HRVATSKA AKADEMSKA I ISTRAŽIVAČKA MREŽA
CROATIAN ACADEMIC AND RESEARCH NETWORK

Metode filtriranja neželjene elektroničke pošte – whitelisting, blacklisting i greylisting

CCERT-PUBDOC-2006-01-146

CARNet CERT u suradnji s LS&S

Sigurnosni problemi u računalnim programima i operativnim sustavima područje je na kojem CARNet CERT kontinuirano radi.

Rezultat toga rada ovaj je dokument, koji je nastao suradnjom CARNet CERT-a i LS&S-a, a za koji se nadamo se da će Vam koristiti u poboljšanju sigurnosti Vašeg sustava.

CARNet CERT, www.cert.hr - nacionalno središte za **sigurnost računalnih mreža** i sustava.

LS&S, www.lss.hr - laboratorij za sustave i signale pri Zavodu za elektroničke sustave i obradbu informacija Fakulteta elektrotehnike i računarstva Sveučilišta u Zagrebu.

Sadržaj

1. Uvod

Neželjena elektronička pošta (eng. *spam*) jedan je od najvećih problema računalnog svijeta. Riječ je o porukama sličnog (ili istog) sadržaja koje se šalju tisućama korisnika. Sadržaj same poruke najčešće je reklama nekog proizvoda, financijska ponuda ili neka druga vrsta usluge. Adrese primaoca najčešće se skupljaju posebno prilagođenim skriptama koje pretražuju *Usenet* grupe i *Web* stranice ili se stvaraju nasumičnim kombiniranjem najčešćih korisničkih imena i poznatih domena.

Statistički podaci govore da je danas između 55 % i 60 % svih *e-mail* poruka *spam*. Iako je stvarni postotak nemoguće odrediti, svaki korisnik elektroničke pošte složiti će se da je ova procjena ispravna, ako ne i konzervativna. Osim ometanja korisnika pri korištenju elektroničke pošte, *spam* troši ogromne količine resursa kako pružateljima internetskih usluga (eng. *Internet Service Providers*) tako i korisnicima istih, stvarajući na taj način financijske i tehnološke gubitke.

Od pojave *spama*, njegov udio u sveukupnom broju *e-mail* poruka neprestano raste. Ta činjenica dovela je do razvitka velikog broja metoda i alata namijenjenih blokiranju takvih poruka. Iako do sada nije osmišljen način potpunog uklanjanja *spam*-a, postoje specifična rješenja koja na razini lokalnih računalnih mreža pokazuju izvrsne rezultate.

Metode filtriranja neželjenih poruka koje se najčešće koriste, a ne baziraju se na analizi sadržaja poruke, jesu: *Whitelisting*, *Blacklisting* i *Greylisting* metode. One se vrlo često kombiniraju radi postizanja boljeg stupnja filtriranja. Kod primjene *Whitelisting* metode prihvaćaju se sve poruke koje se nalaze na *whitelist* popisu dok se ostale tretiraju kao *spam.* Kod *Blacklisting* metode vrijedi pak pravilo da se sve poruke pristigle s *blacklist* popisa proglašavaju *spam*-om dok se ostale propuštaju. *Greylisting* metoda spada u kategoriju specifičnih rješenja. U idealnom slučaju koji podrazumijeva distribuirano korištenje metode na velikom broju poslužitelja elektroničke pošte, ona može izuzetno poskupiti postupak slanja *spam*-a, do te mjere da slanje *spam* poruka pošiljateljima postane neisplativo.

Ovaj dokument opisuje navedene metode filtriranja neželjenih poruka s naglaskom na način rada *greylisting* metode, njene prednosti i nedostatke te mogućnosti implementacije na postojeće poslužitelje elektroničke pošte.

2. *Whitelisting* i *Blacklisting* metode

Whitelisting metoda zasniva se na prihvaćanju svih poruka elektroničke pošte pristiglih s adresa koje se nalaze na *whitelist* popisu. To je najčešće lokalni popis adresa, specifičan za pojedinu domenu, koji provjerenim korisnicima omogućava nesmetanu komunikacija bez nepotrebnih kontrola. *Whitelisting* je sastavni dio većine implementacija *greylisting* metoda. U slučaju da se IP adresa primljene poruke nalazi na *whitelist* popisu, poruka se odmah dostavlja čime se izbjegavaju kašnjenja uzrokovana *greylisting* metodom.

