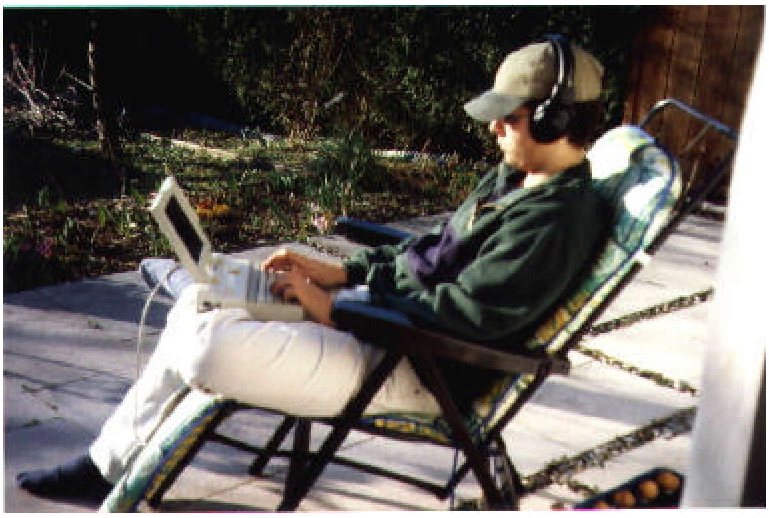
Fighting Spam for fun and profit

by Florian “BlueScreen” Hobelsberger

a user’s guide to reduce spam



”To accept the circumstances like they
are is inappropriate. To change the
circumstances is the solution.”

http://www.IT-Checkpoint.net
http://www.HE-Crew.de

Contents:

1. General Spam Information

1.1. Spam in general

By using the term “Spam” most people think about advertising e-mails invading the own mailbox. It is almost impossible to oversee that the amount of spam arriving in the mailboxes increases every month.
If you widen the term, Spam includes all unwanted e-mails – it doesn’t matter if it is an advertisement or a mail from a person you just don’t like. Also, Spam is usually not sent automated by using programs. The Sender of Spam is from now on referred to as “Spammer” in this document.

1.2. Never trust Spam

You surely sometimes had a look at some Spam-E-Mails, sometimes just because the Subject and the “FROM”-Address mislead you to the conclusion that the received e-mail is an important one. Reading some of the mails sounds really like a good opportunity to earn a lot of money, make really good dealings and so on.
You should never, never never never trust Spam – a person or business e-mailing to you making business proposals without your request is everything but trustworthy. (Of course this is not true if you are in a business and someone requests a product that is provided by you, but i think you know what i mean). Be careful with giving out personal information, including your real name and address. The most important thing is: Do not give out banking information to persons you do not trust or when the connection is not secure). You also wouldn’t trust a person who wants to sell you viagra in dark corner, would you ?

1.3. The Problem with Spam

Spam usually never stays the same – it is “dynamic”. The authors of spam change names, from-adresses, subjects etc very often – this makes identifying spam without opening the message very hard. Sometimes the amount of Spam in a mailbox reaches a level at which the user gets really frustrated. Static Filters usually don’t help. So, how to get rid of the damn Spam ?

1.4. Anti-Spam Software is mostly dumb

As said before, static filters don’t work very well mostly. They help to filter out some of the messages, but many still come through. The Problem with Anti-Spam Software is, that most of it is based on static filters.

1.5. Where did they get my address ?

There exist several methods where obviously addresses could get “stolen”. Mostly people who are very active in the internet by providing a homepage (where they mention their addresses), are active in mailing lists (where their addresses are automatically listed AND which are usually archived on usual homepages), use simple e-mail addresses or sign up at free services. As a conclusion, it could be said that “Spammers” get their addresses by
a) searching the web for e-mail addresses
b) monitoring mailing lists
c) “guessing” e-mail addresses
d) trying every possible combination of characters and numbers (“bruteforce”)
e) buying e-mail address from free service providers.

Since all these “situations” are classified to be “untrusted” locations, you should not give out your real mail address – or you have to fight against a lot of unwanted e-mails. Every enviroment which is not protected by a password or other protection should be considered as “untrustworthy”.

2. Several Approaches to clean your mailbox

2.1. Use “Fake-Mail” Addresses

One approach I started to test is to give out an address i don’t check on all untrusted locations where an e-mail address is needed. Of course this doesn’t make a sense if you want to receive e-mails from persons who get your e-mail address at these locations. But thanks to modern technology, many e-mail providers allow to use “absence messages” – messages that are mailed back as soon as a mail comes in. By using these “absence messages” it is possible to inform the sender of the e-mail about the true mail address. This is based on the fact that usually contributors of Spam don’t read incoming mails. In the bottom line the “Fake-Mail address” is full up with unwanted e-mails that are not read by anyone – the really “interesting” Mails can be sent again to the right E-Mail address. The Problem is that most people won’t be very happy about the need to send the e-mail again.

2.2. Use “E-Mail Forwarders”

Another approach is to use “temporary E-Mail-Forwarders”. You just create temporary e-mail addresses which forward the incoming e-mails to your real account. If the amount of Spam increases too much, you just delete these addresses or deactivate the “forwarder” and begin using the next one. Obviously if you give out “forwarder addresses” and you deactivate them, you will never get e-mails from these Adresses anymore – people sending mails there won’t ever reach you. So, this approach is only useful at registrations where you have to enter your e-mail address once and you are sure that you will never get important mails from the provider.

