Subject: Order of filters
Posted by Markus Schmitz on Sun, 19 Aug 2007 17:26:16 GMT
View Forum Message <> Reply to Message

Hi again,

Are there any guidelines in regards to a reasonable order of the filters?

My current logic goes as follows:

a) SpamTrap filter
This goes first. All email caught here is definitive spam and does neither be protocolled nor further processed

b) Anti Virus Filter

c) Msg Store
All following filters are kind of fuzzy and might have false positives. Therefore I use here a mail store to backup all messages in case a user complains about missing expected emails.

d) any other filters

Does this make sense. Did I miss anything major?

Regards

markus

Subject: Re: Order of filters
Posted by support on Mon, 20 Aug 2007 18:09:34 GMT
View Forum Message <> Reply to Message

If the Spam Trap is the first filter, it will learn messages that SpamAssassin will never see, because they are always removed by other filters (ie. Anti-Virus) that precede SpamAssassin in the pipeline.

To avoid clogging your Bayes DB with virus messages, the SpamTrap should be placed right before the SpamAssassin filter. This is the default position of the SpamTrap filter.

Subject: Re: Order of filters

Posted by Jorge on Mon, 19 Jan 2009 18:47:12 GMT
View Forum Message <> Reply to Message

There are problems with Customer Support's suggested order... According to the documentation, and I agree, SpamAssassin should be at the end of the list because it is very process intensive. This means that email destined for the Spamtrap, which is 100% spam, has to go through all the other filters wasting unnecessary resources. Bad, bad idea.
BytePlant needs to take another look at this one, along with allowing users to freely rearrange the order of filters through the GUI, without being forced to edit text files.

In my many years of dealing with spam with tons of different products, I have learned that the ideal filter order is something like this:

1) Delete all email containing malicious code or viruses. This is important since other filters may accidentally trigger them. You can NOT trust whitelist senders at this point, since most junk mail is spoofed. Of course, you would not know this by looking at NST's default filter order.
2) Use 100% spam from blacklists to train Bayes and immediately delete.
3) Use 100% spam sent to honeypot addresses to train Bayes and immediately delete.
4) Archive all mail if needed. Archiving may be done in step 2 only if you need to keep a nice collection of pure spam for future Bayes training.
5) Accept Whitelist senders and skip any further processing. It is counter-productive to have this filter any earlier or any later. All filters beyond this point should not delete emails, but flag them for redirection to a suspect spam account.
6) DNSBL and RBL filtering
7) The rest is up to your preferences and the load on system resources imposed by the filters used.

By the way, the Spam Trap filter in NST fails to delete messages after they have trained the Bayes database. Sometimes these messages make it all the way to my inbox if the other filters fail to detect them. This should NEVER be. Spamtrap email should always be deleted after Bayes training, unless it is needed for archival purposes.

Jorge

Subject: Re: Order of filters
Posted by support on Mon, 26 Jan 2009 15:46:41 GMT
View Forum Message <> Reply to Message

Jorge wrote:

[...]
>
> 1) Delete all email containing malicious code or viruses. This
> is important since other filters may accidentally trigger them.
> You can NOT trust whitelist senders at this point, since most
> junk mail is spoofed. Of course, you would not know this by
> looking at NST's default filter order.

The AV and attachment filter ignores the whitelist by default. You are right, this is not immediately obvious.

> 2) Use 100% spam from blacklists to train Bayes and immediately
> delete.

Whatever can be filtered by a static blacklist, does not need to clog up the spam database, as the spam filter won't ever see these messages. Static blacklists are constantly out of date and don't help a lot anyway, except in some rare cases.

> 3) Use 100% spam sent to honeypot addresses to train Bayes and
> immediately delete.
> 4) Archive all mail if needed. Archiving may be done in step 2
> only if you need to keep a nice collection of pure spam for
> future Bayes training.

Agreed.

> 5) Accept Whitelist senders and skip any further processing. It
> is counter-productive to have this filter any earlier or any
> later. All filters beyond this point should not delete emails,
> but flag them for redirection to a suspect spam account.

See comment 1).

> 6) DNSBL and RBL filtering

This uses less resources than AV filtering, so I don't agree, this should be one of the first filters.

> 7) The rest is up to your preferences and the load on system
> resources imposed by the filters used.

Agreed.

> By the way, the Spam Trap filter in NST fails to delete
> messages after they have trained the Bayes database. Sometimes
> these messages make it all the way to my inbox if the other
> filters fail to detect them. This should NEVER be. Spamtrap
> email should always be deleted after Bayes training, unless it
> is needed for archival purposes.

With the default configuration, messages are deleted immediately after training. It might be possible to configure the filter in a way so messages aren't deleted, but why should anyone?