Fighting spam for fun and profit
the long road to SpamAssassin 4.0

Giovanni Bechis
<gbechis@apache.org>

Fosdem 2019, Brussels
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:**

- 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
What have SpamAssassin done in 3 years and a half?

- sysadmin team and mass-check work
- 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 spamic(1)
- SSLv3 support removed from spamic(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?

HashBL plugin

Checking Bitcoin scams with HashBL.

Using HashBL plugin (in 4.x release) or using an external plugin you can check if a Bitcoin address has been used for fraudulent purposes or not by asking via a dns query to $bitcoinaddress.bl.btcblack.it.
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. More antiphishing databases will be added soon.
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>
Channels are a set of signed rules, they are important and very effective because standard rules are mostly base on English emails.
New features in SpamAssassin 4.0

"Ole Macro" detection plugin

A plugin have been developed to check if an email contains an Office attachment with a macro, it tries to detect if the attached macro is malicious or not.
New features in SpamAssassin 4.0

”Authentication-Results” parser plugin

A plugin have been developed to check ”Authentication-Results” header fields, it can supply the results obtained to other plugins, to avoid repeating checks that have been performed already.
KAM.cf is a set of "additional/unofficial" rules developed to respond faster to spam. Standard rules take some days to be deployed due to masscheck. They are very effective but there could be "very few" false positives.
New features in SpamAssassin 4.0

SpamAssassin 4.0 and future releases

- full utf-8 support
- GeoDB module for a better geolocalization support
- better TxRep handling
Questions ?