement of not blocking messages.

## 7.2 Principles

Ultra-cheap messaging enables spam (see section 2, The Nature of the Problem in [http://www.hkcs.org.hk/en\\_hk/doc\\_general/as-recommend-final.pdf](http://www.hkcs.org.hk/en_hk/doc_general/as-recommend-final.pdf)) but the problem is that the recipient has no control. The legislation should aim at giving recipients control over the messages they receive and giving senders responsibility for the messages they send. Therefore messages should be:

- i. Solicited
- ii. Appropriate
- iii. From an Easily Identifiable Source
- iv. Legal and Ethical

If a message is sent via a mailing list:

- v. The recipient subscribed to the list
- vi. The list should be Well-Managed

### 7.2.1 Solicited

Messages should be **solicited**: either the recipient (or a qualified representative) asked for the message, or the address was advertised as a contact point for the purpose it is being used. Additionally, the sender should cease to send messages, when requested.

If the recipient is an organisation, then the "qualified representative" would be an Officer of the organisation, or an authorised member of staff. Parents or Guardians would normally have control of their children's addresses. The Postmaster would always be the qualified representative for an account that has been closed.

The concept of a contact point having a purpose is very important for facilitating normal communications. If, for example, a person creates a webpage about their hobby, basket weaving, then including their email address on that page should be taken as an invitation that is unlimited by person (anyone can send a message) by limited by purpose (as long as it is about basket weaving as a hobby). There should be

no need for the page to specifically list acceptable types of message, because there is a clear implied purpose.

Similarly, if a webpage lists a company's products, and gives an address, sales@domain.com, the content of the page and the mailbox name clearly imply that the address should be used for sales enquiries. This could also apply to messages in Electronic Data Interchange (EDI) systems.

Some addresses have a purpose defined in the communication standards, for example, the email addresses postmaster, webmaster and hostmaster are defined as well-known contact points for particular purposes in the Internet standards. A message that matches the purpose should always be permitted, even if the relevant mailbox does not exist.

A specified purpose would add on to any implied purposes. Stating, "please send sales enquiries to postmaster@domain.com" would not stop the address being the contact point for mailing system errors. If the owner of the mailbox sales@domain.com said, "send the holiday pictures to my work address, sales@domain.com", then this would be a permitted use of the address. Of course, the recipient might find they are in breach of their organisation's policy on email usage, but this would be their responsibility, and a matter between them and their organisation, not the sender's problem.

Equally important is that permission can be withdrawn. By default, a reply to the sender asking for no more messages should be obeyed. However, there are many different programs for managing mailing lists, each with a particular method for unsubscribing. It would not be feasible or appropriate for legislation to mandate a particular protocol. Therefore, if the default method is not supported, there should be instructions on how to unsubscribe near the start of the message.

## 7.2.2 Appropriate

Messages should be **appropriate** to the recipient; for example, messages should be in a language that the recipient understands. If the sender does not know what languages the recipient understands, why are they sending them a message? Another example is pornographic messages received by children. Further examples include:

<b>Subject</b>	<b>Recipient</b>
Mortgages in the USA	Outside of the USA
Penis enlargement products	Women
Breast enlargement products	Men
USA "Government Contractor Report"	Outside the USA

### 7.2.3 From and Easily Identifiable Source

The true **source** of the message must be easily **identifiable**. For email, the From: should work, and must relate to the actual message sender. This is necessary for correct identification of the sender, and for unsubscribing, and to allow alternative methods of communication to be found if anything goes wrong.

However, there is an important role for anonymous messages in a free society, in particular, to protect whistle-blowers. Therefore, public interest should be allowed as a defence for anonymising services, so long as there are measures to prevent bulk mail sending.

### 7.2.4 Legal and Ethical

Messages **should not propose** any kind of **illegal or unethical activities**. This will cover invitations to defraud the Nigerian Government, as well as any kind of porn-related message sent to a minor - if the sender does not know the age of the recipient, they should not send age-restricted content.

### 7.2.5 The recipient subscribed to the list

If a message is sent via a mailing list, the recipient must have **subscribed** to the list. Using a purchased bulk mailing list does not absolve the sender from the responsibility of ensuring that the recipients wanted to be sent the messages.

This is not intended to prevent an organisation outsourcing management of its mailing lists, or any part of its messaging systems.

### 7.2.6 The List Should be Well-Managed

Addresses that do not accept messages should be regularly removed from the list.

Changes must be propagated consistently. For example, if two separate departments use a list, a request for removal received by one department should be promptly communicated to the other department. If an organisation provides a list to multiple organisations, it must process and pass on removal requests received via any of those organisations. If such an organisation ceases to exist, the client organisations must either cease to use the address list(s) it provided, or establish an alternative method for communicating removal requests between themselves.

The source of each address on the list should be recorded. In the case of a simple list, this could be implied: "a subscribe request was received from all the addresses on the list". For more complex cases, like the example above, the information will be important for resolving difficult cases.

Some organisations, for example, the Hong Kong Productivity Council, use a large number of mailing lists to communicate with very different interest groups. Understanding which lists they are subscribed to, and which messages came via which lists can become confusing for subscribers. Such organisations might find it useful to have an easy mechanism for people to get a list, with description, of all the lists managed by the organisation, and the lists that they are subscribed to. Such

"meta-list" management should probably be encouraged, but probably not set rigidly in legislation.

## 7.3 Network Control

A lot of current spam is sent via open relays. Although, in the early stages of the Internet, open relays were essential for enabling delivery of email, nowadays it is not necessary. The case study suggests (see section 9.1.2 ) that an open relay connected to the Internet will be discovered and exploited in less than a day, so for a network administrator to do this could be considered negligent. This could be a civil offence, so that victims could seek compensation for the disruption.

However, the vast majority of open relays were probably installed without consent on poorly-secured home user or Small-Medium Enterprise machines, perhaps by a virus or trojan. In these cases, the owner of the machine is also a victim, but the machine is still polluting the Internet for others. Control should probably be exercised via ISPs.

ISPs could block outgoing connections from their customers to Port 25, except to the ISPs mail server. Thus, traffic from unauthorised open relays would be blocked. However, this would limit choice and flexibility, locking customers in to the ISPs services. Why shouldn't, for example, a small charity with a large legitimate mailing list use a cheap broadband connection to push out its mail from its mail server efficiently? A middle road that cuts down unauthorised open relays, but allows responsible use would be to block Port 25 by default, but customers could fill in a form to obtain full access. The form would constitute acceptance of the responsibilities of a network administrator for the relevant IP addresses.

## 7.4 Blocking

We normally design communications systems with the objective of delivering messages, so, as the existing Telecommunications Ordinance recognises, to deliberately block messages is subverting the purpose of the system. Blocking could also constitute censorship on free speech.

