The Minstrel's Articles

Spam, a 21st Century Plague
Part A: All Readers

Version 1.2

26 April 2005

Table of Contents

1 Introduction

Spam is rapidly becoming an all-pervasive blight on our online lives, making e-mail less and less a useful communication mechanism. At the time of writing, it is estimated that around 75% of all e-mail clogging our bandwidth is spam, and this figure continues to rise. When will it stop? Can it be stopped? What can we do to protect ourselves?

In this two-part article, I discuss the problem in some depth – where does spam come from, who is sending it, why they are sending, and move on finally to what can be done, both by the community in general and by the individual Internet user.

Note to the pedantic: of course, I am using the word spam as a colloquialism for Unsolicited Bulk/Commercial E-mail (UBE/UCE), its use arising from the excellent Monty Python sketch. I am obviously not talking about SPAM™, the "delicious luncheon meat" from Hormel Foods Corporation, although they don't contest the use of the name to any great degree any more, having realised that their sales have soared since the term was first used to refer to UBE/UCE. Their product even enjoys something of a cult status now, and I'm sure their merchandising revenue rivals that of the product itself! You can see their statement on the subject at http://www.spam.com/ci/ci_in.htm, which is an interesting enough read in itself!

1.1 Sources and Accuracy

Some of the content of this article is loosely based on the NISCC 'Spam Mitigation' Technical Note (please see the References section at the end of the article), which was coincidentally issued just after I started writing notes on the subject. The remainder is based on personal and professional experience and some research.

It must be noted that I do not necessarily endorse any product or service mentioned, and that I only comment on those products and services of which I have personal experience. There are undoubtedly many others available, and I recommend you perform your own research before selecting any anti-spam product or service.

As with all my articles, I reserve the right to be wrong, but where you do spot an inaccuracy, I would very much appreciate hearing about it! Feel free also to disagree with any of my opinions – it should make for a lively discussion!

1.2 Audience

This series of articles is written primarily with the general Internet user in mind, and this one is no exception. This part of the article, imaginatively entitled 'Part A: All Readers', is aimed at the general Internet user. If you're tired of having to change your e-mail address on a regular basis, or sign up to spam filtering services, I hope this article will help you, as it provides recommendations on reducing or eliminating the spam you receive in your mailbox. If you would like to look at the topic in greater technical depth, 'Part B: Administrators' is intended for anyone looking at implementing blocking, filtering or other defensive mechanisms at a corporate or community level.

1.3 Contact Details

Please come and discuss this or any of my articles on the 'Front Line' discussion list:

http://lists.internetgremlin.com/mailman/listinfo/front-line

Or come and join the forums at Privacysense:

http://www.privacysense.com/

Alternatively, you can contact me privately at:

http://www.minstrel.org.uk/contact/

I look forward to hearing from you!

Finally, on the subject of contact details, here's one for the address harvesters:

nullbox@internetgremlin.com

As the name suggests, anything sent to this address is automatically deleted, but only after being processed by our spam filter, which will learn from the experience!

2 What Is Spam?

We all have an idea of what spam is, and apart from the lucky few, we've all been affected by it to some degree. The word 'spam' is a colloquialism used to refer to Unsolicited Bulk E-mail or Unsolicited Commercial E-mail (UBE/UCE), i.e. any e-mail you receive that fits most or all of the following:

a. You did not request the information being sent to you (hence 'unsolicited')
b. You do not know the sender (whether it appears to be an individual or an organisation)
c. The same message has been sent to hundreds, thousands, even millions of people (hence 'bulk') – in these cases, you will usually be a 'Blind Copy' (BCC) recipient
d. The message is trying to get you to *do* something, which could be:
   - Pay for a product or service, whether or not you're actually interested (hence 'commercial')
   - Read some information
   - Make a 'phone call
   - Send a letter
   - Click on a link
   - Open an attachment
   - Look at an image
   - Even simply open the e-mail itself

The definitions become a little hazy in certain scenarios – for example, you completed a registration form for a conference, and granted permission for the conference organisers to contact you with 'relevant information'. Whose definition of 'relevant' is used, and is this e-mail then 'unsolicited'?

Conversely, you may receive e-mails that you agreed to receive, but they have been produced in a 'spammy' way, making them difficult to identify as legitimate. For example, a large reputable organisation may use a small marketing agency to send out its legitimate marketing material, but no consideration is given by the agency to using secure delivery mechanisms, or to the language used in the material (e.g. "FREE $$$" instead of "Appreciable Saving").

For these reasons, amongst others, legislation against spam has been difficult, as will be discussed later.

