Vylaďte si svůj SpamAssassin

Bohuslav Moučka, ÚVT MU

Se spamem, tedy s nevyžádanými zprávami [1], se pravděpodobně již setkala většina uživatelů elektronické pošty. Pro řadu z nich představuje spam vážný problém, který – bez patřičných protiopatření – ohrožuje vůbec použitelnost elektronické pošty jako nástroje efektivní komunikace. Jednou z úspěšných metod boje proti spamu je používání filtrovacích programů. Ty se snaží „oddělit zrno od plev“, tj. automaticky rozeznat řádné zprávy (kterém mají být uživateli doručeny) od nevyžádaných a obtěžujících zpráv (které lze bez dalšího zkoumání například „zahodit“).

Příkladem úspěšného nástroje pro filtraci spamu je volně dostupný systém SpamAssassin [2], využívaný hojně i na MU. Průměrné hodnoty úspěšnosti filtrace při standardním nastavení se u něj mohou pohybovat kolem 90 – 95%. Tuto úspěšnost lze dále zvýšit speciálním nastavením, které bere do úvahy specifické charakteristiky emailové komunikace u daného konkrétního uživatele. Toto „dolaďování“ filtrace SpamAssassinu však již nemusí být snadnou záležitostí a může vyžadovat poměrně hluboké znalosti z oblasti informačních technologií a systému SpamAssassin samotného.

Cílem tohoto článku je poskytnout základní informace o fungování programu SpamAssassin a nabídnout několik tipů, které zkušeným uživatelům mohou pomoci doladit filtrování jejich osobní pošty.

Jak zapnout filtr SpamAssassin

Systém SpamAssassin musí být nejprve nainstalován na vašem poštovním serveru. Chcete-li jej následně použít pro filtraci osobní pošty, je třeba do souboru `.forward` ve vašem domovském adresáři na poštovním serveru zapsat příkaz pro spuštění programu procmail, například:

```
|/packages/run/links/bin/procmail
```

a do souboru `.procmailrc` přidat řádky pro spuštění SpamAssassinu, např.:

```
:0fw
| /usr/local/spamassassin/spamassassin \
  -c /usr/local/spamassassin/rules
:0:
* ^X-Spam-Status: Yes
Mail/spam
```

Toto nastavení aktivuje spamový filtr a současně udává, že rozpoznané spamy budou ukládány do souboru (poštovní složky) `Mail/spam`. [1]

Po příchodu prvního dopisu se v domovském adresáři uživatele vytvoří podadresář `.spamassassin` a v něm mimo jiné i parametrizační soubor `user_prefs`, do něhož můžete zapisovat vlastní parametry a pravidla programu.

Jak již bylo uvedeno výše, při standardním nastavení rozezná SpamAssassin správně kolem 90% spamů. Jeho úspěšnost můžeme zvýšit třemi způsoby: úpravou některých parametrů, doplněním vlastních pravidel a učením programu.

Skóre zpráv

Program SpamAssassin obsahuje pravidla pro vyhledávání textových řetězců typických pro spamy, a řadu předdefinovaných testů. Pokud se některé pravidlo nebo test při zkoumání dané příchozí zprávy uplatní, je zprávě připočten nebo odečten stanovený počet bodů. Dosáhne-li celkové *skóre zprávy* určené hranice (přednastavené na hodnotu 5), je považována za spam. Tuto hodnotu můžete změnit nastavením příslušného parametru. Například

```
required_hits 4
```

snižuje hranici filtrace, takže více zpráv bude považováno za spamy (může tím ale vzrůst počet selhání, kdy za spam je považována i korektní zpráva).