Blacklisting metoda koristi se popisom IP adresa s kojih su u određenom proteklom periodu pristigle poruke elektroničke pošte klasificirane kao *spam*. Iako se popis može čuvati lokalno, najčešće se ti popisi provjeravaju u stvarnom vremenu, s poslužitelja namijenjenih upravo tome. Takvi poslužitelji imaju visoku frekvenciju ažuriranja i u svakom trenutku sadrže *blacklist* popise trenutno aktivnih pošiljatelja *spam* poruka. Primjeri poslužitelja na kojima se može provjeriti da li se određena IP adresa nalazi na *blacklist* popisu ima mnogo, npr. *dnsbl.info*.

3. *Greylisting* metoda

Greylisting je zamišljen kao *antispam* metoda koja će krajnjem korisniku biti potpuno transparentna a od administratora poslužitelja elektroničke pošte zahtijevati minimalnu količinu održavanja. Može se u grubo opisati kao kombinacija *whitelist* i *blacklist* metoda. U terminima elektroničke pošte, to su komplementarne metode koje se temelje na bezuvjetnom prihvaćanju odnosno odbacivanju sve pošte pristigle s adresa koja se nalazi na listama.

3.1. Pregled metode

Prilikom pokušaja dostave poruke elektroničke pošte, *greylisting* metoda pregledava tri osnovne informacije:

- IP adresu računala koje pokušava dostaviti poruku,
- *e-mail* adresu pošiljatelja („`MAIL FROM`“ polje) i
- *e-mail* adresu primaoca („`RCPT TO`“ polje).

Kombinacija tih triju informacija čini jedan triplet. U slučaju da je određeni triplet prvi puta viđen, odbija se njegova isporuka kao i isporuka svih poruka s istim tripletom koje stignu u određenom vremenskom periodu. SMTP (*Simple Mail Transfer Protocol*) protokol specificira mogućnost privremene nemogućnosti isporuke elektroničke pošte, tako da valjani poslužitelj elektroničke pošte - MTA (*Mail Transfer Agent*) nakon određenog vremenskog intervala pokušava ponoviti isporuku. Ova je činjenica bitna, jer je većina *spam* poruka poslana koristeći aplikacije koje su razvijene samo u tu svrhu. One ne implementiraju u potpunosti SMTP protokol tj. ne pokušavaju ponoviti isporuku. Najčešće koriste privremene, dinamičke IP adrese, što automatski onemogućava ponovni pokušaj slanja poruke.

Važan aspekt ove metode, koji je razlikuje od većine drugih, je činjenica da ne može doći do lažne klasifikacije valjane poruke kao *spam*-a (sve dok MTA potpuno implementira specifikaciju SMTP protokola). Metoda je posebno efikasna po pitanju potrošnje resursa u vidu procesorskog vremena, odnosno mrežnog prometa. Za razliku od heurističkih metoda raspoznavanja *spam* poruka koje se baziraju na analizi sadržaja poruke, kod *greylisting* metode uopće se ne pregledava sadržaj poruke. Štoviše, sadržaj poruke se u slučaju odbacivanja iste niti ne prima, što uvelike pridonosi smanjivanju mrežnog prometa.

3.2. Implementacija *greylisting* metode

Implementacija *greylisting* metode zahtjeva zapisivanje informacija o pojedinim tripletima. Za to se može koristiti neki od postojećih sustava za upravljanje bazama podataka, no zbog jednostavnosti operacija moguća su i jednostavnija rješenja. Struktura baze podataka vrlo je jednostavna, potrebni podaci su sljedeći:

- vrijeme prvog dolaska određenog tripleta,
- vrijeme isteka perioda blokiranja tripleta,
- vrijeme isteka zapisa o tripletu (za stare zapise),
- broj blokiranih pokušaja dostave poruke s određenim tripletom i
- broj poruka sa određenim tripletom koje su uspješno dostavljene.

Implementaciju je moguće upotpuniti s dodatnim informacijama, no navedene su dostatne za ispravni rad *greylisting*-a.