As for the word itself, it came into being in 1937, when the Hormel Foods Corporation ran a competition to name their new "spiced meat" product. The usual explanation for how it came to be used to describe UBE/UCE is based around the Monty Python sketch, in which the word is used (and sung) incessantly, so much so that it begins to drown out normal conversation.

There is a claim, though, that using the word in this context predates Monty Python by a good 10 years – the following is taken from the Free On-line Dictionary of Computing:

> *Correspondent Bob White … cites an editor for the Dallas Times Herald describing Public Relations as "throwing a can of spam into an electric fan just to see if any of it would stick to the unwary passers-by."*

Etymology aside, and whatever the precise definition you agree with, there are some key points over which there is generally no argument:

- It is unsolicited
- It is an annoyance
- There is far too much of it

The actual content of spam messages varies greatly. In my experience, by far the highest volume of spam is that promoting online pharmacies, in particular those selling drugs with alleged sexual benefits (mostly for men, I'm afraid!). Dotted between these advertisements are a general background noise of product and service promotion, chain letters and scams, such as:

- Cheap tobacco
- Septic tanks (!)
- Sure-fire plans to enormous wealth
- Pleas for help in transferring vast sums of money between countries
- Joke faxes you can have sent to your friends
- 'Free' ring tones
- Requests for bank and credit card details
- Cheap Web hosting or e-mail services
- Cheap 'bulk-mail-friendly' hosting services
- Cheap domain registrations
- Cheap computers
- Logo design services
- Cheap inkjet printer supplies
- Your VERY OWN .me/.tv/.name/.biz domain

Most of the above are not of great concern – a little common sense with credit cards and bank details, and a healthy level of cynicism in reading promotional literature, and most people won't be affected to any large degree. I'm sure several of these sound familiar to you!

There is, however, an increasing level of soft and even hardcore pornography, or links to it, being sent using spam techniques, much of which would be classed as illegal in several countries. This area concerns me most, especially when we consider today's level of computer literacy amongst children. It horrifies me to think that some of the material I've seen recently, simply because I have an e-mail address that doesn't bounce, could be viewed by children (or even sensitive adults), and it is with this in mind that I have felt compelled to write this article.

The current estimate is that around 75% of all e-mail traversing the Internet is spam, and the flood shows no real sign of abating. In the next few sections, we will look at why this is the case; if we can understand the motives and mechanisms that encourage and support the sending of spam, we are halfway to being able to protect ourselves against it, and perhaps even help in the fight to eliminate it.

3 Why Should We Care?

To most people, spam is little more than a minor nuisance – you may receive only one or two a day, so why should you worry about it, and do what you can to counter it?

One important reason is that although it may only be a minor problem for you now, but unless you use some level of diversionary tactics, the number you receive will increase. If you are in a position whereby you cannot change your e-mail address when it becomes unmanageable, you should consider protecting your address now to reduce problems later.

On a more 'community-centric' front, if all of us did everything we could to counter spam, the problem would naturally be reduced. If the spammers do not receive any revenue from their activities, I'm sure their business would eventually become unsustainable. Unfortunately, it is difficult for most Internet users to see the 'bigger picture', and realise the impact spam has on society as a whole.

There have been various analyses on the financial and social impact of spam, and conclusions vary considerably, but there is a common theme – it costs us money. It is obvious how this could be the case for someone whose mailbox is 95% spam, in that a great deal of their valuable time will be taken up identifying the legitimate 5%.

It is similarly easy to see how 5 minutes per day spent deleting spam by all 500,000 employees in a large corporation adds up to a large amount of lost productivity! However, even if you only receive a few per day, and it takes you mere seconds to get rid of them, there are still ways in which you are being affected financially.

Consider an ISP whose networks are carrying 60% spam. Without that overhead, they could reduce their total bandwidth and hence costs, and pass savings on to their customers. This also affects users of 'free' dial-up ISPs, since the underlying telephone charges will be indirectly affected by the same problem, and so charges will gradually rise.

Consider also the impact of spam containing illegal or offensive material. Any action necessary to bring the culprits to justice will generally be paid for out of public funds, and hence the taxpayer's pocket.

On a more esoteric note, I have seen it mentioned that the spam problem could indirectly lead to inhibition of free speech, as people become increasingly reluctant to talk online lest they become a target of spam.

Another important social aspect is the potential distress and even more serious problems that could be caused by some spam content.

Whatever your situation, there will be a way in which spam can affect you financially, and for this reason I would urge every single person to take action in whatever way they can. This may be as simple an act as making sure your address is not confirmed (see later), or something more proactive, such as reporting the spam to one of the online services (again, see the recommendations later in this article).

Spam really **does** affect us all, even those users that have never received any! Even if your ISP filters your e-mail, and you never receive spam sent to your

address, it is still worth taking the protective measures mentioned later – your small effort, combined with that of others, could make a big difference.