However, blocking becomes desirable when certain types of unwanted messages threaten to disrupt useful communications. This is certainly the case with today's email spam. Even more clear is the case of mass-mailing email viruses, which are also very common. In this case, the message is a deliberate (by the person who released the virus) attempt to infect other people's computers, which constitutes criminal damage under the Computer Crimes Ordinance. Why not stop this crime-in-progress?

Even "desirable" blocking can be dangerous; this is illustrated by a recent message from technical support at a local ISP, shown in the text box below:

Date: Thu, 19 Aug 2004 11:09:06 UT  
From: Pacific Supernet <service@pacific.net.hk>

Dear Valued Customer,

VIRUS ATTACK WARNING: WORM\_RATOS.A

Thank you for using Pacific Supernet.

It is reported that email worm, WORM\_RATOS.A, is propagating rapidly. The mass-mailing worm arrives as a seemingly harmless email message, with the following details:

Subject: photos

Message body: LOL!;))))

To ensure that our customers are protected, emails with subject only contains "photo" will be blocked. For more information about the virus, please visit Hong Kong Computer Emergency Response Team Coordination Centre:

<http://www.hkcert.org/valert/vinfo/w32.mydoom.s@mm.html>

The problem is that the ISP has unilaterally decided to block messages based on over-simplistic criteria (the Subject of the message is 'photos'). The same ISP has previously announced blocking of messages with the subjects 'Test' and 'Hello'. The criteria would be more accurate if the presence of the virus itself was checked. In the current situation, the ISP is blocking innocent messages, and therefore breaking the strict terms of the Telecommunications Ordinance. There are also limited circumstances when it is important for a message containing a virus to be delivered, for example, when a user is sending a virus sample to an information security specialist.

In extreme circumstances, such blocking might still be acceptable in an emergency situation, if there was no better method for identifying the messages, **and** the volume of the messages threatened to bring down the telecommunications system. In other words; if the choice was between the imperfect blocking and the system failing completely.

One good feature of this ISP's decisions is that they are open: they have taken steps to inform their customers, and an error message is returned to the sender of the messages blocked. Silent blocking would be even worse, because the sender would naturally assume the message had been delivered successfully.

Another tricky situation is where an organisational messaging system is used for personal messages. An organisation can implement blocking of its own messages, because it is the recipient, a legal person (perhaps this should be termed discarding, not blocking, because the recipient has received the message and decided to discard it). However, people also use their work-provided addresses for personal communications, how should organisations deal with the privacy and censorship issues of examining and blocking personal messages? Very few organisations currently have policies covering email usage and similar issues.

We recommend the following principles regarding blocking:

1. Carriers can block in emergency situations, where the blocking is necessary to ensure the delivery of most messages.
2. Carriers are permitted to offer blocking services to recipients.
3. Blocking services should be transparent: the criteria for blocking must be published, and a rejection message must be sent to the message source when a message is blocked.
4. Recipients can choose whether or not to have the blocking service on.
5. Organisations should have a policy on the use of their messaging systems for personal messages. In the absence of a policy, addresses that are clearly function-related (e.g., info, sales, postmaster) are considered organisation addresses. All other addresses are considered personal addresses, and the organisation should behave as if it were a carrier with regard to blocking.

## **7.5 Prevention of Address Harvesting**

Spammers use software, referred to as address harvesting software, to search web pages for email addresses (see section 9.1.1.4.3). Australia's legislation addressed this by making sale and use of harvesting software an offence. This is an area where the value of the legislation is to make clear what practices are acceptable. As the harvesting is simply processing of web pages browsed from the Internet, it would take a Police raid while the software was executing to prove illegal use of harvesting software. Banning sale of such software would be useful in discouraging open marketing, but only minor modification is required to turn any web crawler software into address harvesting software.

## **7.6 Penalties and Prosecutions**

Sending juveniles indecent or obscene messages should have a heavy penalty, including prison.

The effect of prosecutions in other jurisdictions is hard to judge because the laws were only enacted relatively recently. However, it is interesting to note that private action is not permitted by the Australian legislation, and the relevant government agency only has a small budget for prosecutions. In the USA, the focus is also on high-profile prosecutions of the biggest spammers, e.g. :

[http://www.theregister.com/2004/09/24/ms\\_anti-spam\\_lawsuit/](http://www.theregister.com/2004/09/24/ms_anti-spam_lawsuit/)

but tracing and catching the worst offenders is recognised as a difficult problem:

[http://www.theregister.com/2004/09/17/spam\\_bounty\\_analysis/](http://www.theregister.com/2004/09/17/spam_bounty_analysis/)

However, merely focussing on the professional spammers will not control spam in the long term. Spam marketing would still be an attractive option for Small/Medium Enterprises seeking to expand the market for their own products. Individually, an SME sending, say, 10,000 messages to a harvested address list is a minor problem, unlikely to attract a high-profile prosecution. However, 100,000 SME's doing the same will still produce just the same congestion in the messaging system as 10 professional spammers sending 100 million messages each. SME spammers should be dealt with by a combination of education on how not to spam (fair collection of addresses, correct list management) and the threat of a (minor) prosecution.

Thus, the penalty for sending spam should be related to the number of offending messages. Perhaps a fine beyond a linear relationship is appropriate: for example, the fine could be proportional to the square of the number of offending messages (e.g. if the fine for one message is \$X, then the fine for ten messages would be \$100X).

Failing to manage a mailing list correctly (i.e. keeping records of opt-ins, and promptly obeying unsubscribes) should also be fined. However, chasing personal address books of people with a few hundred friends is not appropriate. Perhaps an exemption for lists below 1000, with rising fines for larger lists? But organisations should not be permitted to sub-divide their lists to reduce or avoid fines.

Obfuscation of headers or routing information with the intent of hiding the origin should have a per-message fine.

Victims who have an unauthorised open relay on their system should not be punished - the temporary confiscation of equipment during the investigation should be lesson enough. However, recklessly installing a mail server configured as an open relay should have a minor fine.

## **7.7 Exemptions**

Other jurisdictions have enacted anti-spam laws that have exemptions: Belgium's laws only apply to messages sent to natural persons, i.e., spam sent to companies or other organisations is exempt. Australia's laws exempt political parties, charities and religious organisations from obtaining consent to send messages.

We consider that such exemptions are unnecessary and highly damaging to the effectiveness of the legislation. They also introduce complexity that causes confusion, even amongst those who enacted the legislation, as evidenced by the recent political controversy over messages sent on behalf of Australia's Prime Minister, John Howard:

[http://www.theregister.com/2004/08/27/pm\\_spam\\_slam/](http://www.theregister.com/2004/08/27/pm_spam_slam/)

<http://www.abc.net.au/news/newsitems/200408/s1187075.htm>

The emphasis when drafting the legislation should be on fair principles: anyone or any organisation that operates an electronic mailing list should have a duty to ensure that it only contains addresses of recipients that have consented to receive messages from the list. This is not an onerous restriction, and it does not limit free speech.

