

Anti-Spam Service

User's Guide

CyberSync Consulting
December 3, 2004

Revision History:

This is version 1.2.2 of the Anti-Spam Service User's Guide. All updates to this document will be noted below.

Version 1.2.2	Added Whitelist section
Version 1.2.1	Updated the call to spamc to specify the user via \$LOGNAME. Escape asterisks when identifying the X-Spam-Level when directing to /dev/null.
Version 1.2	Updating to reflect the appropriate changes in the SMT 4.1 interface, as now the SMT 4.1 has control settings for base configurations of the Anti-Spam Service.
Version 1.1.1	Updated the lock file line to have the colon (:) after the <code>fw</code> .
Version 1.1	Updated to reflect documentation for deletion of mail, not just marking spam email and also the frequency of automatic restarts of spamd.
Version 1.0	Original Documentation

Preface:

This document is the user manual for the Anti-Spam Service offered by CyberSync.

Target Audience:

CyberSync Customers who use a web hosting account (**Linux only**).

Table of Contents

1) <u>Background</u>	3
2) <u>Spam Identification Tactics</u>	3
3) <u>Configuration and Configuration Files</u>	3
4) <u>Filtering with an email client</u>	6
5) <u>Deletion of email on the server</u>	6
6) <u>Whitelisting – Allowing Email through filtering</u>	7

Background

The Anti-Spam Service is a mail filter to identify spam. Using its rule base, it uses a wide range of heuristic tests on mail headers and body text to identify "spam", also known as unsolicited commercial email. This document provides basic instruction on how to configure the application as necessary. By default, no configuration is necessary to use this application (once ordered and provisioned).

Spam Identification Tactics

The spam-identification tactics used in this service include:

- **Header analysis:** Spammers use a number of tricks to mask their identities, fool you into thinking they've sent a valid mail, or fool you into thinking you must have subscribed at some stage. The Anti-Spam service tries to spot these.
- **Text analysis:** Again, spam mails often have a characteristic style (to put it politely), and some characteristic disclaimers and CYA text. The Anti-Spam service can spot these, too.
- **Blacklists:** The Anti-Spam service supports many useful existing blacklists, such as mail-abuse.org, ordb.org or others.
- **Razor:** Vipul's Razor is a collaborative spam-tracking database, which works by taking a signature of spam messages. Since spam typically operates by sending an identical message to hundreds of people, Razor short-circuits this by allowing the first person to receive a spam to add it to the database -- at which point everyone else will automatically block it. For the Enterprise version, this is automatically set to grow as necessary. For Shared solution customers, this is capped at 15MB in size as not to contribute towards over disk usage.

Configuration and Configuration Files

The Anti-Spam service requires very little configuration; you do not need to continually update it with details of your mail accounts, mailing list memberships, etc. It accomplishes filtering without this knowledge, as much as possible. The Anti-Spam service typically differentiates successfully between spam and non-spam in between 95% and 99% of cases, depending on what kind of mail you get.

When the Anti-Spam Service is provisioned to an account, it will enable the web-based configuration so it can be used through the SMT 4.1. To utilize the web-based feature, log into the SMT 4.1 (cpanel) at <http://yourdomain/cpanel/> or <http://your.ip.address/cpanel/>. Click on 'Additional Services', then 'Manage', and then 'Anti-Spam Service'. You'll see a screen much like what is below in Figure 3-1.