4 Who's Sending It?

Given that the majority of spam is sent with spoofed sender information, it is difficult to determine for certain whom the senders are. Indeed, if it were easy to identify the sender, we wouldn't have the problem we do now, as they would be shut down quickly! From what I've seen, however, I would say there are three main categories of spam sender:

- Organisations with a product, service or information – this type of message is in the minority, since any organisation sending out their own information and accused of spamming will be forced to change their approach rapidly, either through pressure from their connectivity provider (since ISPs receive most complaints), or through direct complaints from the recipients themselves
- Marketing agencies acting on behalf of a vendor – some are legitimate, and don't appear to be aware that they are using marketing techniques that are frowned upon or even illegal. They may also have bought an address list on the understanding that it contained only 'opt-in' addresses. I have, for example, received unsolicited advertisements for fleet car services from an organisation that clearly thought I was interested, based on information they had been sold. Again, these messages are not large in number, as the agencies concerned will generally become aware of the problem quickly.
- Individuals – I would say this is one of the larger sources of spam. Anybody with an Internet connection has the capability for sending spam, and with the promise of a 'fast buck' from doing so, it is likely that there are vast numbers of individuals with an Internet connection and some downloaded software involved in the business.

There is evidence that there are also larger organised groups of spammers out there, fully aware of what they are doing, and working to defeat any protective measures ISPs or users put in place. This includes some developers of mass-mailing software, which is becoming increasingly able to defeat mail filters. The sophistication of many spamming techniques shows a deep understanding of filtering and blocking technology (as discussed later), and the resources being expended on development of mailer programs and address harvesting techniques imply a larger investment than an individual could manage.

There is one final and important consideration (discussed in more detail later) – whilst you may not be knowingly sending spam out, your machine could be doing so without your knowledge, or laying itself open for others to use it as a mail relay. If you have an 'always on' broadband connection, your machine is an attractive target to virus writers and other malicious software authors that may seek to subvert your machine for their own purposes.

5 Why Are They Doing It?

This is one of the more difficult questions to answer – what's in it for the spammer? Given that they have to go to such lengths not to be identified, and to develop filter-defeating technology, how can they justify the expense?

Put simply, the answer is money. There are a few instances where this is not the case (political agendas, for example), but financial gain is the prime motivator for the vast majority of bulk e-mail.

In some cases, the financial benefit can clearly be seen – where a product or service is being sold, for example, or where bank or credit card details are being gleaned through a swindle. When one considers the huge number of people a single message could be sent to, it becomes clear that only a tiny response rate would be sufficient to turn a profit, especially if those messages are being sent using other people's Internet connections!

Sometimes it is less clear how revenue can be generated from a spam message. It may appear to be garbled gibberish, with no links to take you to a marketing Web site, no response details, no telephone number or anything. However, even messages such as these can generate revenue for the sender, especially if the recipient's e-mail client is HTML-enabled. Simply opening the spam message may load an image or script that pays a visit to a Web or other server. There are plenty of services out there where revenue is paid simply for visits to a Web site.

The pornographic spam that concerns me most is profitable for the senders in exactly the same way. They are looking for subscriptions to services, for revenue from hits to Web sites, and even for a simple unthinking click on a 'Remove Me' link. This confirmation that they have reached a working e-mail address will ensure that address is added permanently to a database somewhere, which will later create further revenue for the spammer when the database is sold.

6 How Did They Get My Address?

It is sometimes not obvious how the spam sender got hold of your e-mail address, especially if you are very protective of it (as recommended towards the end of this article). Indeed, I have had the problem in the past where a **very** obscure e-mail address I was using, but had never published, suddenly started receiving spam messages.

There have been claims that dishonest ISPs will sell address lists to 'interested parties', but I do not intend to discuss that, as there is no evidence that this takes place. A separate, related issue may be that some ISPs do not secure their customer registration systems properly, but again I am not aware of any evidence that points to this.

There are, besides, a number of very effective techniques the spammer can use to obtain lists of e-mail addresses, and some of these are discussed next.

6.1 Address Gathering

The normal approach spammers will take is to collect e-mail addresses from a variety of sources – the more obvious ones are presented here, but there are undoubtedly many more possibilities (I'm sure you can think of several!).

It should be noted that these collection mechanisms could well feed into each other. For example, a spammer's address collection may start with a simple list of domain contacts, and over time filter it of deleted domains, enhance it with conference lists and shared databases, add confirmed addresses from dictionary attacks, etc.