Conversely, relevant Government Departments can assist and encourage efficient electronic communication. For example, nomination forms for elections issued by the Electoral Affairs Commission could require the candidate's official website URL. The Government webpage listing all of the valid nominations for an election could then link to every candidate's website, thus providing a fair, easily located portal for electors to use in researching the candidates. Candidates would be free to solicit subscriptions to their mailing lists on their own websites.

## **8 Other Problems**

Some people, particularly home users, are encountering problems with Spyware, Adware and Popups. In some cases, these may be in violation of existing laws (e.g. if they are installed silently, or as part of some other software, without the fully informed consent of the user and administrator they are unauthorised modification of the installed software, covered by the Computer Crimes Ordinance; or, if they collect and transfer data about websites browsed, there is an issue of fair collection of personal data, under the Personal Data Privacy Ordinance); and popups may present pornographic images to children). However, prosecutions are unheard of. They are often deliberately designed to be difficult to uninstall and they may be linked to spam and spammers in various ways. It might be appropriate to address these as Fair Trading and Consumer Protection issues.

## **9 Annex: Case Studies**

### **9.1 Case 1: Information Security and Trading SME**

The company, referred to as Company A, has been using the Internet since 1993. Initially, it used an overseas dial-up provider, and registered a sub-domain with that provider. In October 1997, it registered its own .com.hk domain, and later registered the corresponding .com domain and other domains within .com.hk, but most email traffic is still through the original .com.hk domain. Some of the email addresses have been continuously active from October 1997 to the present day. The statistics presented here are based on server log files collected between 16 November 2001 and 30 June 2004. During this period, Company A was using a 128Kbps leased line with a Class C IP address allocation.

Company A employs Information Security Specialists. In 2001, Company A was not specifically focussing on spam, but at the time of writing, anti-spam products and services are now a part of its business. This change did not affect the data collection.

Four areas are examined: the mail server logs provide detailed information about the incoming email and effectiveness of anti-spam messages; the firewall logs record searches for mail servers; a snapshot of incoming faxes is categorised; and finally, a mailing list illustrates the benefits of email.

#### **9.1.1 Mail Server Log Analysis**

Company A used a variety of anti-spam measures, changing the mixture in use as conditions changed. The statistics were generated by counting the occurrences of log lines caused by the different events. Care was taken to exclude mail created by internal systems from the count. The events are described below:

### 9.1.1.1 Events

#### 9.1.1.1.1 Incoming Mail, Size

This includes all external mail where the From: phase of the SMTP session was completed, even if the message was rejected before the DATA phase, and the session was terminated (e.g., because the sending server was on a blacklist).

Choosing the log lines to count when measuring spam events was relatively easy: triggering the rule generally gives a recognisable log entry. However, measuring the non-spam in a simple way proved more difficult. Counting the messages transferred from the gateway to the internal mail server was not suitable - it was not feasible to prevent counting of messages between internal systems. This could be avoided by counting the incoming mail to the gateway, because internal sources could be identified and excluded from the count (the "from:" log entries with a "relay=" parameter other than internal servers were counted). However, changes in the version of mail server software in use caused inconsistencies in the logs (in 2002, blacklisted connections resulted in both "Rejected:" and "from:" log entries, but, by 2004, blacklisted connections only produced a "Rejected:" log entry). This was resolved by only counting "from:" log entries which had a non-zero size parameter: blacklisted connections were rejected before the DATA phase, and therefore always had a zero size.

The Incoming Mail statistic is therefore a count of messages with a non-zero size from external systems. It includes non-spam, spam identified by Bayesian filtering, and false negatives, spam that has evaded all the automatic tests. It excludes spam identified and rejected before the DATA phase, and any message with zero size. It also excludes messages to the honeypot addresses, because they are processed within the mail gateway, and are not sent to the internal mail server.

A total of the size of the incoming mail was also calculated.

#### 9.1.1.1.2 Simple Violations

These are a collection of tests made early in the mail processing, including:

Domain of sender	The domain of the sender must exist.
Relaying Denied	An attempt was made to use the mail gateway as an open relay.
No Longer Staff	Initially, when staff leave, their incoming mail is replied to with an unsubscription request, or telling the sender to update their address books, as appropriate. After some months, the remaining incoming mail is lists when repeated unsubscription requests have failed, or other spam, and the messages are rejected before the DATA phase.
NEVER an Address	Some incoming spam arrives for addresses that have been guessed ('accounts', 'sales'), or obvious mis-spellings of existing addresses ('aster', an obvious truncation of postmaster or webmaster). These are rejected before the DATA phase.

### 9.1.1.1.3 Blacklists

A number of DNS blacklists have been used. Some were discontinued because of significant false positives. The lists are not named because they are checking in an order, and a match on one list will lead to immediate rejection, without checking of subsequent lists. Therefore, the data cannot be used to measure the relative effectiveness of the lists.

BL1	An internal blacklist; maintained by Company A. Updating was abandoned when it became too time-consuming.
BL2	A Hong Kong local list.
BL3	
BL4	Overseas lists. Use of BL 3 and BL 4 was discontinued because of too many false positives.
BL5	

### 9.1.1.1.4 Bayesian Filtering

Implemented using the Bogofilter package, this takes a corpus of known spam, and a corpus of known non-spam, and calculates the probability of particular tokens indicating whether a message is spam or not. These probabilities are used to calculate an overall spam probability for each incoming message, and messages are rejected or accepted on this basis.

#### 9.1.1.1.4.1 Invalid Assumption

At the beginning of February 2004 the Bayesian Filter started blocking almost all incoming email. On investigation, this was found to be because a basic assumption in the compilation of the known spam and non-spam corpus was incorrect. It was assumed that mail entering the organisation was similar to mail leaving the organisation; so all outgoing mail was collected and placed into the non-spam corpus. The spam corpus was formed from the mail to the spam reporting address, and the honeypot addresses.

The assumption was incorrect because the organisation does not use the commonest email client in use on the Internet. Thus, most of the incoming spam had headers characteristic of the common email client (whether or not they were actually sent using that client) and none of the outgoing mail had those headers. Eventually, the statistics indicated that any message with headers characteristic of the common email client was spam.

Once the problem was identified, large numbers of non-spam messages with the relevant email headers were added to the non-spam corpus, correcting the probabilities. Apart from this incident, no other false positives have been noted.

#### 9.1.1.1.4.2 'to spam'

Bayesian filtering was added in September 2003, and the changes accompanying it included one very important for the statistics. Staff were encouraged to send spam

that they received to a special email address, so that it could be added to the corpus of spam, and therefore contribute to improving the effectiveness of the filtering. These messages are counted as 'to spam', and they are a measure of the overall false negatives: spam not stopped by the filters.

