Dealing with Spam

What is Spam

Spam is any unwanted, unsolicited email, usually advertising. In some organizations, spam makes up anywhere from 60 to 90 percent of incoming email.

How to Avoid Spam

The best way to avoid getting spam is to protect your email address. Don't put your address on websites. If you need to include your email address, try to spell it out in a way that is human readable, but won't easily be picked up by automated scanning systems, like “kevin AT fnal.gov”. Another trick is to use a picture of your email address on web pages instead of your email address as text.

Avoid Unsubscribe or Opt out Links

Many spam messages will contain links that claim if you click on them will remove you from their mailing list. What you are really doing is verifying that your email address goes to a real person.

Spam Filtering at Fermilab

Fighting spam is a multistage process. All email coming in to Fermilab is first run through our spam filter machines. These machines process the email through a program called Spam Assassin.

Spam Assassin runs a large variety of tests on the email and assigns a score based on the likelihood that an email is spam. This score and some other info is stored in extra headers that are inserted into the email. It does not otherwise modify the email.

From here it passes through two different virus scanners and then on to the final destination on site. For most users, that will be on the IMAP servers. The IMAP servers are configured by default to move any email that Spam Assassin has decided may be spam to a folder called 'Tagged Spam'.

If you don’t see your ‘Tagged Spam’ folder you may have to subscribe to it. With Netscape, Mozilla, or Thunderbird, go to the File menu and select Subscribe. Add a checkmark next to the ‘Tagged Spam’ folder.

You should also be able to access your ‘Tagged Spam’ folder via the webmail interface. It’s a good idea to occasionally delete the spam that collects here so you don’t have to worry about running out of quota.

What Spam Assassin Does

There are two main different kinds of tests that Spam Assassin uses. This first are regular expression tests that checks for various words and phrases that are known to occur in spam. This can include anything from whole phrases like “buy this now” to more esoteric tests like too many $'s or !'s in a message. There are also tests for things like too many html font changes or known methods of pretending an email came from someone's Outlook program instead of a spam sending program.

The other test category is remote blocklists. There are sites that search for and maintain lists of machines that are known to either be spammers themselves, or allow spammers to use their servers, or are hacked or mis-configured to allow relaying of spam. Spam Assassin checks all

received headers in the message against these lists, and if it passed through a suspicious machine, will increase its spam score.

Spam Assassin inserts headers to describe a messages spam status. A message that is not spam will have headers like this:

```
X-Spam-Status: No, score=0.0 required=5.0 tests=none
autolearn=unavailable
        version=3.0.2
X-Spam-Level:
```

Messages that are spam will have headers like this:

```
X-Spam-Flag: YES
X-Spam-Checker-Version: SpamAssassin 3.0.2 (2004-11-16) on hepa1.fnal.gov
X-Spam-Report: *  0.6 UNRESOLVED_TEMPLATE Headers contain an unresolved
template
     *  1.8 HELO_DYNAMIC_ADELPHIA Relay HELO'd using suspicious hostname
(Adelphia)
     *  0.0 RCVD_BY_IP Received by mail server with no name
     *  1.7 MSGID_FROM_MTA_ID Message-Id for external message added
locally
    *  0.2 HTML_90_100 BODY: Message is 90% to 100% HTML
    *  0.0 HTML_MESSAGE BODY: HTML included in message
    *  3.1 HTML_IMAGE_ONLY_04 BODY: HTML: images with 0-400 bytes of
words
Original-recipient: rfc822;kevinh@imapserver1.fnal.gov
X-Spam-Status: Yes, score=7.5 required=5.0 tests=HELO_DYNAMIC_ADELPHIA,
        HTML_90_100,HTML_IMAGE_ONLY_04,HTML_MESSAGE,MSGID_FROM_MTA_ID,
        RCVD_BY_IP,UNRESOLVED_TEMPLATE autolearn=no version=3.0.2
X-Spam-Level: *******
```

The default IMAP server spam filter rule looks for the header “X-Spam-Flag: YES', which is only present if the message is considered probable spam. All messages have an 'X-Spam-Level: ' header with a number of stars that corresponds to a messages spam score. This is handy if you have a mail program that can filter on strings but can't deal with numeric comparisons.

If you wanted to filter messages with a spam score greater than 7, you could have it look for an X-Spam-Level header that contains the string '*******'.

Other ways to fight spam

Many email clients have built in anti-spam features as well. Netscape, Mozilla and Mozilla/Thunderbird all have built in “trainable” spam filters. This means that not only will it filter mail according to built in tests to check for spam, but any mail that you tell it is spam or is not spam will update a word score database so that future emails will be categorized more accurately.

This type of filtering can be extremely effective, if you take the time to correct any mistakes it makes if its filtering. Luckily, with this system, that is usually as simple as clicking on the “Junk/Not Junk” button on the toolbar while reading your email.

For more Info on Fighting Spam

http://computing.fnal.gov/email/spam/ http://spam.abuse.net/
http://www.cauce.org/