

RulesProjMoreInput

Rules Project: Encouraging Contributions

(part of [RulesProjectPlan](#))

Problem description: 'The [SpamAssassin](#) committers are not spending much time writing rules. Attempts to recruit people to become committers to write rules have been somewhat unsuccessful. We could always use more committers and contributors; what can we do to encourage more contribution?'

Here's some ideas.

Sandboxes

See [RulesProjectSandboxes](#).

Mass-checking

[LorenWilton](#) noted 'A big part (perhaps the biggest part) of rules development is the mass check. Most anyone can develop a rule on their home system and see how they **think** it works. Some few (but not many) people can do a mass-check on their home system and see how it **really** works - **for them**. As proposed, this rules project doesn't address the most important part of a rules project - some way to check the rules against a fairly large corpus.'

Nightly Mass-Checks

We currently have the [NightlyMassCheck](#) systems which do this, but turnaround time is too slow for most rule developers.

It does however offer the following good aspects:

- info on how a new rule compares to the full **existing** ruleset
- overlaps between rules, using "hit-frequencies -o"
- collated results across all users' corpora, which can be broken down to view each user separately or all together
- checked rules and their results are kept in a version-control history, so benefits of VC are available
- ongoing visibility of hit-rates of the existing ruleset, against fresh corpora

The ruleqa CGI is now in the [SpamAssassin](#) zone, so this is still visible, even though the automc stuff is disabled. Here it is: <http://buildbot.spamassassin.org/ruleqa/>

'LOAFER': Suggestion: It would be good to know the % runtime figure for a sandbox rule as a missing boundary can take a rule from 1.5% to 0.0% performance hit easily:

```
perl -d:DProf mass-check -j=1 spam:dir:some_reasonable_sample_set_including_hits_and_misses
dprofpp -O 2000 > perf.log
```

[JustinMason](#): Agreed, this would be useful.

Someway of scheduling a small run during the development day would be useful, rather than waiting for the nightly. An email of users completed results would be nice to see too.

[JustinMason](#): the "preflight" system provides the "small run" option – see below!

Preflight Mass-Checks

See [RulesProjBuildBot](#).