#### 9.1.1.1.4.3 Honeypot Addresses

At the same time, some "honeypot" addresses were added, and published in places where spammer's automatic harvesters might find them. As the addresses were clearly labelled, any mail arriving at them is regarded as spam. For example, the Web honeypot address was placed on company web pages thus:

```
<font size="-3">Polite notice to people or organisations wishing to
send email of a commercial nature that we have not solicited: please
use the address <a href="mailto:*****">*****</a></font>
```

The News honeypot address was used as the sender's address in a few HK newsgroup messages advertising a job, with the content of the message clearly stating a different address for job applications. Thus, real job seekers who read the message would send their application to the real address, and spammers harvesting addresses from the newsgroups probably picked up both the real and fake addresses.

Mail to these addresses is automatically added to the spam corpus. They are **not** counted as spam to Company A in these statistics.

#### 9.1.1.1.5 Sender Policy Framework (SPF)

SPF is an Internet draft standard for preventing forging of email from participating domains.

#### 9.1.1.2 Derived Statistics

The counts from the log files were used to calculate the following derived statistics:

Total Mail	The total number of messages; whether they were accepted or blocked. The sum of Incoming Mail, SPF, Simple Violations, and Blacklists. It does not include mail to honeypot addresses.
Spam Total	Total number of spam messages. The sum of Bayesian Filtering, 'to spam', SPF, Simple Violations and Blacklists.
Non-Spam	Total Mail minus Spam Total.
% Spam	Spam Total as a percentage of Total Mail

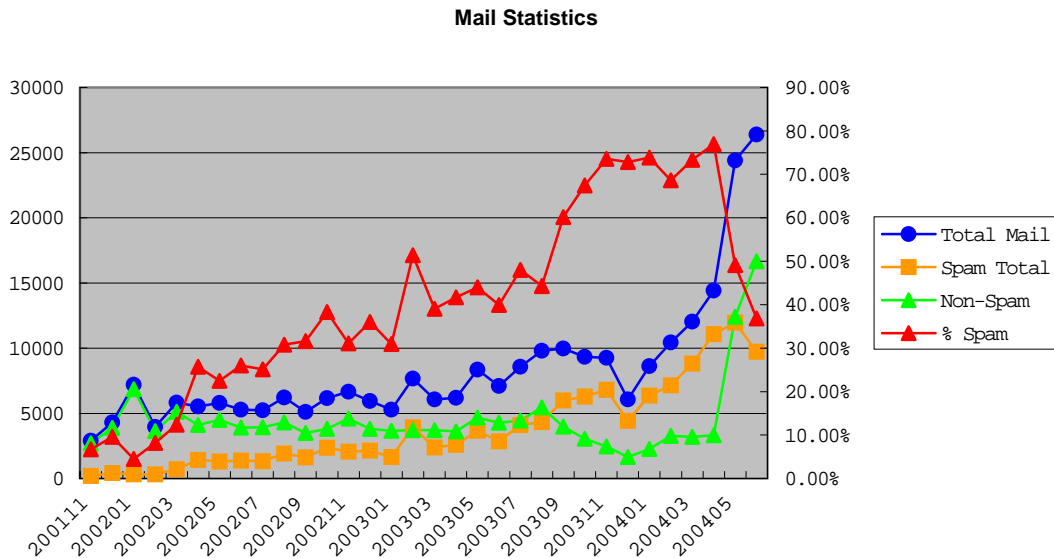


Figure 9.1.1-i: Mail Statistics

### 9.1.1.3 Sources of Error

The statistics are the 'best available' information on the company's mail traffic, but there are various shortcomings:

#### 9.1.1.3.1 Changing Methodology

The anti-spam measures in use, and even the mail system software, changed over time because Company A was reacting to a growing, changing problem. Some of the changes have already been noted and adjustments made where necessary (see 9.1.1.1 for discussion on the log format changes). However, the biggest shortcoming of this type is that before September 2003, there is no measure of the false negatives: how much spam was not identified by the measures in place.

The effects of this are discussed further in the conclusions (9.1.1.4 ).

#### 9.1.1.3.2 Time Delays

All of the statistics are derived from log entries made as the mail is being processed at the mail gateway, with the exception of the 'to spam' measurement of false negatives. The 'to spam' log entries are made when a person forwards a spam that has arrived in their mailbox to the spam-reporting address, thus, there is a variable and unknown delay between that mail being counted in the Incoming Mail, and the identification as unblocked spam. Generally, this will be longer for mail that arrives at the weekend, but, if staff are on leave, the delay may be a week or more. Therefore, the 'to spam' statistic is counting some messages in the wrong month; however, the effect of this overall is probably minor.

#### **9.1.1.3.3 Under-Reporting**

The 'to spam' statistic requires a person to identify a message in their mailbox as spam, and forward it to the spam reporting address. This is more time-consuming than deleting the message, and people may not bother, especially months after the system is installed. Staff were reminded to keep sending the reports, because they are important for ongoing training of the Bayesian filter, but it is suspected that some spam is just deleted.

#### **9.1.1.3.4 False Positives**

Any of the spam-detection methods may incorrectly identify a legitimate message as spam, and there is no specific mechanism in place to detect this. Company A relies on the sender complaining that their message has been rejected (possibly using other channels). A small number of false positives were reported at various times, and the system was updated to avoid them (including by changing the blacklists in use, see 9.1.1.1.3). It is believed that the overall number of false positives is relatively small.

#### **9.1.1.3.5 Viruses**

Email is currently the commonest vector for computer viruses. Many recent viruses forge the sender's address, causing an additional problem: if the virus forges a message from the company's domain, an anti-virus gateway elsewhere that detects the virus might mistakenly send the warning to the forged sender address. The number of such messages, containing viruses, or incorrectly addressed messages containing reports about viruses, has been growing alarmingly in recent months.

How such messages should be classified is a matter of debate: they are all unwanted and junk, but they should probably be dealt with in a different way to unwanted promotional messages. In this case study, although the mail gateway has an anti-virus function the log files cannot be correlated. In general, the mechanisms in place would not recognise the virus messages, or the resulting warnings, as spam, so they will be counted in the Non-Spam total. This probably accounts for the notable rise in non-spam for May and June 2004.

#### **9.1.1.3.6 Other Incidents**

A particular incident occurred in October 2002, where mail system problems led to a mail storm with another domain. This generated around 30,000 log entries that would have been counted in the "Incoming Mail" category; however, the count was specifically adjusted to exclude these. It is possible that relatively minor mail storms have occurred, leading to inflation of the total mail statistics.

### **9.1.1.4 Results and Analysis**

Figure 9.1.1-i shows that amount of non-spam stays at roughly the same level from 2001 to September 2003 - Company A had already been using email for 8 years, four of those using the same domain, so this reflects that staff have established their own patterns of email use, and their amount of correspondence is quite constant. There is a small dip after September 2003, probably a result of having a measure of spam false negatives (see 9.1.1.3.1) and a steep rise in May and June 2004, probably due to outbreaks of mass-mailing, forging viruses (see 9.1.1.3.5).