2.3. Change your Mail Address regularly

This approach is a combination of the first two approaches. If the number of unwanted e-mail increases too much you just will deactivate your primary e-mail address – to inform the Sender an absence e-mail will be sent with the new address. But never forget to “update” all deactivated addresses – it could be quite enervating if you have to follow a trace of several addresses to find the real one.

2.4. Use Filters

If non of these approaches is the right one for you, you have only one choice: Use filters. The most general distinction could be made as this:
a) Blocking E-Mail from specified mail servers (which allow "relaying")
b) Filtering FROM, TO, SUBJECT and BODY

The Problem with Spam is mostly that if you try to automate the process of filtering, you could delete messages you would want to receive – also called "false positives". If you filter out all e-mail coming from servers who allow "relaying" (sending e-mails from this server without authentication) you could also delete really interesting mails. Using filters for "FROM, TO, SUBJECT and BODY" could also remove mails from your mailbox you would have wanted to read.
The result is the following: You always have to find quite a good middle course to eliminate spam and not to kill your friend's e-mails – which results in allowing several spam messages to come through your filters.
But never trust your filters alone if you don't want to loose some valid e-mails – you should at least have a short look at the mails marked as Spam.

3. How to use filters

3.1. “From”

Using “From”-Filters mostly is quite insufficient since “Spammers” almost always change e-mail addresses or use ones that simply don’t exist (to find more about that you should read a bit about SMTP). So, using “FROM”-Filters usually only makes sense if you want to filter non automated e-mails like from people you simply don’t like or if a Spammer uses always the same FROM-Address.

3.2. “To”

Filtering the “To” usually makes more sense than filtering the “From” – but only a few Spam mails can be eliminated by filtering this tag. I am sure almost everyone of you knows e-mails sent to “Undisclosed Recipients” – mostly sent by “Spammers”.

3.3. “Subject”

The Subject is seldom a good indicator for spam since it usually is changed a lot.

3.4. “Message Body”

The message body is the weak point at Spam. Since the Spammers can change almost all other tags but they want to deliver a message, this is where filters become really useful. To configure your filters correctly, you should receive about 10-20 Spam e-mails and “analyse” them – try to find words that are often used in Spam but are very seldom in valid e-mails. For example: Writing down the URLs usually used in Spam mails alone can reduce your spam a lot if you filter for these URLs in the future. I also implemented a few filters searching for the words “Penis, Enlargement, More” (if these 3 words are found in an e-mail it is marked as spam), “Professional, Master, Degree” and “Million, Dollar, Nigeria”. In combination with filtering for about 20 URLs my mailbox is almost Spam-free (not one false positive yet).

4. Which Filter Software to use ?

4.1. Basic Filtering Capabilities in Mail Clients

If you use the filtering capabilities of your mail client you use a (mostly) free (since you already paid for your software or it is free anyway) way to filter out spam. Sadly, not all mail clients support all filter methods or have filtering capabilities at all. Futher, before you can filter with your mail client the whole e-mail has to be downloaded first (at least using POP3) before the filter can start.

4.2. Mailwasher

Mailwasher (http://www.Mailwasher.net) replaces your Mail Client in the first place – it first downloads some information about the e-mails (like Sender, Recipient, Subject and – if you want – also the message body) – and allows you to filter out Spam while not loading it into your e-mail client itself. Further, you don't waste time downloading attachments. For each mail exist three options: Delete (delete the mail at the server), blacklist (add the sender to a list of people whose e-mails will automatically be deleted) and bounce (an e-mail will be rendered which looks like an "this address does not exist"-mail). Further, it allows to add addresses to a "friends"-list whose emails will automatically be accepted in the future.
After checking your incoming emails with Mailwasher and deleting unwanted mails you can start your mail client either automatically or by clicking a button in the mailwasher program.

4.3. "GCF DeSPAM Tunnel"

GCF DeSPAM Tunnel (http://www.gcf.de/start.php?show=projects) uses quite a different approach: It acts almost like a proxy server. After setting up the Tunnel you have to change your configuration data at the mail client. From now on, your mail client first connects to the DeSPAM Tunnel which then catches the mails and also searches for "keywords" – every keyword has a value assigned to it. If the total value reaches a predefined level to the subject will automatically be added an "[SPAM]" in front of the usual Subject. The actual filtering process has again to be made by the mail client. Further, GCF DeSPAM also allows to add addresses to a "friends"-list.

4.4. Spamnet

Spamnet (http://www.cloudmark.com/products/spamnet/) sadly only works with Outlook 2000/XP but seems to be the most efficient approach to kill Spam. Spamnet simply uses a “central list of spam” – every user of Spamnet can add a Spam mail to this list by simply clicking “delete” – a unique fingerprint of this mail will be created, in all mailboxes will this mail from now on be marked as spam.

5. Conclusion

Trying to find out which of the received e-mails is a valid one can be quite enervating or even depressing sometimes. Using my proposed methods can help you to reduce Spam, but it won't solve the problem fully.