
Subject: Teaching SpamAssassin

Posted by [Mark Haller](#) on Tue, 17 Aug 2004 10:20:27 GMT

[View Forum Message](#) <> [Reply to Message](#)

Hi there,

A question which is really bugging me at present and am posting here in the hope that someone else has found an answer as I continue to scour the net ...

I want to teach SpamAssassin of ham eMails that were incorrectly flagged as SPAM.

All tools/utilities to do this point at a mail folder.

I want to point a text file at the sa-learn, as I am using Lotus Domino and can easily spit out a text file (or continuous text string) to the sa-learn.

Has anyone come across this?

Regards

Mark Haller
LogicSpot
UK

Subject: Re: Teaching SpamAssassin

Posted by [support](#) on Mon, 23 Aug 2004 13:35:10 GMT

[View Forum Message](#) <> [Reply to Message](#)

> A question which is really bugging me at present and am
> posting here in the hope that someone else has found an answer
> as I continue to scour the net ...

>

>

> I want to teach SpamAssassin of ham eMails that were
> incorrectly flagged as SPAM.

>

> All tools/utilities to do this point at a mail folder.

>

> I want to point a text file at the sa-learn, as I am using
> Lotus Domino and can easily spit out a text file (or continuous
> text string) to the sa-learn.

Teaching a single message as spam:

```
cd [NoSpamTodayInstallationPath]
```

sa\salsa-learn -C sa\ruleset --spam

with containing the message you want to teach.

Subject: Re: Teaching SpamAssassin

Posted by [Mark Haller](#) on Mon, 23 Aug 2004 13:54:55 GMT

[View Forum Message](#) <> [Reply to Message](#)

Hey there,

Thanks for the reply - sounds great so far but you said with MessageFile containing the message ... do you have a sample format of such a file? or can you point me to a URL that I can learn more about this?

Sorry to bother you - I am not sure how your support extends to sa learn and so on, or whether you guys only support your core tool?

Regards

Mark Haller
LogicSpot

Subject: Re: Teaching SpamAssassin

Posted by [Mark Haller](#) on Fri, 27 Aug 2004 05:49:56 GMT

[View Forum Message](#) <> [Reply to Message](#)

Hi guys,

I know you're busy but I could really use a reply to the above message crack this baby and we're purchasing! Just got to work out license type then! So many mailboxes in use!

Mark

Subject: Re: Teaching SpamAssassin

Posted by [support](#) on Fri, 27 Aug 2004 08:49:33 GMT

[View Forum Message](#) <> [Reply to Message](#)

> I know you're busy but I could really use a reply to the
> above message crack this baby and we're purchasing! Just

> got to work out license type then! So many mailboxes in use!

The messages have to be in the common RFC 2822 format, see <ftp://ftp.rfc-editor.org/in-notes/rfc2822.txt> for details.

Subject: Re: Teaching SpamAssassin

Posted by [InforMed Direct](#) on Mon, 13 Sep 2004 10:04:30 GMT

[View Forum Message](#) <> [Reply to Message](#)

> with MessageFile containing the message ... do you have a

The message file is the plain internet ASCII version of the email. But as you have found, it's often darn hard to get this format of file out of a mail system. In the Outlook/Exchange world, there isn't a simple "save-as erl" function. We have to go around the houses by asking users to move (not forward!) spam emails into a shared folder. Then once a day, an admin person downloads these into Outlook Express and then we use a free utility DBextract which can export an Outlook Express inbox back into the right format for feeding through sa-learn.

The same process could be used for ham although we find that as NST/SA automatically passes non-spam email through the learner, we don't have to worry about ham too much.

I think that the latest version (paid upgrade) of NST has a mechanism whereby users can forward spam email to a special address which NST then automatically passes through the sa-learn process as spam. We've not got this version so can't comment more.

Cheers, Rob.

Subject: Re: Teaching SpamAssassin

Posted by [Mark Haller](#) on Mon, 13 Sep 2004 10:13:32 GMT

[View Forum Message](#) <> [Reply to Message](#)

Hi Rob,

Wow, sounds like a convoluted way that you have to manually intervene. Real shame :(Thinking on your feet there!

Can someone from Byteplant confirm Rob's final point - the new version has this functionality available - as if that's the case, the developer working on this can ditch his work! He's having real

problems trying to extract a valid format as well - simple mails are easy but more complex ones that have come in, in different formats, are proving tough. We've just bought the full version after careful thought so hope this in!

Cheers,

Mark

Subject: Re: Teaching SpamAssassin

Posted by [InforMed Direct](#) on Mon, 13 Sep 2004 10:16:24 GMT

[View Forum Message](#) <> [Reply to Message](#)

> Can someone from Byteplant confirm Rob's final point - the
> new version has this functionality available - as if that's the
> case, the developer working on this can ditch his work! He's

I infer it's the honeypot mechanism mentioned on the release notes:

<http://www.byteplant.com/company/news/04081901.html>

Cheers, Rob.

Subject: Re: Teaching SpamAssassin

Posted by [Mark Haller](#) on Mon, 13 Sep 2004 10:19:15 GMT

[View Forum Message](#) <> [Reply to Message](#)

Hey Rob,

You know what - you'll be right there, then!! :)

I was thinking Winnie the Pooh and the Te of Piglet (great book by the way) among others and didn't even realise!

The beta testing starts a week on Monday so we'll see how well it fairs! :)

Thanks again,

Mark
LogicSpot

Subject: Re: Teaching SpamAssassin

Posted by [support](#) on Mon, 04 Oct 2004 15:22:16 GMT

[View Forum Message](#) <> [Reply to Message](#)

Now, to finally say something on this subject:

Learning forwarded mails is at best useless, at worst even counter-productive:

Upon forwarding, both mail clients (and mail servers) modify mails in unforeseeable ways. Both sa-learn or NoSpamToday! are unable to undo the changes your mail client, and the forwarding option you've chosen, may have introduced.

With version 2 we intend to introduce a mail storage. This mail storage can be used to cache incoming mails, in a pristine and unmodified condition, and in a suitable format for learning.

To learn messages, the mail administrator would collect storage IDs of messages to learn as either ham or spam, and use those IDs to sa-learn the offending mails from the cache.

This is not yet the whole way, but at least you no longer have to trouble yourself with getting mails exported and converted into MIME format from proprietary database formats.

Subject: Re: Teaching SpamAssassin

Posted by [InforMed Direct](#) on Tue, 05 Oct 2004 15:54:01 GMT

[View Forum Message](#) <> [Reply to Message](#)

> Learning forwarded mails is at best useless, at worst even
> counter-productive:
> Upon forwarding, both mail clients (and mail servers) modify
> mails in unforeseeable ways. Both sa-learn or NoSpamToday! are

That's why we don't forward spam, we *move* it.

Cheers, Rob