Although the amount of non-spam has stayed roughly constant, the total mail volume has increased steadily, reflecting a growth in identified spam from less than 7% to over 70% of total mail. Part of the growth can be accounted for by improvements in the methods of identifying spam and, finally in September 2003, the mechanism for users to report spam that was not identified automatically (see 9.1.1.1.4.2), so the early data under-estimates the total spam. A better estimate for spam in November 2001 would probably be between 10% and 20% of total mail.

#### 9.1.1.4.1 Costs

Calculating the monetary cost of spam for Company A is almost impossible and would not reflect the costs for an 'average' company: Company A specialises in Information Security, so, on occasion, time was taken to investigate the techniques being used where an 'average' user in an 'average' company would delete and forget. Company A is now treating the problem of spam as a business opportunity: what it learns from the spam it receives is used to help other organisations with their spam problem.

A simple calculation for the month of June 2004 will be based on the following assumptions:

- i. The average hourly wage of a user is \$60
- ii. It takes a user 5 seconds to identify and delete a spam message from their in-box.
- iii. It takes a user 30 seconds to identify a spam message in their in-box, and forward it to the spam reporting address.

##### 9.1.1.4.1.1 User Cost of Not Blocking Spam

The total spam for June 2004 was 9721 messages, so, if none of that had been blocked, users would have spent 13.5 hours deleting it, at a cost of **\$810**.

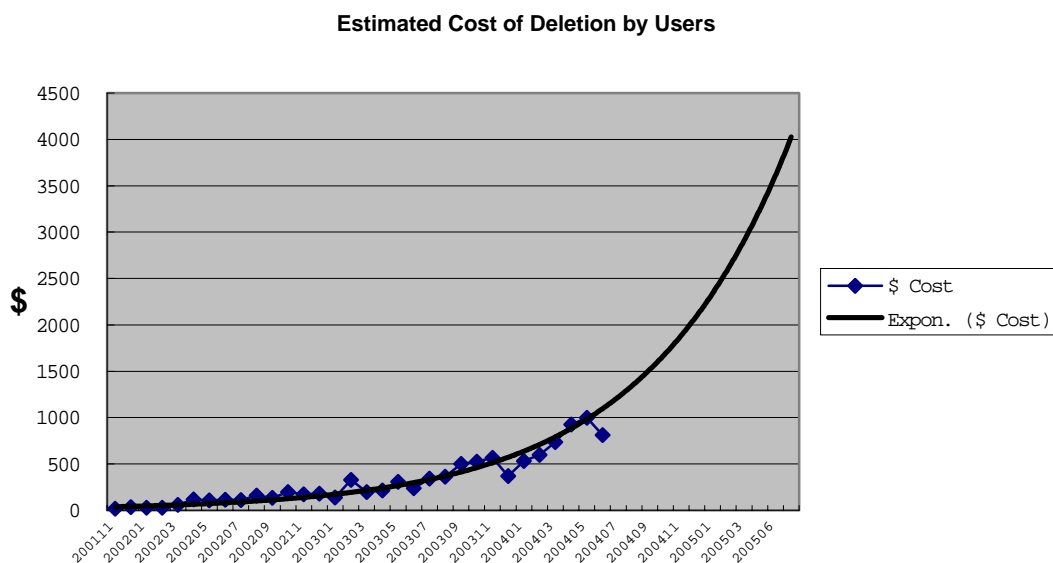


Figure 9.1.1-ii: Estimated Cost of Not Blocking Spam

Figure 9.1.1-ii shows the estimated costs since November 2001, and extrapolated to June 2005. If the exponential trend continues, the cost per month of Company A doing nothing stop spam would be about **\$4000**.

#### 9.1.1.4.1.2 User Cost of Reporting Unblocked Spam

The spam that was reported to have reached users inboxes for June 2004 was 298 messages, so, reporting it is estimated to have taken 2.5 hours, at a cost of **\$149**. This is the cost to users of maintaining the accuracy of the Bayesian Filter.

#### 9.1.1.4.2 Anti-Spam Techniques

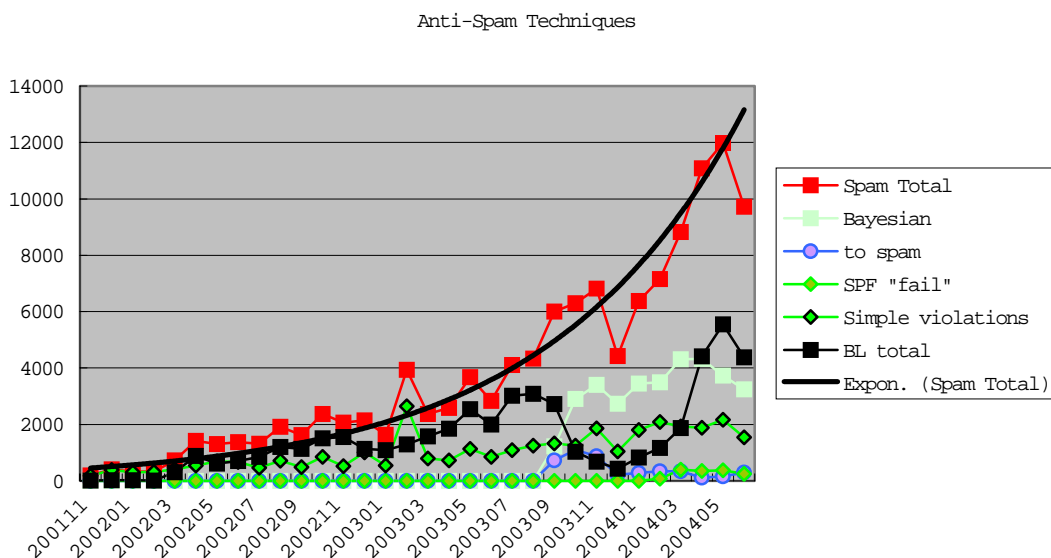


Figure 9.1.1-iii: Anti-Spam Techniques

Figure 9.1.1-iii shows the contribution made by the different anti-spam techniques, and also an exponential trend line fitted to the spam total data.

The 'to spam' statistic is particularly interesting because it represents the failure rate of all the other methods in blocking spam, and also the work being done to train the Bayesian Filter. It was introduced in September 2003, and quickly rose, then fell as the Bayesian Filter became more accurate. The Bayesian Filter has proved very effective for Company A, after training it stopped over 90% of the spam it examined (note that other studies have shown that Bayesian Filters are much less effective for groups of over 100 recipients).

The relatively new SPF technique has made little contribution so far, but it depends on domains publishing which servers are permitted to send their mail. As it is still only an Internet draft standard, not many domains have created the necessary DNS records, so it is limited, but adoption is growing rapidly. Its main advantage will be not in preventing spam directly, but in preventing forging of the sender's address, which will at least put an end to the address forging of the current crop of viruses.



