Fighting spam for fun and profit

Giovanni Bechis
<gbechis@apache.org>

Open Source Summit Europe 2018, Edinburgh

SNB Information Technology & Web Solutions

About Me

- sysadmin and developer @SNB
- OpenBSD hacker for $\sim$ 10 years
- Apache SpamAssassin developer
- random patches in random open source software (amavisd-new, courier-imap, cyrus-sasl, memcached, ...)

Different types of Spam

Spam is different for everyone

- "Nigerian" scam
- plain text messages
- advertising messages
- messages created by bots
- phishing messages

Different types of Spam

Spam is different for everyone and it depends on:

- spoken languages
- personal interests
- social networks used
- web sites visited (and subscribed)

SA as a framework

SpamAssassin should be seen as a framework, not as "plug & play" software

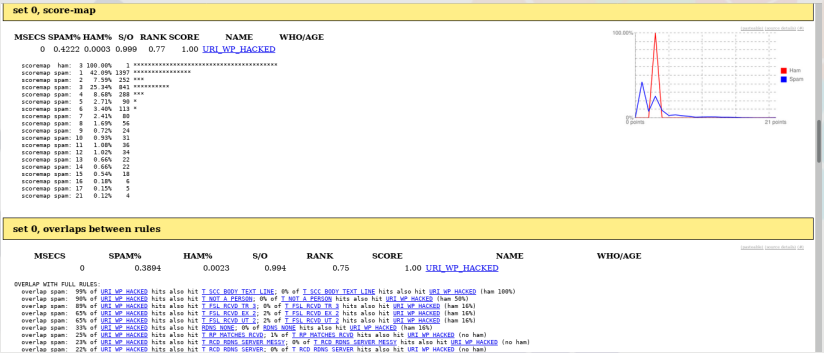
- if you follow HOWTOs you will not take the best out of any software
- to take the best out of SA:
  - write your "simple" rules
  - participate to "masscheck"
- SA is a general purpose antispam framework, it's used to filter spam in some webforms and it's even integrated in a not-so-famous cms

What's Masscheck ?

- a tool to test rules for accuracy and hit-rate
- a good way to check how rules are performing
- mass-check is run nightly based on users corpora submission, from those data, scores are assigned to rules and new rules are promoted

Checking how rules are performing

RuleQA:



set 0, score-map

MSECS	SPAM%	HAM%	S/O	RANK	SCORE	NAME	WHO/AGE
0	0.4222	0.0003	0.999	0.77	1.00	URI_WP_HACKED	

```
scoremap  ham:  3 100.00%    1 ****************************************
scoremap spam:  1  42.09% 1397 ****************
scoremap spam:  2   7.59%  252 ***
scoremap spam:  3  25.34%  841 **********
scoremap spam:  4   8.68%  288 ***
scoremap spam:  5   2.71%   90 *
scoremap spam:  6   3.40%  113 *
scoremap spam:  7   2.41%   80
scoremap spam:  8   1.69%   56
scoremap spam:  9   0.72%   24
scoremap spam: 10   0.93%   31
scoremap spam: 11   1.08%   36
scoremap spam: 12   1.02%   34
scoremap spam: 13   0.66%   22
scoremap spam: 14   0.66%   22
scoremap spam: 15   0.54%   18
scoremap spam: 16   0.18%    6
scoremap spam: 17   0.15%    5
scoremap spam: 21   0.12%    4
```


set 0, overlaps between rules

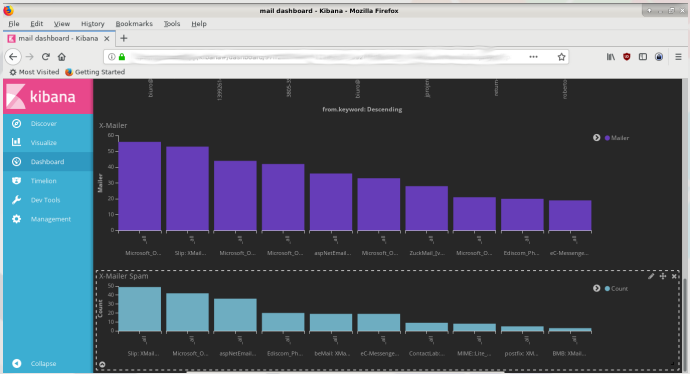
MSECS	SPAM%	HAM%	S/O	RANK	SCORE	NAME	WHO/AGE
0	0.3894	0.0023	0.994	0.75	1.00	URI_WP_HACKED	

```
OVERLAP WITH FULL RULES:
  overlap spam:  99% of URI_WP_HACKED hits also hit T_SCC_BODY_TEXT_LINE; 0% of T_SCC_BODY_TEXT_LINE hits also hit URI_WP_HACKED (ham 100%)
  overlap spam:  90% of URI_WP_HACKED hits also hit T_NOT_A_PERSON; 0% of T_NOT_A_PERSON hits also hit URI_WP_HACKED (ham 50%)
  overlap spam:  89% of URI_WP_HACKED hits also hit T_FSL_RCVD_TR_3; 0% of T_FSL_RCVD_TR_3 hits also hit URI_WP_HACKED (ham 16%)
  overlap spam:  65% of URI_WP_HACKED hits also hit T_FSL_RCVD_ES_2; 2% of T_FSL_RCVD_ES_2 hits also hit URI_WP_HACKED (ham 16%)
  overlap spam:  65% of URI_WP_HACKED hits also hit T_FSL_RCVD_UT_2; 2% of T_FSL_RCVD_UT_2 hits also hit URI_WP_HACKED (ham 16%)
  overlap spam:  38% of URI_WP_HACKED hits also hit RDNS_NONE; 0% of RDNS_NONE hits also hit URI_WP_HACKED (ham 16%)
  overlap spam:  25% of URI_WP_HACKED hits also hit T_RP_MATCHES_RCVD; 1% of T_RP_MATCHES_RCVD hits also hit URI_WP_HACKED (no ham)
  overlap spam:  23% of URI_WP_HACKED hits also hit T_RCD_RDNS_SERVER_PRESS; 0% of T_RCD_RDNS_SERVER_PRESS hits also hit URI_WP_HACKED (no ham)
  overlap spam:  22% of URI_WP_HACKED hits also hit T_RCD_RDNS_SERVER; 0% of T_RCD_RDNS_SERVER hits also hit URI_WP_HACKED (no ham)
```

