**#1. Message subject contains email address or recipient name**: Either the complete email address or part of the email address (the part before the domain) is added to the subject in order to personalize the message and trick the recipient into thinking that it is a legitimate message. For legitimate mails there is no reason to enter the recipient’s email address in the subject, so the presence of this is a pretty sure sign of spam.

**#2. Message body is base64 encoded**: Spammers use base64 to encode the message headers and body so that spam filters are not able to read the content and perform any filtering. Most email clients will decode the message so that the message can still be read by the recipient.

**#3. Sender address contains number or character sequence**: Spammers use automated programs to register thousands of email addresses. Since they are generated in bulk, they often include number or character sequences such as FRfJIrqOpV@hotmail.com or bob36189624@gmail.com. At first spammers used number sequences but when most spam filters started to block these types of addresses they changed to using character sequences which are harder to detect.

**#4. From: and Reply To: address are different**: This is a common feature of spam mails, but it is also very common with newsletters. The importance of this characteristic should be minimized since it is also found in legitimate emails.

**#5. Message contains only HTML body**: HTML messages usually include a plain text version of the email so that recipients with email clients that cannot read HTML can still view the message in plain text. However, many spammers tend to send HTML messages without this plain text body part. This is done to save on size and to force recipients to read the HTML version which automatically opens an image and connects to a web site when the message is opened. Newsletters also tend to send messages without a plain text body part, so it is important to use a white list of allowed newsletters so as not to catch any false positives.

**#6. Recipient’s email address is not in the To: or Cc: fields**: Red Earth Software found this to be the most commonly found characteristic in current spam messages. The reason for this is that the recipient’s email address is hidden in the Bcc: field or X-receiver field, along with a substantial number of other email addresses. Spammers do this in order to conceal the fact that the mail was sent to a large number of recipients, and presumably so as not to publish their email list. Some persons might add recipients to the Bcc: field for sending out â€˜legitimateâ€™ mailings, but these will tend to be of a more personal nature (which you might wish to block anyway) since most professional companies do not use this method for sending newsletters or mailings. Note however that if you do block emails without a local recipient in the To: or Cc: field, you will be blocking all bcc: messages.

**#7. Empty To: field:** This is also typical for spam messages. Because spammers send out bulk emails by entering all recipients in the Bcc: field or X-receiver header, the To: field is often empty. According to the RFC 822, Paragraph A.3.1. (<http://www.w3.org/Protocols/rfc822/> Overview.html), the worldwide standard for the format of email messages, every message is required to have at least one email address in the To: field. Therefore, if this field is empty, this must indicate ‘shady practices’.

**#8**. **Missing To: field:** Emails that have no To: field at all, can quite definitely be considered as spam since this can only happen if done on purpose for spamming reasons.

**#9**. **Missing From: field:** Again the reasoning behind this is to disguise the actual sender of the message.

**#10**. **More than 10 recipients in To: and/or Cc: fields:** Many spam messages contain more than 10 recipients in the To: and/or Cc: fields. This can however also be used for ‘legitimate mailings’ but again these will tend to be of a personal nature (which you might wish to block anyway) since most professional companies do not use this method for sending newsletters or mailings.

**#11**. **Bcc: header exists:** In normal email messages, a Bcc: header does not exist since this is stripped from the mail.

**#12. X-UIDL header exists:** Incoming messages should not have an X-UIDL header since they are only intended for the mail server to stop it downloading messages more than once, for instance when ‘leave messages on server’ is checked. This header would normally be stripped when the message is received. Spammers add an X-UIDL header to try to get the recipient’s mail server to download multiple copies of their message and therefore increase the chance that the message will be read.

**Spam characteristics % of researched mails**

Recipient address not in To: or Cc: field 64% .64

To: field is missing 34% .34

To: field contains invalid email address\* 20% .20

No message ID 20% .20

Cc: field contains more than 15 recipients 17% .17

From: is the same as the To: field\* 6% .06

Cc: field contains more between 5-15 recipients\* 3% .03

To: field contains more between 5-15 recipients\* 2% .02

To: field contains more than 15 recipients 1% .01

Bcc: field exists 0% .00

To: field is empty 0% .00

From: is blank or missing 0% .00

Net=1.67