6.1.1 Public Sources

There are quite a few sources of e-mail addresses available, quite legitimately, to the public, including (this is not an exhaustive list):

- **Domain Registries** – do you own a domain name, or are you listed as the contact for one? If so, your e-mail address is very likely to be listed for access by anyone that cares to retrieve it. For example, take a look at http://www.dnsstuff.com/tools/whois.ch?ip=example.com - three addresses are listed (all of which happen to be the same. Depending on your registrar, you may not be able to obscure this information (for .uk domains, Nominet kindly do obscure e-mail addresses in some cases).
- **Web Site Contact Lists** – is your e-mail address listed on a Web site? Perhaps it's in a list of addresses of members of a community, or in the archives of a discussion list or forum, or perhaps it's on your own Web site in a 'mailto:' link, or even simply in a comment in the HTML source. See the next section, which discusses how addresses like this are trawled.
- **Directory Services** – directories have always been a popular service, and as interconnectivity increases, so huge quantities of contact information end up in a central location. 'People Finder' services, such as those provided by Yahoo!, 192.com, Lycos and others can be a valuable source of address lists.

6.1.2 Spiders and Bots

Retrieving e-mail addresses from thousands of Web sites would be a long and tedious process if done manually and is clearly not a cost-effective option for the spammer. Hence, software designed to visit Web sites and follow links, collecting information in the process, are used extensively for address gathering. These 'Spiders' or 'Bots' are generally legitimate software used by search engines to

index Web sites but, since most are Open Source (i.e. developed by and provided to the public), they can be rewritten and used to gather address lists very easily.

In this way, huge numbers of addresses, most of which will be active (perfect for the spammer!) can be collected, stored and passed on very quickly. This is why (as mentioned above), a simple 'mailto:' link to your e-mail address on a Web page will very rapidly ensure you start receiving more spam than you can handle.

This technology is by no means limited to Web content. Similar scripts can be used to 'crawl' through newsgroups (Usenet), monitor IRC channels, etc.

Please see later in this article for ways you can protect your e-mail address from misuse.

6.1.3 Purchased and Community Lists

Another valuable source of e-mail addresses is a legitimate one. If you have ever signed up for anything on a Web site, completed a questionnaire online or on paper, attended a conference, or provided your e-mail address to a marketing agency through any other means, you are likely to already be on a spammer's database. This is not necessarily through any fault of the agency itself, as it is relatively easy for a spammer to obtain a copy of these lists, if you have agreed to receive 'relevant information' by e-mail.

Lists of this kind are sold between marketing organisations, quite legitimately if you have agreed to receive targeted marketing literature, and frequently under quite tight control. There is no way for the agency to confirm, however, that the list they sell will not end up in a mass-mailer's possession, and they regularly do.

Another theory, although I know of no evidence to prove or disprove this, is that there exist enormous databases of 'confirmed' (see below) e-mail addresses created, shared and sold by the spammer community. I am confident such databases exist, as it would make simple financial sense to those involved, and would potentially be far more dynamic than the CD-based lists that may be available from marketing agencies. E-mail addresses these days are generally, in large part because of the spam issue, transient – a huge number of users switch addresses regularly as spam levels in their mailbox increase.

Taking this one step further, although potentially a moot point – are these databases secured against theft?

6.2 Dictionary/Brute Force Attack

One effective approach spammers can take is, rather than to gather e-mail addresses from obscure sources, to *generate* them using an exhaustive 'guessing' process.

People (and organisations) like their e-mail addresses to be memorable, and so they commonly consist of firstname@domain.com or first.last@domain.com. So, taking a dictionary of first names, perhaps in combination with common surnames, and combining this with public information on domains, and you can pretty quickly hit an enormous number of active e-mail addresses. This is known as a Dictionary Attack, commonly used for guessing passwords when attempting to break into a system, but equally applicable here.

Of course, many e-mail addresses are less simple – for example, bob67543@hotmail.com - and so another technique is necessary to hit upon

these. Another method used to break passwords is the Brute Force attack, which involves running through all possible combinations of letters and numbers, starting at aaaaaa, then aaaaab, and so forth.

Combining the two techniques would easily catch our HotMail address, as well as many other people named Bob using the same e-mail system – take a common first name, and add all combinations of numbers after it.

Of course, this causes the occasional problem for those of us using 'wildcard' mail forwarding – i.e. anything@domain.com is forwarded to a single mailbox. I've stopped using this on most of my domains now for precisely this reason – when a Dictionary or Brute Force attack is run against such a domain, a mailbox can become swamped in minutes!

The descriptions here are really quite oversimplified, as a simple Dictionary or Brute Force attack would be quickly and easily spotted, and the channel would be shut down rapidly. More sophisticated variations of the techniques seem to be used, such as trying 10 dictionary names at a time for each domain, or distributing the Dictionary or Brute Force efforts across multiple machines to make the attack less noticeable.