- score assigned to messages that has been hit by a rule
- ham/spam hit by a single rule
- rules that overlaps on a particular rule

Checking how rules are performing

ELK:



- any info that could be detected from log files
- ham/spam hit by a single rule
- check how a single header is evaluated as spam/ham

What have SpamAssassin done in 3 years and a half ?

- sysadmin team and mass-check work
- security fixes for PDFInfo plugin and core modules
  CVE-2017-15705, CVE-2016-1238, CVE-2018-11780 & CVE-2018-11781
- perl bug triggered by SA on RedHat distros

What have SpamAssassin done in 3 years and a half ?

Assorted improvements:

- faster startup code and free(3) fixes for spamc(1)
- SSLv3 support removed from spamc(1)
- freemail antiforge improvements
- added possibility to score based on continents in geo-aware plugins
- improvements in URILocalBL plugin
- TxRep file descriptor leak fixes
- better check for http[s] mismatch plugin
- regression tests switched to Test::More

What have SpamAssassin done in 3 years and a half ?

HashBL plugin

The HashBL plugin is the interface to The Email Blocklist (EBL).

The EBL is intended to filter spam that is sent from IP addresses and domains that cannot be blocked without causing significant numbers of false positives.

What have SpamAssassin done in 3 years and a half ?

GeoIP2 support

Starting on 04/01/2018 Maxmind legacy geoip databases have been discontinued.

GeoIP2 support has been added to RelayCountry and URILocalBL plugins. In addiction RelayCountry supports also IP::Country::DB_File as an option.

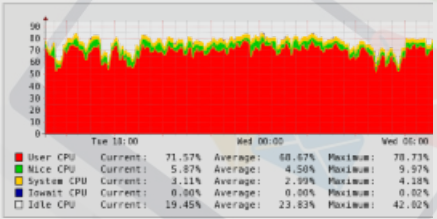
What have SpamAssassin done in 3 years and a half ?

Anti phishing plugin

A new anti phishing plugin has been developed, it searches phishing uri in a database downloaded from PhishTank or from OpenPhish.

What have SpamAssassin done in 3 years and a half ?

Resource limits plugin



A new plugin that uses BSD::Resource perl module to assure your spamd child processes do not exceed specified CPU or memory limit.

What have SpamAssassin done in 3 years and a half ?

"From Name" spoof plugin

A new plugin that perform various tests to detect spoof attempts using the From header name section.

```
From:  "safeaddress@paypal.com" <hacked@hacked.eu>
```

What have SpamAssassin done in 3 years and a half (unofficial) ?

"Ole Macro" detection plugins

Two new plugins have been developed to check if an email contains an Office attachment with a macro. They will probably be integrated in one official plugin.

What have SpamAssassin done in 3 years and a half (unofficial) ?

"Url shortener" plugin

A plugin that detects the presence of url shorteners and checks in blacklists for their targets has been developed.

KAM.cf rules: respond faster to spam

KAM.cf rules

KAM.cf is a set of "additional/unofficial" rules developed to respond faster to spam, standard rules takes some days to be deployed due to masscheck.
They are very effective but there could be "very few" false positives.

International channels

International channels

Channels are a set of signed rules, they are important and very effective because standard rules are mostly base on english emails.

SpamAssassin: the future

SpamAssassin 4.0 and future releases

- full utf-8 support
- GeoDB module for a better geolocalization support
- better TxRep handling
- check uris inside attachments
- ole macro and url shorteners plugins

How to help ?

Help needed for better spam handling

- participate in masscheck
- write to users@ ml if you encounter false positives
- open bug reports if you find bugs
- test new code

Bibliography

Some useful links

- Masscheck
  - https://wiki.apache.org/spamassassin/NightlyMassCheck
- RuleQA
  - https://wiki.apache.org/spamassassin/RuleQaApp
  - https://ruleqa.spamassassin.org
- KAM rules
  - https://www.pccc.com/downloads/SpamAssassin/contrib/KAM.cf
- Rules in SpamAssassin developers sandboxes
  - https://svn.apache.org/repos/asf/spamassassin/trunk/rulesrc/sandbox

Bibliography

Some useful links

- International channels and rules
  - http://sa.zmi.at
  - https://spamassassin.snb.it
  - http://lemat.priv.pl/pliki/sa_body_test_pl.cf
  - https://www.michelinakis.gr/Dimitris/spamassassin/gr_domain.cf
- Unofficial plugins
  - https://github.com/smfreegard/DecodeShortURLs
  - https://github.com/fmbla/spamassassin-olemacro
  - https://github.com/bigio/spamassassin-vba-macro