The Blacklist graph, Figure 9.1.1-v, mainly shows how the choice of blacklists changed over time. Company A's own blacklist, BL 1, was found to give a poor return for the effort involved in updating it. The overseas blacklists BL 3 and BL 4 were reasonably effective, but they did occasionally have false positives, notably for major local domains, such as netvigator.com, that had a direct effect on communication with customers. The cost of a single false positive (in terms of customer inconvenience and administrator workload, even if not in direct monetary terms of lost business) far outweighs the inconvenience of receiving many spam messages, particularly for a service-oriented organisation. The overseas list BL 5 was used as a replacement, it is less effective, but has had no reported false positives. The local list BL 5 was tested recently and found to be effective, but some false positives have been reported. The more a blacklist blocks, the more likely are false positives. This probably stems from the nature of the blocking: either an IP is blocked or allowed, so they are good at dealing with spam from dedicated spam sending hosts, but unsuitable for dealing with spam from large ISPs, where there are many well-behaved users, but a small number of users are sending as much spam as they can before their accounts are terminated.

### 9.1.1.4.3 Honeypot Addresses

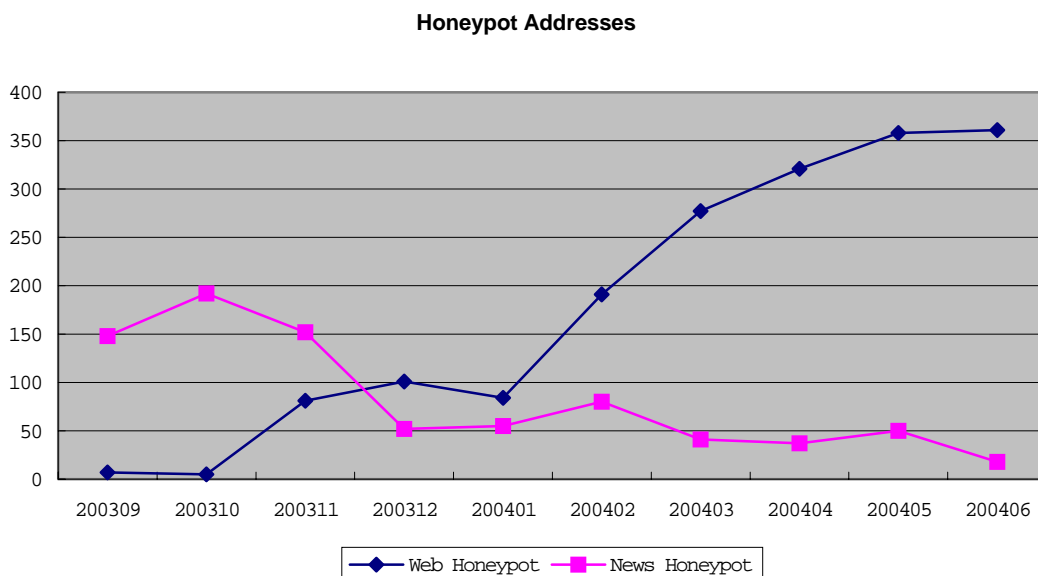


Figure 9.1.1-vi: Honeypot Addresses

The honeypot addresses are both a way of improving the corpus of known spam for the Bayesian filter, and an experiment to investigate the techniques used by spammers to collect addresses. The results show that publishing an address on a website is a good way of getting onto spammer's lists.

### 9.1.2 Port 25 Probes

SMTP email is transferred on TCP port 25, and a lot of spam is sent via open relays. A possible way for spammers to locate open relays is to attempt connecting to port 25 on random IP addresses, or a sequence of IP addresses. This type of scanning was blocked by Company A's firewall, and recorded in the log files. Figure 9.1.2-i shows the total number of such probes per month.

Probes per Month on Port 25

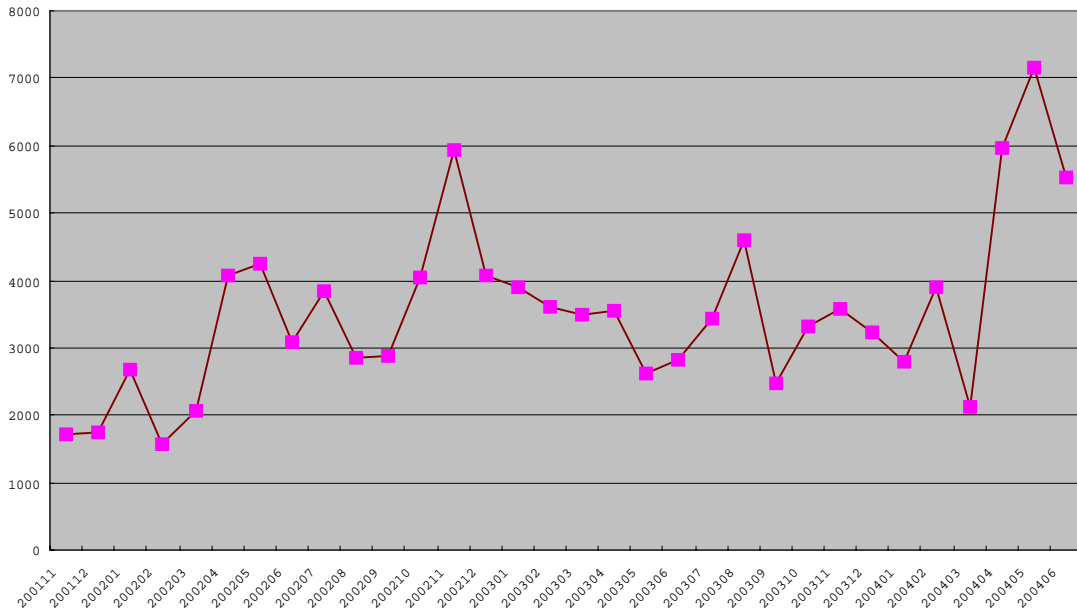


Figure 9.1.2-i: Port 25 Probes

The number of destination IP addresses being blocked and logged was dependant on the sub netting of the Class C, which changed during the logged period. To correct for this, Figure 9.1.2-ii shows the number of probes per destination IP address. Also of interest is the number of unique source IP addresses - in many instances, a single source will exhaustively probe every IP address in the destination range in a short period of time, so this statistic gives an idea of the number of sources conducting searches over the monitored range.