Síťové testy

Kromě hledání vzorů ve zprávách může SpamAssassin také spolupracovat s několika servery, které shromažďují zprávy označené jako spamy některými z mnoha tisíců uživatelů po celém světě. Program zašle serveru kontrolní součet zprávy a dostane odpověď, zda jde o

[1] Pro konkrétní informace o aktuálních nastaveních a možnostech využívání filtru SpamAssassin se obraťte na fakultního správce elektronické pošty.

známý spam. SpamAssassin může takto spolupracovat se 3 servery: Vipul's Razor (`http://razor.sourceforge.net`), Pyzor (`http://pyzor.sourceforge.net`) a DCC (`http://www.rhyolite.com/anti-spam/dcc`). Každý server používá vlastní klientský program, který musí být nainstalován na poštovním serveru před spuštěním SpamAssassinu. Následující parametry pak určují, zda SpamAssassin bude tyto servery pro identifikaci spamu využívat (1 značí, že server bude využíván, 0 značí, že nebude):

```
use_pyzor 0
use_razor2 1
use_dcc 1
```

Důvěryhodné sítě

Při analýze zprávy prohlíží SpamAssassin hlavičky „Received" od poslední, zapsané poštovním serverem (relay) na němž běží SpamAssassin, směrem zpět a určuje, zda příslušná adresa je důvěryhodná. Důvěryhodná je poslední relay, celá podsíť typu B (o rozsahu 65 tisíc adres) v níž tato relay leží a privátní sítě (neveřejné adresy). Seznam důvěryhodných adres můžete rozšířit; následující příkaz doplní do seznamu důvěryhodných síti celou síť 147.229.*.* :

```
trusted_networks 147.229/16
```

Nedůvěryhodné adresy jsou hledány na serverech černých listin (black-lists). Je-li adresa nalezena, skóre zprávy bude zvýšeno.

Analýza zpráv

SpamAssassin používá testy pro kontrolu jednotlivých částí zprávy. Testuje hlavičky (header), tělo zprávy bez HTML značek (body), tělo s HTML značkami (rawbody), tělo zprávy bez dekódování MIME částí (full), URI v těle zprávy (uri) a adresy v URI (uridnsbl). Uživatel si může vytvářet i vlastní testy a zapisovat je do souboru `.spamassassin/user_prefs`.

Ukažme si příklad, jak vytvořit pravidlo, které hledá v těle zprávy text „Wysak Petroleum"[2]:

```
body  WYSAK /Wysak Petroleum/
describe WYSAK Includes Wysak Petroleum
score WYSAK 3.5 3.2 2.8 2.5
```

První řádek pravidla uvádí, kde se bude hledat, název pravidla a hledaný výraz. Druhý řádek obsahuje popis, který bude uveden v popisu zprávy, pokud se dané pravidlo uplatní. Třetí řádek obsahuje skóre přičtené zprávě při nalezení výrazu; zde může být uvedena jedna hodnota pro všechny případy nebo 4 pro následující možnosti:

- Bayesovské ani síťové testy se nepoužívají,
- Baysovské testy se nepoužívají, ale síťové ano,
- Baysovské testy se používají, ale síťové ne,
- Baysovské i síťové testy se používají.

V hodnocení zprávy se při uplatnění tohoto pravidla objeví text:

```
*    20 WYSAK BODY:
     Includes Wysak Petroleum
```

Některé základní zásady při sestavování pravidel:

- jméno pravidla má délku maximalně 22 znaků (písmen, číslic a „_"),
- jména začínající „T_" jsou vyhrazena pro testovací pravidla,
- jména začínající „__" jsou vyhrazena pro subtesty metatestů.

Metatest je pravidlo, které kombinuje výsledky několika dalších testů pomocí logických operátorů. Testy začínající „__" jsou subtesty, které nemají skóre a nejsou uvedeny v seznamu pravidel při testování zprávy.