U nastavku slijedi primjer SMTP sjednice koji prikazuje djelovanje *greylisting*-a, „K:“ je prefiks podataka poslanih od klijenta a „P“ od poslužitelja:

```
K: HELO domenaA.com
P: 250 Hello domenaA.com
K: MAIL FROM: <posiljatelj@domenaA.com>
P: 250 2.1.0 Sender OK
K: RCPT TO: <primatelj@domenaB.com>
P: 451 4.7.1 Please try again later
```

Prikaz sjednice koja obustavlja dostavu poruke pokazuje da se prekidom odmah po primitku `RCPT` naredbe uvelike smanjuje mrežni promet. Sadržaj poruke šalje se nakon `DATA` naredbe, što je prikazano u sljedećem primjeru sjednice. Ovaj puta radi se o sjednici čiji je ishod uspješna dostava poruke:

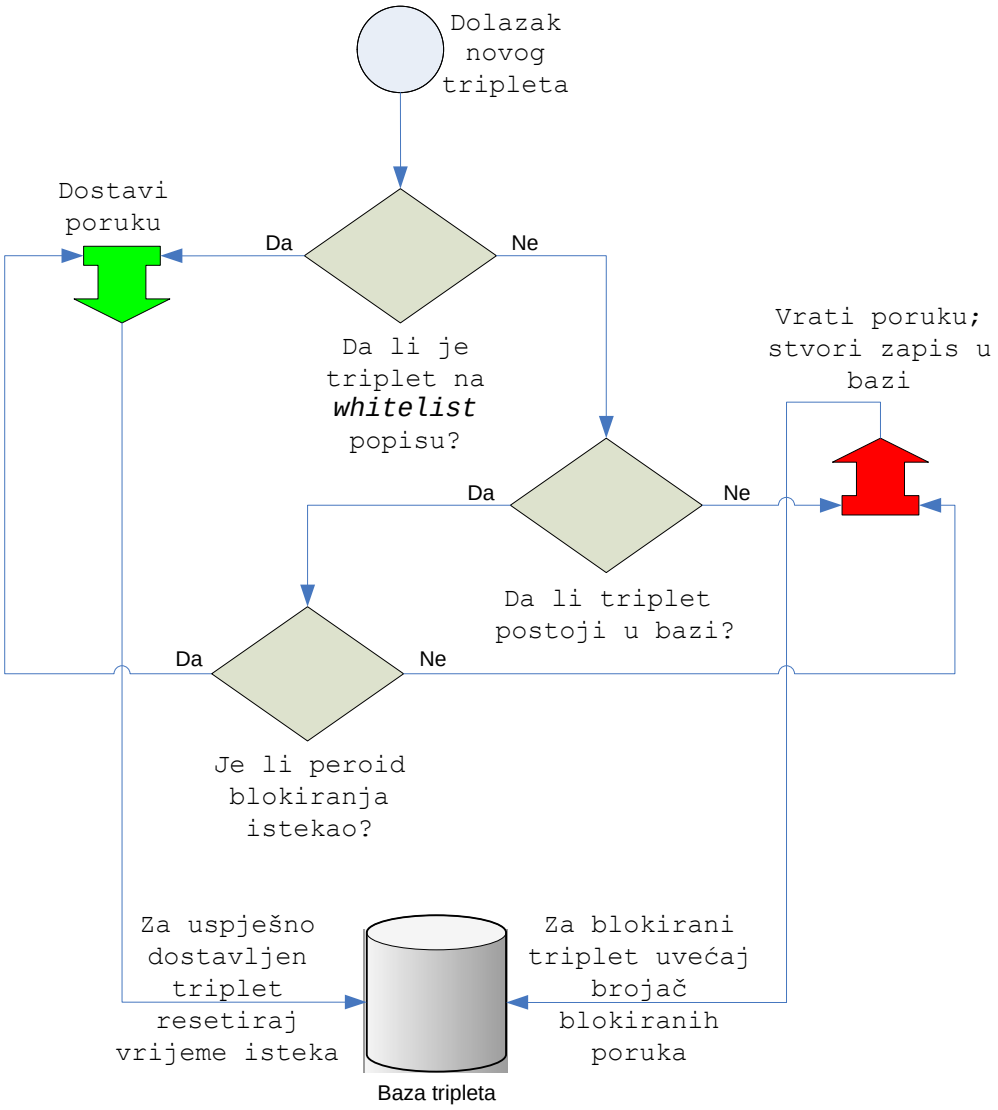
```
K: HELO domenaA.com
P: 250 Hello domenaA.com
K: MAIL FROM: <posiljatelj@domenaA.com>
P: 250 2.1.0 Sender OK
K: RCPT TO: <primatelj@domenaB.com>
P: 250 2.1.0 Recipient OK
K: DATA
P: 354 Enter mail
... (slanje podataka)
P: 250 2.0.0 Message accepted for delivery
```

Sustav je poželjno nadopuniti *whitelisting* metodom kojom se zaobilazi nepotrebno kašnjenje kod slanja poruka unutar iste domene ili komunikacije s domenskim MX (eng. *Mail exchange*) poslužiteljima.