You might think that the quantity of bounced e-mails generated by this kind of attack would be colossal, and you'd be right – one of my servers currently bounces hundreds of messages per day! To the spammer, however, this is not a concern – remember that most of these messages have spoofed sender addresses, and so the real sender will never see a bounce.

It is important to spammers, though, that they receive confirmation when they reach a valid e-mail address, so that they can build their databases, and target more accurately the people that may open, read and react to their messages. This leads us on to the next section, on 'Address Confirmation'.

6.3 Address Confirmation

Once an address has been gathered or guessed, and a spam message (or even an almost blank test message) has been sent, how does the spammer receive confirmation that they have reached a valid address? It is important to the spammer to obtain this confirmation simply because sending to invalid addresses costs time and, therefore, money. Increasing the accuracy of their databases will directly increase their revenue.

It is important that we recognise the various techniques used on this front so that we can take steps to protect our e-mail addresses from being confirmed as active to the spammer. Some of these steps we can take, but unfortunately, some are down to our ISPs and their good practices.

Imagine you are sending e-mail you know you won't receive a reply to, since the 'From' address is not real, making 'Read' and 'Delivery' receipts unusable. How would you determine whether it bounced or was delivered, whether it was opened or not, whether it was read or deleted?

The first (and surprisingly the most common) technique is to quite simply ask the recipient's mail server whether the address is real or not! This is a supported option in the underlying mail delivery protocol, SMTP, and poorly configured mail servers will quite happily tell the spammer whether the address they are sending to exists or not. Best practice these days suggests that these parts of the SMTP

protocol (the EXPN/VRFY commands) are disabled on Internet-facing mail servers. Unfortunately, far too many servers are still configured to respond to such requests.

Assuming all the mail servers involved in getting the spammer's message from them to the recipient are well configured (apart, obviously, from the first one in the chain!), and no confirmation has yet been sent back to the sender, the next target is the recipient him/herself. There are a number of methods used to obtain a response of some kind from the recipient, including:

- Most blatantly, the 'Opt Out' link – if you are invited to click on a link in an e-mail to 'remove your address from the database', you can almost guarantee that clicking on the link will do precisely the opposite. The address will be confirmed as valid, and become a permanent fixture in the database.
- HTML-formatted e-mails – the majority of spam messages are HTML-formatted, in general containing images or other special tags (whether you can see them or not). On loading these images from the spammer's Web server (or, rather, the latest server they have hijacked), the e-mail address is confirmed as valid. Even worse, it has been confirmed that the recipient accepts HTML-formatted mail, and is inclined to open all their messages. **Note**: if you have an e-mail client with a 'Preview' function, it is likely that images in HTML will be loaded when using it – you are still opening the e-mail when 'Previewing' it!

There are other, less-common techniques in use to solicit a response from recipients, but in general, the two above are the ones you will see. Note that an HTML message that appears to be empty may still be a probe for valid e-mail addresses.

7 Why Can't Spammers Be Traced?

I have already discussed the fact that if a real 'From' address were used in these messages, the sender would quickly be shut down. You will also be aware that all e-mails contain 'hidden' header information identifying the machine that originally sent them – therefore, the spammer wishing to avoid detection would probably not use their own personal machine to send the messages, or must find a way to fake this information as well.

One method that used to be more common than it is now was to insert a large number of fake mail headers into the e-mail being sent, so that it became difficult to identify which was the real sender. Unfortunately for the spammer, the anti-spam community has gotten wise to that trick, and it is rarely used now. Instead, the prevailing method now is to use services deliberately or accidentally provided by others, or even to illegally take control of insecure machines.

7.1 Open Mail Relay

When you send e-mail, in general it will be sent to your local ISP's mail server (e.g. smtp.yourisp.com or mail.yourisp.com), which will then forward it to the final recipient's mail server. It's as simple as that, and it is the simplicity of this process that has contributed to the widespread use of e-mail for communication. Simplicity, unfortunately, is also one of the reasons spam has become such a problem.

It is easy to set up a mail server, but far more difficult to prevent it being abused.

> To present an analogy – most people have a landline telephone in their house. The only people authorised to use it are the people in that household, and even then, some of them may be subject to certain restrictions (do you have teenage children?). Your house is generally not accessible to outsiders that may make calls on your line, and you trust that the local exchange is secure enough that the line can't be abused.
>
> Imagine, though, if you put a telephone extension in a booth on the street outside your house, with a sign above it advertising free calls to anyone. I suspect your telephone bill would rocket very quickly, and you would soon have your line cut off as news got around and passers-by started making nuisance calls.