Probes on Port 25

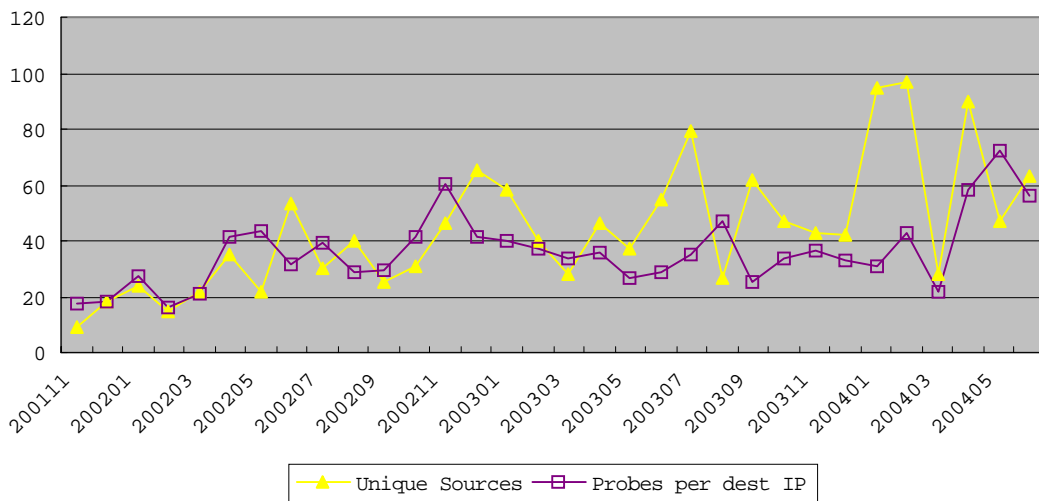


Figure 9.1.2-ii: Port 25 Sources and Probes per Destination

### **9.1.2.1 Sources of Error**

The log files are not a complete record: at various times the servers involved were down for power failures, maintenance or other reasons. Network disruptions could also prevent the probes, or the logging of events. These would all reduce the number of recorded events, but the resulting recording error is believed to be small, probably less than 1%.

Nothing is known of the intent behind the probes. Most are assumed to be a deliberate search for open relays to be exploited, but there are other possibilities. Other researchers may be surveying the number of mail servers, or open relays, on the Internet. Configuration mistakes at other sites could also cause such traffic: if a DNS MX entry for a domain was mistyped to point at the target range, other mail servers around the Internet would attempt to deliver mail for that domain to a monitored IP address, inflating the number of recorded probes, and the number of unique sources. However, such mis-configuration would probably be discovered and corrected quickly, as the domain would stop receiving mail. No such events are known to have occurred.

### **9.1.2.2 Discussion**

Overall, there appears to be a moderate upward trend in this type of scanning, but the volume of the traffic is small, and not a concern. However, if this small IP range is typical of the Internet as a whole, there are a large number of systems searching for open relays at any time. Significantly, in terms of Information Security, if a mail server is (inadvertently) configured as an open relay and connected to the Internet, but not advertised in any way, it will be discovered, on average, in less than a day. Once discovered, it will be exploited for transmitting large quantities of spam. Spammers are known to trade the locations of open relays; indeed, some spam offers lists of open relays for sale.

## **9.1.3 Faxes**

Company A uses a single-line fax server, and a fax machine to receive faxes. The fax server receives the faxes, unless it is unavailable (busy with an outgoing fax, or down), when the fax machine receives the faxes. Detailed log files are not available, so the faxes received by the fax server for a short period (1 August 2004 - 9 August 2004) were examined manually and categorised as non-spam, uncertain or spam.

### **9.1.3.1 Sources of Error**

Some faxes could have been wrongly categorised by the manual process. The data only represents a single period of 9 days, there is no information about the development of the junk fax problem over time. Only faxes received by the fax server were examined, but there is no reason to believe this introduced a sampling bias to the data.

### **9.1.3.2 Results and Analysis**

167 faxes were examined; almost two-thirds (65%) were definitely spam, see Figure 9.1.3-i. Based on the assumption that opening a fax, identifying it as spam and deleting it would take 20 seconds (spam email can usually be identified by a user

from the subject and sender, but for spam faxes, an image file must be opened), and the same hourly wage as in 9.1.1.4.1, deleting 10.9 spam faxes a day costs **\$123** per month. If the spam faxes are received on a fax machine, the user time to identify the spam will be lower, but the environmental cost of using more paper will be higher.

**Faxes, 1 - 9 August 2004**

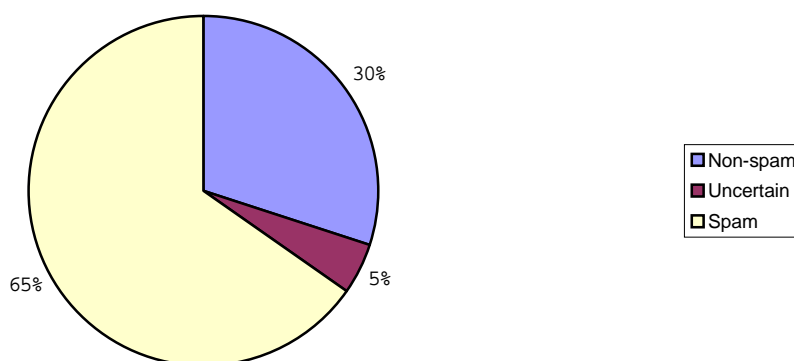


Figure 9.1.3-i: Faxes, 1 - 9 August 2004

### 9.1.4 Mailing List

Company A sends a monthly newsletter by email to a mailing list of subscribers. The mailing list is advertised on the company website, and has been publicised by other methods in the past. The mail server handles subscription and unsubscription requests directly, an email "from" an address can be used to subscribe or unsubscribe that address. This method is vulnerable to falsified subscription requests, but no incidents have been reported.

The number of subscribers at the time of writing was 940, but some subscribers report that they distribute it within their organisation. After each monthly mailing, the list manager deletes the addresses that have been bounced because they no longer exist.

#### 9.1.4.1 Cost Comparison

As Company A already has a leased line, and is maintaining the mail server, the incremental cost of distributing the newsletter by email is zero.

If the newsletter (usually 6 pages) was distributed by physical mail, the costs would be:

Item	Cost per recipient (HK\$)	Cost (HK\$)
A4 paper (6 pages, 3 cents/page)	0.18	169.20

Printing / copying (6 pages, 7.5 cents per page)	0.45	423.00
Postage	1.40	1316.00
<b>Total</b>	<b>2.03</b>	<b>1908.20</b>

The staff time to fold and label the newsletters is not estimated.

#### **9.1.4.2 Discussion**

If email spam is a big problem, why do companies persist in using email at all? The cost comparison shows that a small company can realise significant savings on a small, monthly newsletter.

The number of recipients for the list is quite small, mainly because addresses are only added on request. There are many mailing lists operating in Hong Kong, and some have been observed using much more aggressive methods to harvest addresses and add them to the lists without permission. When a company realises the cost savings of email distribution, there is a strong temptation to use unsolicited mass mailing for marketing purposes, without regard for the damaging effect on the usefulness of email, or the reputation of the company or Hong Kong.