Uzevši u obzir navedene činjenice, bazična implementacija izgledala bi na sljedeći način:

- Provjeri je li adresa računala s kojeg dolazi poruka na *whitelist* popisu, ukoliko je - dostavi poruku.
- Provjeri da li baza podataka sadrži informaciju o tripletu:
  - Ako triplet ne postoji u bazi, stvara se zapis i poslužitelju elektroničke pošte sa čije je adrese stigla poruka šalje se dojava o privremenoj nemogućnosti isporuke poruke.
  - Ako triplet postoji u bazi ali period blokiranja istog još uvijek nije istekao, poslužitelju elektroničke pošte sa čije je adrese stigla poruka šalje se dojava o privremenoj nemogućnosti isporuke poruke.
  - Ako triplet postoji u bazi a njegov je period blokiranja istekao, dostavi poruku.
- Ako je dostava poruke uspješna, u zapisu o tom tripletu uvećaj brojač uspješno pristiglih poruka, a vrijeme isteka zapisa postavi opet na nulu.
- Ako je dostava poruke privremeno onemogućena, u zapisu o tom tripletu uvećaj brojač blokiranih pokušaja dostave poruka. U slučaju da je pošiljatelj poruke tzv. *null* pošiljatelj (posebni slučaj), poruka o grešci ne javlja se odmah nakon `RCPT` naredbe, već se čeka `DATA` naredba.

U nastavku slijedi blok dijagram *greylisting* metode.



Slika 1: Blok dijagram *greylisting* metode

3.3. Preporučeni parametri *greylisting* metode

Greylisting metodu moguće je implementirati koristeći različite vremenske parametre sustava. Autor metode [3] ipak sugerira sljedeće vrijednosti:

- period blokiranja novog tripleta – 1 sat
- životni vijek tripleta koji je blokiran – 4 sata
- životni vijek propuštenog tripleta – 36 dana

Period blokiranja novog tripleta postavljen je na 1 sat iz više razloga. Vrijednost ne smije biti prevelika de ne dolazi do prevelikog kašnjenja ili čak isteka vremena nakon kojega poslužitelj elektroničke pošte odustaje od pokušaja ponovnog slanja poruke. S druge strane, pošiljatelj *spam* poruka najčešće koriste privremenu IP adresu koja brzo biva otkrivena i objavljena na *blacklist* poslužiteljima. Istraživanja pokazuju da većina (preko 90 %) pošiljatelja *spam* poruka ne koristi istu IP adresu duže od nekoliko minuta.

Životni vijek tripleta koji je blokiran postavlja se na duže vrijeme (4 sata) zbog činjenice da većina poslužitelja elektroničke pošte ima definiran kraći interval nakon kojega se ponovno šalje poruka koja nije uspješno dostavljena. S druge strane, ta vrijednost mora biti što manja tako da se u bazi ne čuvaju nepotrebni tripleti, koji bi u slučaju velikog broja poruka mogli zauzeti veliku količinu resursa.

Životni vijek propuštenog tripleta postavlja se na puno veći vremenski interval (36 dana) zbog činjenice da taj period pokriva slučajeve kada je korisnik pretplaćen na određenu mjesečnu *mailing*

listu. Tim je intervalom pokriven i slučaj kada se poruke s liste šalju na određeni dan u mjesecu pa intervali mogu biti veći od 30 dana (do maksimalno 35 dana).

4. Implementacije

U nastavku dokumenta prikazani su osnovni podaci o postojećim implementacijama *greylisting*-a za nekoliko poslužitelja elektroničke pošte. Za svaki od tih poslužitelja odabrana je i opisana jedna implementacija, iako je njihov broj veći.