A silly analogy, perhaps, but very close to the situation with mail servers. In theory, the only people that should be using any particular mail server (or relay) are the customers of that ISP, or the users on that local network, etc. A badly configured mail server, however, can be used by anybody.

It is also surprising how quickly open mail relays can be discovered. A colleague of mine once inadvertently left their mail server open after upgrading their software, and on reviewing his logs later, it only took 30 minutes for the server to be identified and to start relaying large quantities of spam!

Poor configuration is one issue, and a very serious one, but there are several other reasons why open mail relays might be available on the Internet:

1. Deliberate provision – ironically, some of the spam messages I've seen floating through my overloaded mail server are advertising 'bulk mail friendly hosting services'. In other words, there are ISPs now providing services

targeted at the spammers. Open mail relays, Web hosting of any kind of content, who knows?
2. Poorly written software – there are a large number of software packages that enable system administrators to relay mail for their users. Unfortunately, some of the software I have personally seen, particularly freeware and shareware, provides no facility whatsoever to *stop* mail relaying from unauthorised users!
3. Malicious software – see below for a more detailed discussion, but several recent viruses and Trojans have actually been seen to install open mail relays on infected machines, and even publish their existence to the spamming community.

7.2 Open Proxy

Rather than present here a detailed description of what a proxy server is and how it works, suffice it to say that they are designed as 'go-betweens' and caches for Web or other access. If poorly configured, they can allow a spammer to anonymously use other poorly configured Web-based contact forms or even free e-mail services (like HotMail or Yahoo!) to effectively anonymise their activities.

7.3 Malicious Software

The (gradually) increasing awareness of correct configuration of mail and proxy servers is slowly reducing the 'anonymous' channels available to the spammer. For this reason, other surreptitious techniques have developed, generally focusing around the idea of *creating* these channels where they do not currently exist.

It is still frighteningly easy, even in these days of increased awareness, to propagate a virus, Trojan or worm. Indeed, as I write this, a worm is spreading across the Internet, despite the fact that a recipient has to open an unknown executable attachment received via e-mail from an address they do not recognise!

There have been very recent (and successful) examples of viruses and worms that will turn your machine into an open mail relay or proxy server, and in some cases even announce themselves to the spammer community as 'ready to send'.

There will be more (much more!) on virus protection in a future article.

7.4 Address Hijacking

This technique is very much more obscure, but there is evidence that it is happening. Using various low-level routing protocols, it is possible to temporarily hijack IP addresses that are registered and valid, but not currently in use. With this spoofed address, spam messages can be sent, and then the address released again, with very little possibility of tracing the source of the abuse.

8 What Can Be Done About It?

So far, this article has concentrated on identifying the problem the reasons and technology behind its escalation. This section now presents the methods users can employ to combat it. It must be noted here, as it is at various points throughout this section, that the fight against spam is not something we can win in isolation – it affects everybody, and so everybody must be involved in countering it.

This section is broken into three sub-sections:

- Action you can take to avoid receiving spam in the first place if you don't currently, or have recently started using a new e-mail address
- Action you can take to reduce the amount of spam you receive once it begins (and it usually does!)
- What you can do with any spam you do receive, despite your efforts

8.1 Avoidance

There are a number of preventative measures you can take to avoid receiving any spam at all (although the Dictionary or Brute Force Attack may still get to you) – if you have never received any spam, or have just started using a new e-mail address, then following the guidance in this section will (hopefully) put off that inevitable day.

If you have already started receiving spam, you should move on to the next section, but do still continue to follow these avoidance steps as well.

- Never post to a newsgroup, bulletin board, Web-based forum or similar facility. If you cannot avoid doing so, you should never use your real e-mail address. Instead, configure your newsreader, profile or whatever is appropriate to use a fake address – you could use nullbox@internetgremlin.com if you like, which silently ditches messages it receives! If you require a response to your messages, then try to think of a way to present your address in a way that a human could understand, but an automated address harvester would not recognise. Examples include:
  - user at domain.com
  - user(at)d-o-m-a-i-n.c-o-m
  - user@nospam.domain.com

  There is a risk in using this approach (address 'munging' or obfuscation) is that address harvesters are becoming increasingly sophisticated, and are learning to interpret addresses such as these.
- If you do post to a newsgroup or forum with a spoofed address, don't forget to remove any e-mail signature you may be using, unless it only contains innocuous information.
- If you use an e-mail based discussion list or some other form of 'closed' forum, check whether the list has open Web-based archives, or any kind of feed into a newsgroup. If so, ensure that the technology hides sender's e-mail address, or at least mangles it in some way.
- An alternative approach is to use a 'temporary' throwaway e-mail account (e.g. Yahoo!, HotMail, etc.) to post to facilities such as this, and stop using it as soon as the mailbox becomes filled with spam.
- Following on from this idea, in particular if you want to try to take action against illegitimate sale or distribution of e-mail addresses, you could create an e-mail address 'per purpose. For example, if you visit a trade fair, and give your e-mail address out to people you meet there, you might create an address called trade-fair-2004@somedomain.com. In the event that this

mailbox becomes filled with spam, you can narrow the source down quite easily. For the masochistic, how about a separate e-mail address per contact! This approach also allows you to 'validate' each person you communicate with before giving them your real e-mail address.

