Use and Abuse of Anti-Spam White/Black Lists

September 26, 2006

White and Black lists are standard spam filters. Their typically simple interface, provide a way to quickly identify emails as legitimate or spam.

Any piece of information within the email, or provided by the SMTP protocol, could be useful on identifying emails. The information used for this purpose often includes the originating host IP, sender address, recipient addresses, email subject and the email body.

In general configuring these lists is very intuitive. It just involves identifying a piece of data based on which an email may be classified with a reasonably high degree of certainty. This is then inserted in the appropriate white/black list. However it is best not to get carried away. It is not unusual to come across administrators trying to block all spam through black lists. Today we look at how to use these filters in an effective manner. We identify the criteria to follow and some things to avoid.

Fundamentals

There are some very basic rules that are worth remembering when configuring white and black lists.

1. If we white list all legitimate emails, everything else is spam.

2. Safe guarding the delivery of legitimate emails is more important than ensuring no spam reaches the inbox.

3. The effectiveness of a white/black list is dependent on the reliability of the email information against which it is applied.

White listing all legitimate emails is obviously not practical. However the first point stresses the ripple effect that is achievable. These lists are typically the first stage of a multi-layered filtering setup. Let’s say we are able to immediately identify a good proportion of legitimate emails and spare them from going through further filtering. The filtering layers that follow will see less legitimate emails. Thus there is immediately a lowered risk of false positives. In turn this allows for more aggressive spam filtering.

The second point sets clear our priorities. A spam filtering system must keep false positives (misclassification of legitimate emails) to a minimum. This is true even at the cost of leaving some spam unfiltered.

The last point reminds us of spoofing. Spammers are always trying to give a legitimate look to their emails. A white/black list needs to target the information that leads to the most reliable matching results.

White/Black Listing by IP

IP white listing is great when dealing with hosts known to only generate legitimate emails. There are various applications that use emails for reporting, to manage workflows, to deliver faxes etc. These are typically perfect candidates for this list.

The IP is widely regarded to be the most reliable piece of information available to us. It is directly determined from the connection established between sending and receiving hosts. However this is only possible as long as the sending host directly connects to the host enforcing the IP white/black list. Otherwise IPs may be determined through the Received email header.

Each host involved in routing the email from source to destination adds a Received header to the email. This header includes the IP of the last host from which the email was sent. Thus a complete set of Received headers builds up, tracing the delivery route.

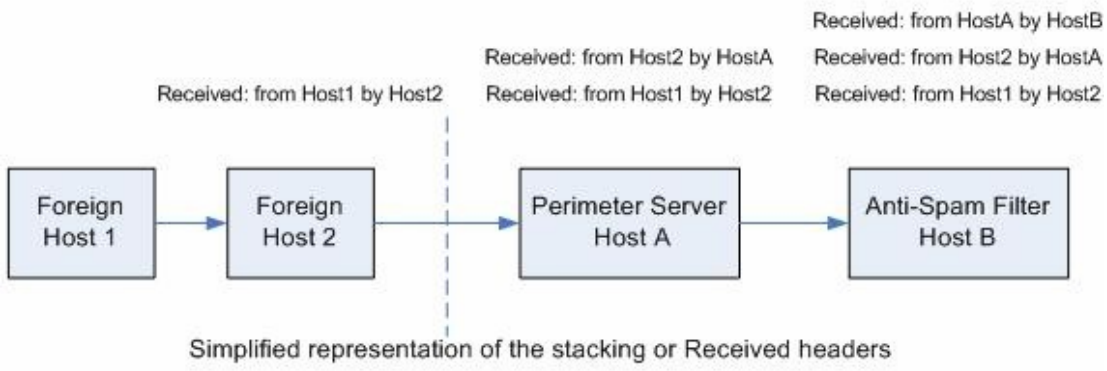
As an example here are the headers taken from a spam email. The first header is the one last inserted.

Received: from pcp0010520431pcs.prshng01.fl.comcast.net ([69.244.215.149]) by serv-box1.exchangeinbox.com with Microsoft SMTPSVC(6.0.3790.211); Thu, 10 Feb 2005 20:02:30 +0100

Received: from gateway-r.comcast.net by pcp0010520431pcs.prshng01.fl.comcast.net with HTTP; Thu, 10 Feb 2005 11:38:26 -0600

Received: from 154.5.87.115 by pcp0010520431pcs.prshng01.fl.comcast.net [69.244.215.149] with Microsoft SMTPSVC(5.0.2195.6713); Thu, 10 Feb 2005 11:37:30 -0600

One may wonder about the trustworthiness of Received headers. After all can't headers be forged? As the following diagram illustrates, this is not a problem:



Simplified representation of the stacking or Received headers

Here the spam filter (Host B) sits behind a perimeter server that handles internet originating email. The connection IP will always be that of Host A and is of no use to us. We are instead interested in the IP of the last foreign host (Host 2) connecting to our perimeter server. This is what Host A sees and inserts in its Received header. Furthermore, since Host A is our own perimeter server, the Received header is certainly reliable.