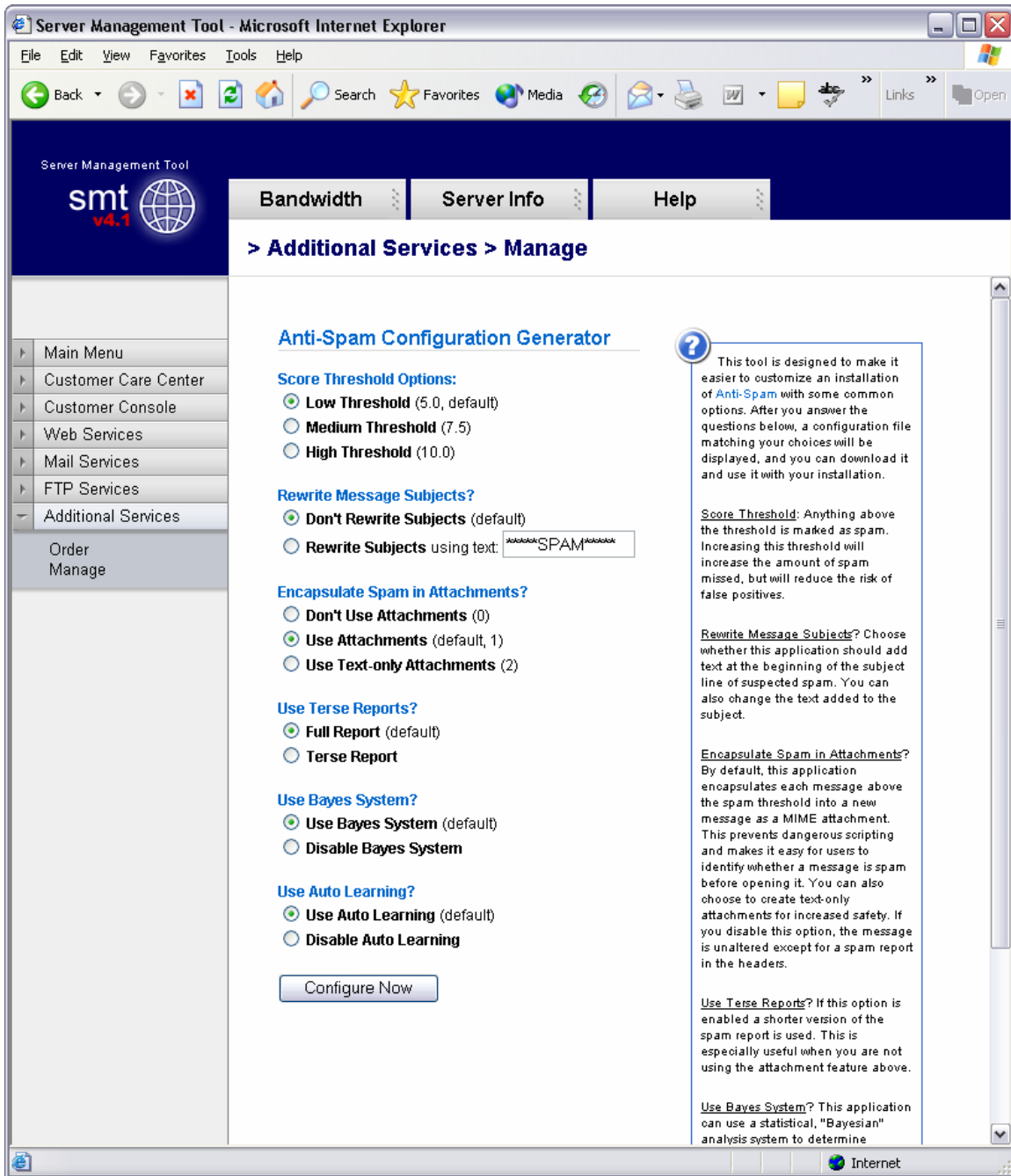


Figure 3-1

The configurations that can be changed are as follows:

- **Score Threshold** - Anything above the threshold is marked as spam. Increasing this threshold will increase the amount of spam missed, but will reduce the risk of false positives.
- **Rewrite Message Subjects** - Choose whether this application should add text at the beginning of the subject line of suspected spam. You can also change the text added to the subject.
- **Encapsulate Spam in Attachments** - By default, this application encapsulates each message above the spam threshold into a new message as a MIME attachment. This prevents dangerous scripting and makes it easy for

users to identify whether a message is spam before opening it. You can also choose to create text-only attachments for increased safety. If you disable this option, the message is unaltered except for a spam report in the headers.

- **Use Terse Reports** - If this option is enabled a shorter version of the spam report is used. This is especially useful when you are not using the attachment feature above.
- **Use Bayes System** - This application can use a statistical, "Bayesian" analysis system to determine whether each message is spam based on previous examples of spam and non-spam.
- **Use Auto Learning** - This application can automatically train its Bayes database by analyzing messages that have a score that strongly suggests that they are spam or non-spam.

For those who wish not to use the web-based interface, you will need to know the file structure of the Anti-Spam Service. On a default installation, the following directories are provisioned:

```
/.spamassassin
/etc/mail/spamassassin
```

Files located in `/etc/mail/spamassassin`:

a.) **bogus-virus-warnings.cf** - This rule set tries to pick out collateral spam caused by viruses. The goal is to catch warnings generated by virus scanners along the lines of "you sent us virus", which are sent to the (usually faked) senders of virus-infected e-mails.

b.) **chickenpox.cf** - Rule set designed to catch spam like "l.ooks f|or th.is kind of sy-ta.x"

c.) **local.cf** - The file in which you are able to customize your settings, if you have the Enterprise version. Below is an explanation of the settings in this file:

- **rewrite_subject 1** (Whether to change the subject of suspected spam)
- **report_safe 1** (Encapsulate spam in an attachment)

```
# [CUSTOMER_IP] is replaced with IP of customer's server
trusted_networks [CUSTOMER_IP] 127.0.0.1
```