- If you have a Web page or even an entire site on the Internet, never put your e-mail address on there, even if it's not hyperlinked, or is in the comments of the HTML, as these are simple targets for address harvesters. It is advisable to use a feedback form of some kind, but please take advice on securing these, as poor forms can become open mail relays in themselves! Also ensure you use a form that does not require you to put your e-mail address in a hidden field in the form, as this defeats the purpose!
- If your mail client, anti-virus or personal firewall software contains built-in functionality for filtering spam, consider enabling it (although do be aware that false positives are always a possibility)

8.2 Reduction

If you're reading this section, you've already started receiving spam because your address has been harvested, has been guessed in some way or some unscrupulous individual has sold it. My heartfelt sympathies go out to you.

If, like me, you are not in a position to ditch the e-mail address concerned and start using a new one, there are some steps you can take towards reducing the tide, which, unless you take action, will increase rapidly.

- If you can avoid it, do not open a spam message, or view it in any sort of preview window or pane – many e-mail clients will render HTML immediately, perhaps loading graphics, and confirming your e-mail address to the spammer. Then, if you don't have the time, energy or inclination to take action against spam (see the next section), then delete the message.
- Never, ever reply to a spam message – there are two possible effects:
  - Malicious configuration of the e-mail headers could cause your reply to be addressed to a large group of people
  - If the sender address is not completely spoofed, and will be received by a malicious server, your e-mail address will be confirmed and added to a spam database.
- If you are in a position to do so (i.e. not using a pre-configured company workstation), configure your e-mail client not to respond to 'Read Receipts', or at least to prompt you before doing so, for the same reasons as the previous point.
- Never click on a 'Remove Me' link, or follow any other instructions of that nature – again, you will most likely be confirming your e-mail address to the spammer, and the flood will increase.

8.3 Reaction

Once you start receiving spam messages, what you do with them depends on your outlook, available time, energy and inclination. If you simply want to get rid of the stuff as quickly as possible, read the first of the following two sections. If you want to try to help in the fight against spam, please also read the second.

8.3.1 Delete and Ignore

For most people, this is going to be the approach to take. Simply delete the e-mail, preferably without opening or previewing it as discussed above (the subject lines are usually easy to spot!), and forget all about it.

The problem with this is that it does nothing to help in fighting the plague – it implies a level of trust that 'somebody else' is taking action somewhere. However, if everybody did that, nothing would change! I strongly recommend at least taking the first step described in the next section, even if you do not have the time or the energy to do any more.

8.3.2 Take Action

There are a number of things you can do to report spam you receive, and try to get some action taken. How far you want to go depends on how much the problem affects you personally, how much time you have available and (at the far end of the scale) your technical expertise.

Unfortunately, I have to tell you that you are unlikely to see the benefit very often – in several years of reporting and taking action against spam (I must be into the hundreds of thousands by now!), I have only had two or three accounts shut down, and one or two open mail relays secured. It's a thankless task, but even that occasional small victory, if you multiply it by the number of people with e-mail addresses, can make a difference.

There are three approaches you can take for run-of-the-mill spam, and these are listed below in ascending order of effort involved. For anything you receive that may be illegal or involve children, there is a different route you should take – please see the next section.

1. The simplest and easiest tactic is to report the message to your ISP or e-mail service provider. Almost all ISPs will provide an 'Abuse' contact e-mail address, and they will usually require you to forward the message with all the headers included. Unfortunately, many ISPs will respond that the message didn't originate from their networks, and so they won't take any action. If your ISP or mail provider tends to respond in this way, you should try the next approach.
2. Using a spam reporting service is the next best thing to personally tracking down abusers (see the next approach). I only have personal experience of SpamCop (www.spamcop.net), but there are other similar services in existence. The idea is that you copy and paste a spam message including headers into a Web-based form (or forward it directly to the service, or use their e-mail services, and so forth), and the system will then analyse the message, strip forged headers, determine the network(s) and/or Web site(s) being used, and report to the relevant administrators. SpamCop offers a limited-functionality free interface, so give it a go – your reports are anonymous if you prefer them to be.
3. The final option for those with a lot of time, energy and technical expertise is to perform the investigation yourself, and make personal contact with the administrators of systems involved in the transmission of any particular message. This involves very careful analysis of mail headers, otherwise you could falsely accuse the innocent. There is also the added complication that the system administrator maybe involved with the spammers, making you a potential target for further abuse. However, some will feel strongly enough to take action in this way, and will find ways to deal with the adverse consequences.

