A method for spam blocking

Les Hatton
CISM, University of Kingston*

July 15, 2008

Abstract

Spam (unsolicited e-mail) can be split into two forms. The first is where the address of the spam is explicit and accurate, in other words the originator of the e-mail is prepared to be visible and contactable directly. This is very similar to the unsolicited mail which arrives by normal mail delivery methods. The second and by far the most common, is where the originator masquerades as a legitimate company in order to sell products or services through a third address embedded in the body of the e-mail either as an e-mail address or as a web reference.

This paper proposes a method of adjusting Mail sending and receiving software to prevent the second most numerous class of spam whereby a legitimate user's e-mail identity is hijacked.

1 Overview

This paper splits spam into two forms:-

1. Spam in which the originating address is genuine and belongs to the person proposing a product or service.

2. Spam in which the originating address is genuine but does not belong to the person proposing a product or a service. Instead the address has been hijacked to attempt to get through spam blocking procedures and the real response address is hidden within the body of the message either as an e-mail or as a web-site reference. The vast majority of spam received by the author (more than 99%) falls into this category and usually involves selling products which would be considered dubious at best and very likely offensive to most.

This paper deals specifically with a method to block the second category above by an encyphered verification system built on top of the normal e-mail protocol. Blocking the first category will not be considered further here.

*L.Hatton@kingston.ac.uk, lesh@leshatton.org

2 The method

Let S be the genuine e-mail address of a sender and R be the genuine e-mail address of a receiver. Typically a spammer will masquerade as S in representing itself to R in order to introduce a third contact point T, the address of the service or product being foisted on the receiver. The masquerading is done to make it more difficult to detect using simple blacklists. Furthermore, let K(P,M1,M2) be the one-way encyphered address of two e-mail addresses M1, M2 and a password or phrase, P. In this sense one-way means that the mapping P,M1,M2 -¿ K(P,M1,M2) for a given mail-address or addresses is as is used for example in verifying passwords on operating systems such as Linux.

Let Ps, Pr be the sender's and receiver's keys. The method works by generating an encyphered key using the From: and To: e-mail addresses and some password P locally defined and known only to the sender as follows.

2.1 Typical transaction

e-mail sender Prefix K(Ps,S,R) to the address in From: field.

e-mail receiver

```
if key is absent in From: field
  Process as normal e-mail.  Ultimately, reject.
else
  Generate K(Pr,S,R) and check against each key
  if 1 key in From:
    if match
      Read mail
    else
      Add K(Pr,S,R) to existing key in From: field and return to sender
    endif
  elseif 2 keys in From:
    if match
       remove matching key and return to sender
    else
       trash mail
    endif
  else
    trash mail
  endif
endif
```

Here, matching a key means encoding the passphrase and the S and R e-mail addresses and generating the same key.

An example is shown in Figure 1. A simple transaction requires 3 mails rather than one but this would be a small overhead if it prevented category 2 spam as defined above.

Figure 1: An example of a simple sender - receiver mail transaction.

Note also that this method is immune to man-in-the-middle attacks, so although a spammer T could extract an authentication key from R, this key would not work in any transaction between S and R. Note finally that this method works with CC and BCC files as it composes keys on the From: line only.

2.2 Key generation

The quality of the key function K() need be of much lower cryptographic grade than those used to hide content. All that is necessary here is to make it too expensive to crack the keys.