Příklad metatestu[3]:

```
body CLICK_BELOW_CAPS
     /CLICK\s.{0,30}(?:HERE|BELOW)/s
describe CLICK_BELOW_CAPS
    Asks you to click below
body __CLICK_BELOW
    /click\s.{0,30}(?:here|below)/is
meta CLICK_BELOW
    (__CLICK_BELOW && !CLICK_BELOW_CAPS)
describe CLICK_BELOW
    Asks you to click below
```

CLICK_BELOW_CAPS je standardní pravidlo, které je pravdivé, pokud se ve zprávě vyskytují

[2] Takovýto či obdobný text se vyskytoval delší dobu v řadě spamů orientovaných na nabídku nákupu akcií.

[3] řádky v příkladu jsou v dvousloupcové sazbě rozděleny (pozn. editora)

slova CLICK HERE nebo CLICK BELOW zapsána velkými písmeny. __CLICK_BELOW je podtest bez skóre, který je pravdivý, když jsou ve zprávě výše uvedená slova v libovolné kombinaci velkých a malých písmen. Metatest CLICK_BELOW je pravdivý, pokud __CLICK_BELOW je pravdivý a CLICK_BELOW_CAPS nepravdivý, tedy pokud jsou hledaná slova v libovolné kombinaci písmen, kromě všech velkých. Vedle logických operátorů je možné v metatestech používat i operátory aritmetické a porovnávací.

Černé a bílé listiny

SpamAssassin používá černé a bílé listiny adres. Pokud je odesílatel na některé listině, je jeho zprávě zvýšeno nebo sníženo skóre. Přidat do listiny adresu odesílatele, jehož dopisy nepovažujeme nikdy za spam, můžeme příkazem

```
whitelist_from certs@cesnet.cz
```

Tato metoda ovšem není příliš bezpečná, protože adresy odesílatelů jsou ve spamech často podvržené. SpamAssassin však umožňuje spojit adresu odesílatele s důvěryhodnou relay. Chceme-li na listinu přidat všechny adresy, ve tvaru *@muni.cz, které přijdou ze stroje v doméně muni.cz, zapíšeme příkaz

```
whitelist_from_rcvd *@muni.cz muni.cz
```

Učení programu

Kromě statických testů se SpamAssassin učí ze všech zpráv, které zpracoval a svoje chování přizpůsobuje, aby maximalizoval přesnost rozeznávání spamů. Používá dvě metody učení: první jsou automatické bílé listiny, druhou jsou Bayesovské filtry.

Automatické bílé listiny

Na rozdíl od „ručních“ černých a bílých listin uvedených výše, jsou automatické bílé listiny (AWL, Auto-Whitelists) založeny na průměrování: pokud Vám někdo pošle dopis, který po vyhodnocení SpamAssassinem získá skóre 20, a následně Vám pošle druhý dopis, který získá skóre 2, pak AWL tyto dvě hodnoty zprůměruje, takže výsledné skóre druhého dopisu je zvýšeno na 11 (tzv. auto blacklisting, založený na „spamovské“ historii). Funguje to i naopak: jestliže tentýž odesilatel pošle dopis se skórem 0 a následně dopis se skórem 7, pak druhému dopisu je skóre sníženo na 3.5 (auto whitelisting, založený na nespamovské historii).

SpamAssasin využívá v systému AWL automatického učení následujícím způsobem: po každé přijaté zprávě je její skóre přičteno k celkovému skóre odesílatele a je zvýšen čítač jeho zpráv. Průměrné skóre je použito k modifikaci aktuální zprávy. Rozdíl průměrného skóre a skóre aktuální zprávy je vynásoben váhou a přičten ke skóre aktuální zprávy. Hodnota váhy je nastavitelná v rozsahu 0 až 1 příkazem

```
auto_whitelist_factor 0.7
```

Defaultní hodnota je 0.5. Nastavíme-li vyšší hodnotu, bude mít větší význam historické skóre, hodnota 1 znamená, že výsledné skóre zprávy se bude rovnat historickému skóre. Hodnota 0 způsobí, že historické skóre bude ignorováno.