4.1. Implementacija za *Sendmail* poslužitelj

Sendmail omogućava obogaćivanje funkcionalnosti poslužitelja elektroničke pošte dodavanjem tzv. *milter* komponenti, tj. filtra za elektroničku poštu. Implementacija *greylisting*-a za *Sendmail* naziva se *RelayDelay* i implementirana je kao *milter* komponenta. Korištenje komponente zahtjeva *Sendmail* inačice 8.12.x (ili novije) s instaliranim *milter* programskim bibliotekama, Perl prevodilac inačice 5.8.0 (ili novije) s DBI modulima i modulom DBD:mysql te MySQL sustav za upravljanje bazama podataka inačica 3.23.xx (ili noviji). U nastavku su prikazane SQL naredbe za stvaranje osnovne baze odnosno tablica koje se koriste u *greylisting* metodi (implementacija sadrži i dodatne tablice namijenjene stvaranju izvještaja).

```
CREATE DATABASE relaydelay;
USE relaydelay;
create table relaytofrom    (
        id              bigint           NOT NULL auto_increment,
        relay_ip        char(16),
        mail_from       varchar(255),
        rcpt_to         varchar(255),
        block_expires   datetime         NOT NULL,
        record_expires  datetime         NOT NULL,
        blocked_count   bigint default 0 NOT NULL,
        passed_count    bigint default 0 NOT NULL
        aborted_count   bigint default 0 NOT NULL
        origin_type     enum('MANUAL','AUTO') NOT NULL
        create_time     datetime         NOT NULL
        last_update     timestamp        NOT NULL,
        primary key(id),
        key(relay_ip),
        key(mail_from(20)),
        key(rcpt_to(20))
);

create table dns_name
(
       relay_ip      varchar(18)       NOT NULL,
       relay_name    varchar(255)      NOT NULL,
       last_update   timestamp         NOT NULL,
       primary key(relay_ip),
       key(relay_name(20))
);
```

4.2. Implementacija za *Postfix* poslužitelj

Postfix poslužitelj elektroničke pošte svoju funkcionalnost obogaćuje tzv. poslužiteljima sigurnosnih pravila (eng. *policy server*). Implementacija *greylisting* metode za *Postfix* poslužitelj naziva se *Postgrey* a realizirana je upravo kao poslužitelj sigurnosnog pravila te je uključena u sve distribucije Postfixa (u direktoriju `/examples/smtpd-policy/greylist.pl`). Da bi *Postgrey* ispravno radio potreban je Perl prevodilac inačice 5.6.0 (ili novije), Perl moduli Net::Server, IO::Multiplex, BerkeleyDB te programska biblioteka Berkeley DB inačice 4.1 (ili novije). U datoteci `greylist.pl` moguće je odrediti lokaciju datoteke koja služi kao baza podataka *greylisting* metode te period blokiranja novih tripleta. Početne postavke su sljedeće:

```
$database_name="/var/mta/greylist.db";
$greylist_delay=60;
```

4.3. Implementacija za *Exim* poslužitelj

Implementacija za *Exim* poslužitelj elektroničke pošte naziva se *Greylistexim* a realizirana je C programskim jezikom. Za pohranu podataka koristi se MySQL sustav za upravljanje bazama podataka, pri čemu je struktura korištenih tablica jednaka onoj prikazanoj za *RelayDelay* implementaciju.
Početni vremenski parametri, koje je moguće promijeniti putem konfiguracijskih datoteka, postavljeni su na vrijednosti 55 minuta (period blokiranja novog tripleta), 8 sati (životni vijek tripleta koji je blokiran) i 36 dana (životni vijek propuštenog tripleta).

5. Korisnici *greylisting* metode

Greylisting je danas implementiran na velikom broju poslužitelja elektroničke pošte. Između ostalih, korisnici *greylisting* metode jesu:

- SoftHome.net
- Sneakemail
- University of Bergen
- Texas A&M University
- RWTH
- PowWeb
- MailSnare.net
- Tiger Technologies
- Leibniz Rechen Zentrum
- The Roller Network
- Svenska Universitet
- APNIC (Asia Pacific Network Information Centre)

Postoje, međutim i poslužitelji elektroničke pošte koji zbog nepotpune implementacije SMTP protokola u trenutku pisanje ovog dokumenta ne podržavaju *greylisting* metodu. To su:

- Novell Groupwise, inačice 6.0,
- ISMail, inačice 1.7.1,
- InterMail, inačice 4.0,
- Kerio MailServer, inačice 5.0.5.

6. Zaključak

Rješenje koje bi u potpunosti uklonilo problem primanja neželjenih mail poruka još uvijek nema. *Greylisting* je metoda koja ne može zaustaviti slanje *spam* poruka, ali sigurno može pošiljateljima *spam*-a taj proces učiniti vremenski zahtjevnijim a samim time i financijski neisplativim. Metoda ima i svoje nedostatke, u prvom redu to je unošenje vremenskog kašnjenja kod „legalnih" poruka. Iako se kašnjenje događa samo prilikom prve interakcije dviju stranka koje koriste servis elektroničke pošte, nekim je korisnicima to neprihvatljivo pa stoga niti ne koriste ovu metodu.

7. Reference

[1] Greylisting.org,
http://www.greylisting.org/
[2] RFC specifikacija SMTP protokola,
http://www.faqs.org/rfcs/rfc2821.html
[3] The Next Step in the Spam Control War: Greylisting, Evan Harris,
http://projects.puremagic.com/greylisting/whitepaper.html