By Falko Timme
Published: 2007-11-21 11:24

Adding And Updating SpamAssassin Rulesets With RulesDuJour

Version 1.0
Author: Falko Timme <ft [at] falkotimme [dot] com>
Last edited 11/14/2007

This article explains how you can download additional SpamAssassin rulesets resp. automatically update these rulesets with a shell scripts called **RulesDuJour**. These additional rulesets can increase your spam recognition rate dramatically. Most of the rulesets that RulesDuJour supports can be found on the **SA Rules Emporium web site**.

I do not issue any guarantee that this will work for you!

1 Preliminary Note

I assume that you have already set up SpamAssassin (it doesn't matter if it's a standalone daemon or called through some other daemon such as amavisd - RulesDuJour supports all these configurations).

2 Downloading RulesDuJour

I'd like to store the RulesDuJour script in the */usr/local/sbin* directory, so I go there and download the script:

```
cd /usr/local/sbin

wget http://sandgnat.com/rdj/rules_du_jour

chmod 750 rules_du_jour
```

3 Configuring RulesDuJour

Whenever the RulesDuJour script is called, it tries to read the configuration file */etc/rulesdujour/config*. Therefore we create that file now:

```
mkdir /etc/rulesdujour

vi /etc/rulesdujour/config
```

```
TRUSTED_RULESETS="TRIPWIRE SARE_EVILNUMBERS0 SARE_RANDOM"; # TRIPWIRE, SARE_EVILNUMBERS0, SARE_EVILNUMBERS1, SARE_EVILNUMBERS2, BLACKLIST, BLACKLIST_URI, RANDOMVAL, BOGUSVIRUS, SARE_ADULT, SARE_FRAUD, SARE_BML, SARE_SPOOF, SARE_BAYES_POISON_NXM, SARE_OEM, SARE_RANDOM, SARE_HEADER, SARE_HEADER0, SARE_HEADER1, SARE_HEADER2, SARE_HEADER3, SARE_HEADER_ENG, SARE_HTML, SARE_HTML0, SARE_HTML1, SARE_HTML2, SARE_HTML3, SARE_HTML4, SARE_HTML_ENG, SARE_SPECIFIC, SARE_OBFU, SARE_OBFU0, SARE_OBFU1, SARE_OBFU2, SARE_OBFU3, SARE_REDIRECT, SARE_REDIRECT_POST300, SARE_SPAMCOP_TOP200, SARE_GENLSUBJ, SARE_GENLSUBJ0, SARE_GENLSUBJ1, SARE_GENLSUBJ2, SARE_GENLSUBJ3, SARE_GENLSUBJ_ENG, SARE_HIGHRISK, SARE_UNSUB, SARE_URI, SARE_URI0, SARE_URI1, SARE_URI3, SARE_URI_ENG, SARE_WHITELIST, SARE_WHITELIST_RCVD, SARE_WHITELIST_SPF, ZMI_GERMAN, SARE_STOCKS
SA_DIR="/etc/mail/spamassassin";                # Change this to your SA local config
                                # directory, probably /etc/mail/spamassassin.
                                # For amavisd chrooted, this may be:
                                # /var/amavisd/etc/mail/spamassassin
MAIL_ADDRESS="your@yourdomain.com";
SINGLE_EMAIL_ONLY="true";                       # Set this to "true" to send only one notification
                                # email per RDJ run with "interesting"
                                # activity. Set to "" to send a separate
                                # for each interesting activity.
EMAIL_RDJ_UPDATE_ONLY="";                       # Set this to "true" to send notifications only
                                # when an update for RDJ has been retrieved.  Set
                                # to "" (default) to send notifications whenever a ruleset
                                # has changed. (Has no effect unless SINGLE_EMAIL_ONLY is set)
SA_LINT="/usr/bin/spamassassin --lint";         # Command used to lint the rules
SA_RESTART="/etc/init.d/amavisd restart";       # Command used to restart spamd
                                # May be /etc/rc.d/init.d/spamassassin restart
                                # For amavisd, may be /etc/init.d/amavisd restart
```

```
                                # For minedefang, may be /etc/init.d/mimedefang restart
CURL_PROG="/usr/bin/curl";                  # Location of the curl program
CURL_OPTS="-w %{http_code} --compressed -O -R -s -S -z"; # Parameters of the curl program
CURL="${CURL_PROG} ${CURL_OPTS}";           # Curl program with parameters
WGET_PROG="/usr/bin/wget";                  # Location of the wget program
WGET_OPTS="-N"                              # Parameters of the wget program
WGET="${WGET_PROG} ${WGET_OPTS}";           # Wget program with parameters
PERL="/usr/bin/perl";                       # Location of the perl program
GREP="/bin/grep";                           # Location of the grep program

TAIL="/usr/bin/tail -n 1";                  # Location (and parameters) for 'tail -n 1'
HEAD="/usr/bin/head -n 1";                  # Location (and parameters) for 'head -n 1'
MAILCMD="/bin/mail";                        # Location of the mail program
                                # that takes and understand the -s flag
# DEBUG="true";                             # Uncomment this to force debug mode on (or use -D)
```

The *TRUSTED_RULESETS* line contains all rulesets that you want to use (make sure you test these before using them on production systems!); I've listed all available rulesets in a comment at the end of the line.

The *SA_DIR* line must contain your SpamAssassin configuration directory; usually that's */etc/mail/spamassassin*.

The *MAIL_ADDRESS* should contain an email address to which you want RulesDuJour sent notifications about the download/update process.

All other options are explained (as comments) in the above script. The *SA_RESTART* should be the command that is used to restart SpamAssassin. If you run SpamAssassin as a standalone daemon, it's probably something like */etc/init.d/spamassassin restart* or */etc/init.d/spamd restart*; if SpamAssassin is called through amavisd, you must specify the command used to restart amavisd (e.g. */etc/init.d/amavisd restart*).

It's a good idea to use full paths to all programs in the above script (e.g. */usr/bin/spamassassin* instead of *spamassassin* or */usr/bin/curl* instead of *curl*). You can find out the full path of each program with *which*, e.g.

```
which spamassassin
```

```
which curl

which wget

which perl

which grep

which tail

which head

which mail
```

```
[root@server1 sbin]# which spamassassin
  /usr/bin/spamassassin
  [root@server1 sbin]# which curl
  /usr/bin/curl
  [root@server1 sbin]# which wget
  /usr/bin/wget
  [root@server1 sbin]# which perl
  /usr/bin/perl
  [root@server1 sbin]# which grep
  /bin/grep
  [root@server1 sbin]# which tail
  /usr/bin/tail
  [root@server1 sbin]# which head
  /usr/bin/head
  [root@server1 sbin]# which mail
  /bin/mail
[root@server1 sbin]#
```

4 Running RulesDuJour

If you are in the */usr/local/sbin* directory, you can run RulesDuJour like this:

```
./rules_du_jour
```

In an other directory you can call it like this:

```
rules_du_jour
```

Of course, you can always use the full path as well:

```
/usr/local/sbin/rules_du_jour
```

5 Creating A Cron Job

Of course, you don't want to run RulesDuJour manually each time; therefore we set up a cron job like this:

```
crontab -e
```

```
0 3 * * * /usr/local/sbin/rules_du_jour 2&>1 > /dev/null
```

The above cron job would run RulesDuJour each night at 3:00h.

6 Links

- SpamAssassin: **http://spamassassin.apache.org**

- RulesDuJour: **http://sandgnat.com/rdj/rules_du_jour**
- SpamAssassin Rules Emporium: **http://www.rulesemporium.com**