#### **9.1.5 Conclusions**

It is estimated that not blocking email spam would cost Company A \$810 per month, and the figure is rising. By using a variety of control methods, this has been reduced to \$149 per month. Junk faxes cost the company an estimated \$123 per month, and controls have not been implemented. It would be dangerous to extrapolate these figures to to approximately 300,000 SME's in Hong Kong.

The Port 25 scanning activity suggests that an unadvertised open relay connected to the Internet will be discovered, and therefore probably exploited, in less than a day.

Electronic messaging is an important business enabler for SMEs, but it must be used responsibly.

### **9.2 Case 2: User at a Large non-Profit Organisation**

The user, referred to as User B, is an employee of a large organisation that is heavily reliant on IT in its operation. User B's position does not require publication of his email address, and the address could not be found by using some common search engines, but neither is the address secret: it appears on User B's business card and is used in communication with job-related and function-related contacts.

User B counted his junk mail and normal mail for 18 days, 9 August 2004 - 26 August 2004. The statistics are shown in Figure 9.1.5-i: User B's Email. Overall, 99 out of 513 messages, i.e. 19% were junk. This is 5.5 junk messages per day, using the same assumptions as 9.1.1.4.1 this is \$167 per year.

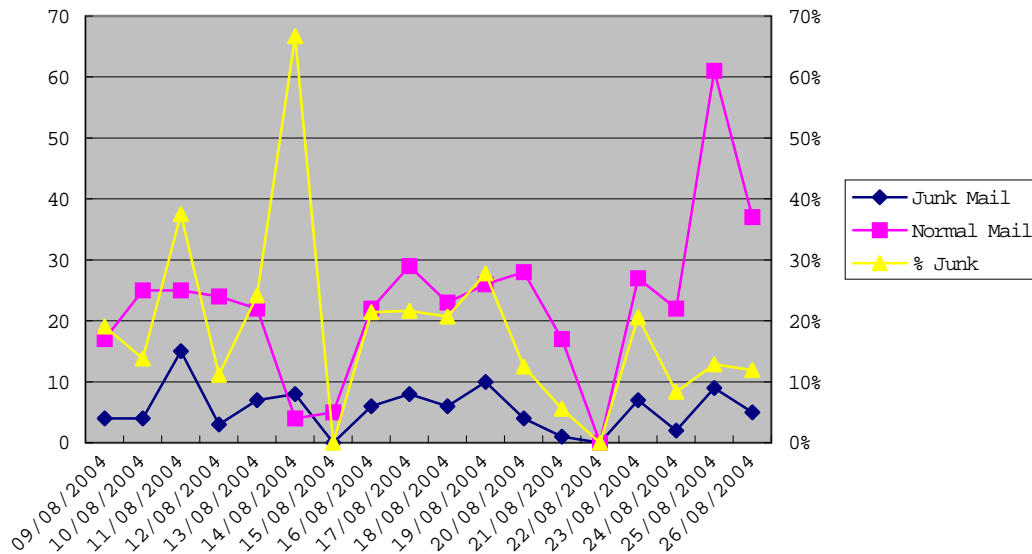


Figure 9.1.5-i: User B's Email

### 9.3 Case 3: User at a Small Consultant Company

The use, referred to as User C, is the head of a small company that provides IT Audit and Security consulting services. User C is in a prominent position, within a specialist field. User C has a commercial end-user anti-spam product installed.

User C has about 12 email addresses, and he examined one weeks' email to all of those addresses. The total number of emails was approximately 3600, the commercial product identified 169 as spam, and the total number of spam was approximately 480. Thus about 13% of User C's email is spam, and the anti-spam software catches a disappointing 35% of the spam, this is illustrated in Figure 9.1.5-i: User C's Email. Using the same assumptions as 9.1.1.4.1, deleting the junk messages that get past the installed filtering software costs \$1348 per year.

### User C's Email

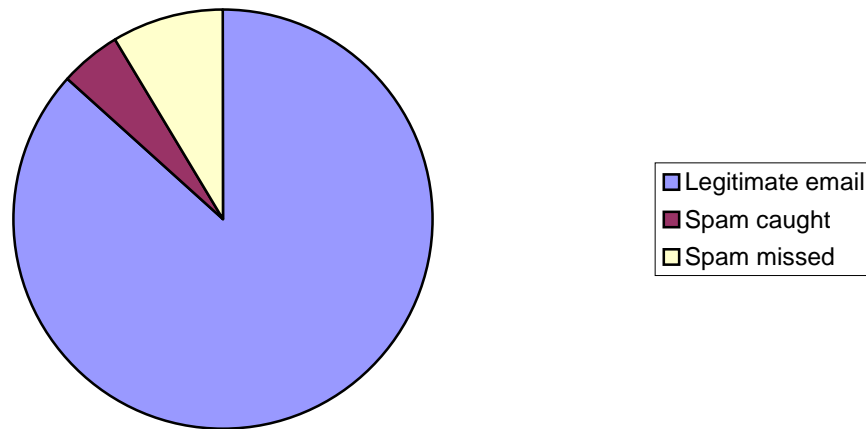


Figure 9.1.5-i: User C's Email

## 9.4 Other Spam Studies

A Visual History of Spam:

<http://weblogs.asp.net/oldnewthing/archive/2004/09/16/230388.aspx>

## 10 Annex - Glossary and Abbreviations

The following terms are widely used on the Internet and when discussing messaging systems. However, they are not all clearly defined and standardised. This paper uses these definitions, but other sources may use some terms in different ways.

- Anonymous List** A mailing list that removes the senders identification.
- Blacklist** Usually refers to IP blacklists, which are lists of IP addresses that are recognised as spam sources, used for blocking. See section 6.3
- ISP** Internet Service Provider
- Mailing List** A list of addresses on a mail server: a message sent "to the list" will be replicated and sent to every address on the list. Lists are usually established with a certain purpose (e.g. discussions on software X) or for a particular group. See also Anonymous List, Moderated List, Public List, Restricted List. Note that not all of these sub-categories are mutually exclusive: a list could be Public and Restricted (only list members can send to the list, but anyone can join).
- Moderated List** A mailing list that only accepts messages from certain sources - called the moderators. People wanting to send a message to the list send it to the moderators, who decide whether the message is acceptable.

Opt-In	A management regime for mailing lists: addresses are only added on request, <i>c.f.</i> Opt-Out.
Opt-Out	A management regime for mailing lists: addresses may be added without permission, but will be removed on request, <i>c.f.</i> Opt-In.
Postmaster	The technical person responsible for maintaining a messaging system.
Public List	A mailing list that anyone can subscribe to, either by completing a subscription form, or sending a message to the list management software.
Restricted List	A mailing list where only messages from list members are accepted.
Whitelist	A list of IP addresses or email addresses known to not send spam. See <a href="http://www.hkcs.org.hk/en_hk/doc_general/as-recommend-final.pdf">http://www.hkcs.org.hk/en_hk/doc_general/as-recommend-final.pdf</a> section 5.1.2.