# SpamAssassin: pravila (rules), kako ih promijeniti i kako izbjeći lažne pozitive

 Mogućnost da sasvim legitimni mailovi vaših korisnika budu prepoznati kao spam (*false-positive*) smo već spominjali u člancima u našoj E-knjizi o SpamAssassinu na adresi http://sistemac.carnet.hr/node/486 [1]. Po, u IKT industriji omraženom Murphyjevom zakonu, to će biti upravo mail vašeg pretpostavljenog, ili nekog korisnika koji jednostavno na svaki problem burno reagira i uvijek eskalira problem na više razine.

U redu, našli ste problematični mail i po članku na Portalu "Amavisd-release: oslobodite svoj mail! [2]" vratili mail iz karantene. No, kako sprječiti da se slično više ne ponovi?

Prva stvar je analizirati što je SpamAssassin našao problematično kod specifičnog maila. Ovo nije problem, jer se sve nalazi u SpamAssassin izvješću (ako je tako konfigurirano), ali i u zaglavlju maila u karanteni. Evo primjer iz stvarnog života:

```
X-Spam-Flag: YES
X-Spam-Score: 7.007
X-Spam-Level: *******
X-Spam-Status: Yes, score=7.007 tag=2 tag2=6.31 kill=6.31 tests=[FROM_HAS_ULINE_NUMS=
0.291,
    HTML_30_40=0.374, HTML_MESSAGE=0.001, RAZOR2_CF_RANGE_51_100=0.5,
    RAZOR2_CF_RANGE_E8_51_100=1.5, RAZOR2_CHECK=0.5, UNDISC_RECIPS=0.841, URIBL_BLACK
=3]
```

Sve informacije koje nam trebaju su ovdje. Različite aplikacije rabe različita zaglavlja, a ukratko značenja su:

**X-Spam-Flag** Mail je spam, jer je prešao određenu razinu (vidi se u zaglavlju)
**X-Spam-Score** Točna numerička završna razina (score) nakon provođenja svih testova
**X-Spam-Level** Završna razina zaokružena na cijeli broj, označena brojem zvjezdica
**X-Spam-Status** Najzanimljiviji redak, koji govori sve što je SpamAssassin zaključio o mailu

U ovom retku možemo vidjeti koja je točna razina spama, koja je razina "rezanja", te razina označavanja (primjerice, želimo vidjeti score maila, iako sam mail neće završiti u karanteni).

U ovom se zaglavlju nalaze svi provedeni testovi, te koju ocjenu su doprinijeli u ukupnom rezultatu. Ova na izgled kriptična imena imaju točno svoja značenja, te se u njihovoj definiciji može vidjeti zašto su se uopće uvažila. Gdje se nalaze te definicije? Većina se nalaziu direktoriju /usr/share/spamasassin, i najbrže ih je naći sa:

```
# cd /usr/share/spamassassin/
# grep -r FROM_HAS_ULINE_NUMS *
20_head_tests.cf:header FROM_HAS_ULINE_NUMS     From =~ /_\S?(?:[a-
z]+\w*?\d+|\d+\w*?[a-z]+)\w*\@/i
20_head_tests.cf:describe FROM_HAS_ULINE_NUMS   From: contains an underline and numbe
rs/letters
...
50_scores.cf:score FROM_HAS_ULINE_NUMS 0.744 0.217 0.310 0.291
```

Test se sastoji od provjere postoji li u "From" polju, osim slova, i brojevi i podvlaka (_). Zašto je ovo bitno? SpamAssassinova pravila se baziraju na statističkoj analizi milijuna spamova. Kako se u nekom određenom broju spamova pojavljuju baš ovakav oblik, to je označeno kao relevantno i ocjenjeno s koeficijentom statistički određene razine. Konkretno, radi se o imenu "Helena_sofia_dionisio".

NOVO:

Često, problematičnim se mogu pokazati i dodatna SARE pravila, koja se automatski nadograđuju preko cron skripte /etc/cron.daily/spamassasin-cn. Kako se SARE skripte više ne nadograđuju (što znači da je mogućnost preoštre ocjene veća), mnogima će prestanak uporabe SARE pravila biti jedini način da spriječe lažne pozitive. SARE, i drugi dodatni rulesetovi se nalaze u /var/lib/spamassassin/<inačica>, te ih od tamo možete obrisati. Obrišite i cron datoteku.

Kako je dosta teško iz šturog objašnjenja uvijek shvatiti o čemu se radi, najbolje je poslužiti se Googleom. Na taj način ćete, osim detaljnijeg opisa konkretnog pravila, dobiti i načine rješavanja problema s tim pravilom. Naime, nisu sva pravila jednakovrijedna, a i mogu jednostavno biti izbrisana iz distribucije SpamAsassina (npr. ovo se pravilo uopće ne pojavljuje u inačici 3.2.5, ali postoji u inačici 3.1.7).

Pravila se u svakoj novoj inačici SpamAssassina re-evaluiraju, dodaju se nova i nestaju stara. Dakle, Google je u ovim slučajevima "vaš prijatelj". Možda najpoznatije pravilo koje je dosta često znalo praviti probleme je FORGED_MUA_OUTLOOK, koje je označavalo da su zaglavlja krivotvorena tako da podsjećaju na Outlook. Zbog mnogih inačica Outlooka SpamAssassin jednostavno nije prepoznavao zaglavlja i označavao ih krivo kao lažna.

Zbog ovog legitimnog razloga vrijedi smanjiti koeficijente koje ovo pravilo dodjeljuje, ili ga čak u potpunosti anulirati. Jednostavno, u /var/lib/amavis/spamassassin/user_prefs upišite:

```
score FORGED_MUA_OUTLOOK 0
```

i uspješno ste spriječili da Vam ovo pravilo povećava ocjene mailova. Nakon svake promjene unutar SpamAssassina, pogotovo ukoliko mijenjate ili dodajete pravila, poželjno je napraviti provjeru sintaktičke korektnosti:

```
# su amavis -c 'spamassassin --lint'
```

Nikakavih poruka o greškama ne bi trebalo biti, a ukoliko se pojave treba istražiti u čemu je problem i nastajati ga popraviti. O tome, drugi puta.

Na kraju, moramo napomenuti da "petljanje" po SpamAssassinu i mijenjanje pravila napravite samo ukoliko je zaista nužno. Prije toga probajte otkloniti razloge zašto se uopće ta pravila uključuju. Provjerite DNS i reverzni DNS, podesite parametre internal_networks i trusted_networks, zamolite korisnike da ne rabe HTML mail i slično. Nakon pravilnog podešavanja sustava, lažni pozitivi će se događati u puno manjem broju.

- Logirajte [3] se za dodavanje komentara

sub, 2009-03-28 15:25 - Željko Borš**Kategorije:** Software [4]
Spam [5]
**Vote:** 0

No votes yet

**Source URL:** https://sysportal.carnet.hr./node/548

**Links**
[1] https://sysportal.carnet.hr./node/486
[2] https://sysportal.carnet.hr./node/526
[3] https://sysportal.carnet.hr./sysportallogin
[4] https://sysportal.carnet.hr./taxonomy/term/25
[5] https://sysportal.carnet.hr./taxonomy/term/34