- **ok_languages en es de** (Mail using languages used in these country codes will not be marked as being possibly spam in a foreign language)
- **ok_locales en** (Mail using locales used in these country codes will not be marked as being possibly spam in a foreign language)
- **use_bayes 1** (Enable the Bayes system)
- **auto_learn 1** (Enable Bayes auto-learning)
- **required_hits 8.0** (How many hits before a message is considered spam)
- **whitelist_from** (Allow email or domain – From header)
- **whitelist_from_rcvd** (Allow email or domain – Received header)

d.) **sa-blacklist-uri.cf** - Contains a large set URI's (universal resource identifier) that have been found in spam messages. A URI is the generic set of all names and addresses which are short strings which refer to objects (typically on the Internet). The most common kinds of URI are URLs and relative URLs.

Filtering with an email client

Once identified, the mail can then be optionally tagged as spam for later filtering using the user's own mail user-agent application (i.e. Microsoft Outlook or Outlook express, etc.) The email is marked up with the subject line like follows:

```
*****SPAM***** Official Prize Claim Form
```

The "*****SPAM*****" part of the subject line is added by CyberSync's Anti-Spam Service. User's can setup a filter to move emails with this subject line into a temporary or deletable folder within the mail client. The "Official Prize Claim Form" is the original subject line of the email.

Deletion of email on the server

This function is currently only available on our managed dedicated server offering (I-MHIP Enterprise, Business Enterprise, etc). The function is there to delete suspected spam (that fits the criteria set forth by your spam engine), but remember that the Anti-Spam Service is an engine, not just a filter. It is recommended to mark the potential spam, and put those email in a different directory (named something like Junk or Junk Mail) in your email client. An example of this is listed below in Figure 4-1.

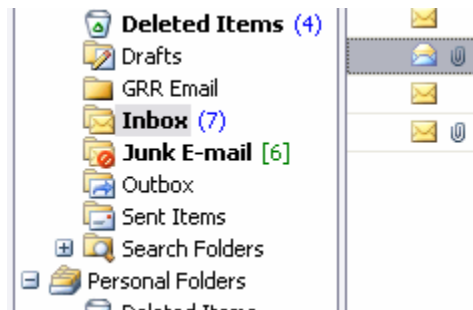


Figure 4-1

This means keeping a copy of all or most of your mail, separated into spam and good email piles, and review your spam email to ensure it is learning properly. As for the actual configuration of deletion of email, which is not recommended, but can be done, follow the following instructions.

1. Edit the following file:

```
/etc/procmailrc
```

2. Find the following section of text in the file that appears to be like what is listed below:

```
# Use spam assassin to do filtering
:0fw: spamassassin.lock
* < 256000
| /usr/bin/spamc -f -p 8061 -u $LOGNAME
```

3. Update the file to look like the configuration below. The lines in red are the important changes:

```
# Use spam assassin to do filtering
:0fw: spamassassin.lock
* < 256000
| /usr/bin/spamc -f -p 8061 -u $LOGNAME
:0
* ^X-Spam-Level: \*\*\*\*\*\*\*\*\*\*\*
/dev/null
```

(Note: We have found that in some cases the asterisks must be escaped with back slashes, in some cases they will work without being escaped)

4. Upload the `/etc/procmailrc` file back to its original location. Once the updated file has been uploaded, your server will recognize the change and begin deleting email that fit the spam criteria. If your mail services don't change after the update, please check the following:

Permissions on the `/etc/procmailrc` file are set to 640 (The permissions can be check with the `ls` command, and changed with the `chmod` command via ftp).

Make sure the new rule to delete the spam follows the rule to call `spamc` in `/etc/procmailrc`.

If the change still doesn't seem to be taking affect, contact CyberSync's Customer Service Department.

Whitelisting – Allowing Email through filtering

Customers with the Enterprise version of the Anti-Spam services may modify their `local.cf` file, with whitelist definitions.

Definitions such as:

```
whitelist_from_rcvd CyberSync.com CyberSync.com
```

Will allow all incoming mail, with received headers from `CyberSync.com` to be allowed through the anti-spam service.

Note: The first parameter is the address or domain you wish to whitelist, and the second parameter is the domain name which must be present in the Received: headers. So in case the first is not unique enough, the second pattern can be used to be more specific.

The following definitions will allow incoming email from specific domains, to be allowed to be received with a higher score by the anti-spam service:

```
whitelist_from @CyberSync.com
whitelist_from @ CyberSync.com
whitelist_from @support.CyberSync.com
whitelist_from @ CyberSync.net
```

It is recommended that these entries be added to your system; if they are not already to allow your account to receive mail from CyberSync and our support system for processing ticket requests.