Bayesovské filtry

Druhou metodou učení jsou bayesovské filtry. Zde je již zapotřebí osobní zásah konkrétního uživatele; pro využití učení je třeba programu předložit velké množství zpráv obou druhů – nechtěných zpráv (spam) i dobrých normálních dopisů (ham) a tím ho „doučovat“ pro správné rozpoznávání. Aby program pracoval efektivně, měla by každá z jeho databází (spam i ham) obsahovat alespoň tisíc zpráv (čím více, tím lépe). Minimální množství, po němž program začne využívat bayesovskou databázi, je 200 zpráv typu spam a 200 zpráv typu ham (důležité je trénovat SpamAssassin na obou typech zpráv).

Program si podle předložených výukových zpráv vytvoří databázi symbolů (řetězců délky 3–15 znaků) nalezených ve zprávách. Ke každému symbolu si zapíše počet jeho výskytů ve spamu a hamu a čas posledního použití při vyhodnocení zprávy. Symboly, které nebyly použity dlouhou dobu, jsou z databáze vymazány, aby se zvýšila efektivita. V druhé databázi je seznam zpráv, z kterých se program učil. Program učíme příkazy:

```
sa-learn --mbox --spam reklamy
sa-learn --mbox --ham  mojedopisy
```

V prvním příkaze předkládáme programu soubor reklamy obsahující spamy, ve druhém soubor mojedopisy obsahující hamy. Parametr `-mbox`, označuje, že předkládána je poštovní schránka s více zprávami (tyto schránky jsou obvykle uloženy u uživatele v podadresáři Mail nebo mail). Při kontrole je zpráva rozdělena na symboly, které jsou hledány v databázi. Podle výsledku je zprávě přiřazena pravděpodobnost, že jde o spam, což je v hlavičce vyznačeno tím, že zpráva vyhovuje pravidlu např. BAYES_80, tedy, že s pravděpodobností 0.8 - 0.9 jde o spam. Bayesovským pravidlům označujícím pravděpodobnost menší než 0.5 je přiřazeno záporné skóre, pro pravděpodobnost větší než 0.5 kladné. Po vybudování počáteční databáze symbolů je třeba ji aktualizovat. Buď'to můžeme programu předložit všechny zprávy nebo jen zprávy, které chybně označil.

Používání bayesovských filtrů můžeme ovlivnit několika parametry. Pokud bychom chtěli bayesovské filtry zcela zakázat, nastavíme parametr:

```
use_bayes 0
```

Následujícím parametrem určíme, že program se bude automaticky učit (což je implicitní nastavení, 0 znamená, že se učit nebude)

```
bayes_auto_learn 1
```

Nastavíme, že program se bude automaticky učit ze spamů s vyšším skóre než 8 a z hamů s nižším skóre než -2. (Do tohoto skóre se nepočítá skóre z bayesovských pravidel a z bílých a černých listin.)

```
bayes_auto_learn_threshold_spam 8.0
bayes_auto_learn_threshold_nonspam -2
```

Program může vyloučit ze zpracování zadané pole záhlaví zprávy, které by mohlo být při učení zavádějící (zpravidla jde o pole generované jiným antispamovým nebo antivirovým programem), například:

```
bayes_ignore_header X-Muni-Spam-TestIP
```

Možností, jak dolad'ovat přesnost filtrace SpamAssassin u své osobní pošty, je samozřejmě mnohem více, než bylo možné popsat v tomto krátkém informačním článku. Zájemce o další podrobnosti můžeme odkázat na domovské stránky systému SpamAssassin [2] a na čtivou dokumentaci na SpamAssassin-Wiki [3].

Literatura

[1] M. Kolaja, M. Bartošek. Jemný úvod do (anti)spamové problematiky. Zpravodaj ÚVT MU. ISSN 1212-0901, 2002, roč.12, č.5, s.1-6

[2] Domovská stránka projektu SpamAssassin. `http://spamassassin.apache.org/`

[3] SpamAssassin-Wiki. `http://wiki.apache.org/spamassassin/` □