8.3.3 Illegal or Obscene Material

If you receive or see any online content (e-mails, Web sites, etc.) that you feel is potentially illegal, you should immediately report it to the relevant authority. Where exactly you go with it depends on the legislation in your country, and the

organisations that exist, but in the UK, there are several organisations you could approach for advice:

1. Contact the Internet Watch Foundation (www.iwf.org.uk) - the IWF was formed specifically to take action against particular types of illegal content; they focus primarily on (in their own words):
   - Images of child abuse, anywhere in the world
   - Adult material that potentially breaches the Obscene Publications Act in the UK
   - Criminally racist material in the UK

   They will, however, offer advice on what to do with any other type of potentially illegal content, and are likely to recommend you approach one of the next few organisations.
2. Local police – many local police services have specialist computer crime teams.
3. National police – the Metropolitan Police have the national expertise in Computer Crime
4. National Criminal Intelligence Agency – this group liaises with authorities in other countries, and will probably become involved where content is hosted in, or originated from, abroad.

My recommendation would be to use the IWF as your first port of call. If they are unable to help, they will be able to put you in touch with the right people. If the content falls within their remit, however, they react quickly and effectively, contacting the relevant authorities and shutting down the source of the problem.

9 Summary

The situation is grave, but there is some hope. We will probably never achieve total elimination of spam until the world's systems achieve 100% security (!), but all of us can help in reducing the flood to manageable levels.

I feel the problem is best tackled by taking steps to increase the ratio between cost and profit to the spammer to such a degree that it becomes unprofitable to do. Technical work is taking place to tackle the problem from two opposite directions:

1. From the source – securing of open mail relays and proxies, reduction in the impact of virus attacks, additional measures in existing protocols, etc. are aimed at making it difficult (and therefore costly) for the spammer to send their messages in the first place.
2. At the target – filtering and blocking technologies aim to prevent any spam message being received. If they are not read, then the spammer's revenue stream is cut off.

Separately, and perhaps less effectively, legislation is focusing on responsibility and accountability – if we can make the spammer accountable for their actions, there will be less inclination to mass-mail in an untargeted manner. Bringing the law to bear on illegal practices and content will also help in the battle.

I urge all readers to take as many of the actions described in this paper as possible. At the very least, protect your e-mail address at all costs.

10 Further Reading and References

10.1 Other Analyses and Documentation

- NISCC – Technical Note on Spam Mitigation (contains an extensive references section): http://www.uniras.gov.uk/l1/l2/l3/tech_reports/NTN0204.pdf
- Evan Harris' Greylisting Concept: http://greylisting.org/
- Paul Graham: *A Plan for Spam* – http://www.paulgraham.com/spam.html
- Sender Policy Framework (SPF): Internet Draft – http://spf.pobox.com/spf-draft-20040209.txt
- MessageLabs: a paper on Sender Authentication, with something of a sales angle – http://www.messagelabs.com/intelligencenewsletter/march2004/
- On the December 2003 Legislation:
  - BBC: http://news.bbc.co.uk/1/hi/technology/3120628.stm
  - Sophos: http://www.sophos.com/spaminfo/articles/ukspamlegi.html

10.2 Organisations

- SpamCon Foundation (Anti-spam) – http://spamcon.org/
- The Spamhaus Project (Anti-spam) - http://www.spamhaus.org/
- Internet Watch Foundation (legal action) – http://www.iwf.org.uk/

10.3 Products and Services

- Service: SpamCop (filtering and reporting), also includes an exhaustive FAQ and a free limited-functionality service – http://www.spamcop.net/
- Service: MessageLabs (Enterprise filtering) – http://www.messagelabs.com/
- Product: SpamAssassin (filtering) – http://www.spamassassin.org/

10.4 Legislation

- EU Directive 2002/58/EC on Privacy and Electronic communications – http://europa.eu.int/information_society/topics/ecomm/all_about/todays_framework/privacy_protection/index_en.htm
- DTI: Implementation of EU Directive 2002/58/EC in UK law: http://www.dti.gov.uk/industries/ecommunications/directive_on_privacy_electronic_communications_200258ec.html
- CAN-SPAM (US): http://www.spamlaws.com/federal/108s877.html