The inability to forge IPs is one of the reasons why so many anti-spam technologies rely on this piece of information. However there is a catch worth keeping in mind, IPs may change. In case of white listing, this can be an issue especially when the white listed host is not under our control.

The issue of changing IPs is a lot more acute in case of IP Black listing. Today spammers are making use of zombie machines. These are hacked machines utilized for the delivery of spam and malware. There are so many such machines, that spammers may not reuse the same one again for a very long time.

White/Black Listing by Sender

The sender address is one of the least reliable email information elements. To see how easy it is to spoof it, just configure an Outlook Express email account with whatever sender address.

Spoofing certainly minimizes the usefulness of sender black listing. However there are still cases when blocking by sender is effective. Often this filter is employed to block the broader class of unwanted emails rather than just spam. For example, sender blocking does a good job in filtering “legitimate” newsletters. By definition if the newsletter follows a proper subscription/cancellation procedure, it is not spam. However this is a minor detail once an organization wants to disallow its distribution.

Another similar example is when emails are blocked by domain suffix through wildcards. If an organization decides not to be interested in emails from certain countries, it can block all senders whose address ends with the country suffix. Again this blocks a whole class of emails rather than targeting spam directly.

The bottom line is that when dealing with sender white and black listing, we are often identifying senders who are not hiding their identity. Spammers usually do not fall in this category.

White/Black Listing by Recipient

In general we can categorize recipients in two:

- Valid addresses for which our organization hosts a mailbox or allows delivery
- Invalid addresses and addresses to other foreign domains

Clearly the first step is to eliminate the second recipient address category. This is not the job of classic recipient black listing. Instead these are filtered by blocking addresses not present in Active Directory. This functionality is supported out of the box in Exchange 2003. Furthermore foreign domains are blocked by disallowing relaying.

Recipient lists are meant to deal with addresses falling under the first category. White listing a valid recipient disables spam filtering, allowing all emails to reach the inbox. This is sometimes done for mailboxes receiving very critical emails. If we don't want to risk a single false positive, not even one in a million, then this list does the job.

Likewise a recipient black list blocks the delivery of all emails addressed to a specific recipient. This is commonly used for mailboxes that are only meant for internal use.

Recipient black lists may also be handy when used in combination with other white lists. Consider a mailbox that is only meant to receive emails generated by a web hosted feedback form whose subject is fixed. White lists should always take priority over black lists. Thus by white listing the email subject and black listing the mailbox address, we effectively block all emails except those matching the feedback form subject. Of course it would be best if the subject were to be fairly unique.

White/Black Listing by Subject/Body

There are no limitations to what the email subject and body may contain. This is the key characteristic that must be dealt with when populating lists targeting the email content.

Trying to black list all possible permutation of a single keyword is a hopeless feat. A quick look at how the word pharmacy is being expressed in the latest spam wave should clarify this point (PHAxquRMACY, PHoizARMA etc).

Furthermore, spammers can count on a whole slew of other tricks including images, hidden text etc. This is why complex technologies like SmartScreen, the one behind the Microsoft Intelligent Message Filter (IMF), were developed. Thus the bulk of content based filtering should certainly be left to engines specifically built for this purpose. The role of black listing should be that of fine tuning the core filter.

Content white listing can be useful when spammers start targeting your business sphere. If you happen to sell the same products spammers are pushing, white listing could instruct the core content filter to allow those emails through. Otherwise filtering technologies not based on content provide an alternative solution.

Content white listing may also be employed in a precautionary manner. For example we could simply list the product names and services our organization provides, without waiting for any false positives. This works quite well as long as you don't sell the big brand names targeted by spammers.

The most important fact about content white/black listing is certainly the selection of keywords and phrases to be matched. When using Google it is quite obvious that short generic phrases are to be avoided. The same rule applies here:

- Multiple word phrases should be preferred to single keywords
- Single keywords should be longer than 5 characters
- Many short words are also sub-strings in other words, use whole word matching whenever possible

These are the basic keyword selection rules. Of course each filter may provide additional functionality, allowing for more accurate content matching. It is certainly worth checking the filter documentation. Getting these keywords wrong is a lot easier than many think.

Final Tips

The different pieces of information extracted from an email have a varying level of reliability. Understanding these characteristics allow us to avoid the tricks employed by spammers.

Armed with this knowledge, white lists become an effective tool for legitimate emails to bypass filtering. Black lists allow us to quickly trap any spam that would otherwise manage to go through unfiltered.