Sistemi di filtraggio SPAM

Seminario Sicurezza

Daniele Venzano

mailto:venza@libero.it

23 dicembre 2003

Sistemi di identificazione dello SPAM

I messaggi SPAM hanno dei costi che si riflettono su chi riceve i messaggi e non su che li manda.
Questi costi sono sia in termini di banda (che in certi casi si paga, es. GPRS), sia in termini di tempo perso per cancellare i messaggi.
L'obiettivo dei sistemi di identificazione è quello di riconoscere i messaggi SPAM in maniera automatica.

Sistemi di identificazione dello SPAM

I messaggi SPAM hanno dei costi che si riflettono su chi riceve i messaggi e non su che li manda.
Questi costi sono sia in termini di banda (che in certi casi si paga, es. GPRS), sia in termini di tempo perso per cancellare i messaggi.
L'obiettivo dei sistemi di identificazione è quello di riconoscere i messaggi SPAM in maniera automatica.

- Filtri basati su hashing dei messaggi

Sistemi di identificazione dello SPAM

I messaggi SPAM hanno dei costi che si riflettono su chi riceve i messaggi e non su che li manda.
Questi costi sono sia in termini di banda (che in certi casi si paga, es. GPRS), sia in termini di tempo perso per cancellare i messaggi.
L'obiettivo dei sistemi di identificazione è quello di riconoscere i messaggi SPAM in maniera automatica.

- Filtri basati su hashing dei messaggi
- Filtri Bayesiani (o Statistici)

Sistemi di identificazione dello SPAM

I messaggi SPAM hanno dei costi che si riflettono su chi riceve i messaggi e non su che li manda.
Questi costi sono sia in termini di banda (che in certi casi si paga, es. GPRS), sia in termini di tempo perso per cancellare i messaggi.
L'obiettivo dei sistemi di identificazione è quello di riconoscere i messaggi SPAM in maniera automatica.

- Filtri basati su hashing dei messaggi
- Filtri Bayesiani (o Statistici)
- Filtri basati su punteggi (o Euristici)

Sistemi di identificazione dello SPAM

I messaggi SPAM hanno dei costi che si riflettono su chi riceve i messaggi e non su che li manda.
Questi costi sono sia in termini di banda (che in certi casi si paga, es. GPRS), sia in termini di tempo perso per cancellare i messaggi.
L'obiettivo dei sistemi di identificazione è quello di riconoscere i messaggi SPAM in maniera automatica.

- Filtri basati su hashing dei messaggi
- Filtri Bayesiani (o Statistici)
- Filtri basati su punteggi (o Euristici)
- Filtri Challenge-Response

Sistemi di identificazione dello SPAM

I messaggi SPAM hanno dei costi che si riflettono su chi riceve i messaggi e non su che li manda.
Questi costi sono sia in termini di banda (che in certi casi si paga, es. GPRS), sia in termini di tempo perso per cancellare i messaggi.
L'obiettivo dei sistemi di identificazione è quello di riconoscere i messaggi SPAM in maniera automatica.

- Filtri basati su hashing dei messaggi
- Filtri Bayesiani (o Statistici)
- Filtri basati su punteggi (o Euristici)
- Filtri Challenge-Response
- Filtri basati su blacklist

Spamassassin 1/3

Spamassassin è un sistema di filtraggio basato su punteggi che può utilizzare anche delle blacklist, dei filtri basati su hashing (razor, pyzor) e un filtro bayesiano.

Messaggio di SPAM innocuo, riconosciuto da Spamassassin:

```
Subject: Hello, it's again me.
To: "venza@libero.it" <venza@libero.it>
Content-Type: text/html;

Good day!
I's again me, i wrote you few days ago, do you remember me?
I live in a small city and i do not meet very many people, i found
great site, look at it
http://218.15.192.225/index.php?a=3D000344&b=3D001&c=3D20
also "http://218.15.192.225/ugly.php?campaign=3D000344
there you can ask me not send more emails to you.

Anna ssltame
```

Spamassassin 2/3

Spamassassin segna in modo particolare i messaggi riconosciuti come SPAM aggiungendo degli header:

```
X-Spam-Flag: YES
X-Spam-Checker-Version: SpamAssassin 2.60 (1.212-2003-09-23-exp)
X-Spam-Status: Yes, hits=5.5 required=4.0 tests=BAYES_90,HTML_MESSAGE,
        NORMAL_HTTP_TO_IP,SUB_HELLO,TO_ADDRESS_EQ_REAL autolearn=no
        version=2.60
X-Spam-Level: *****
```

Questo messaggio ha totalizzato 5.5 punti, ne bastavano 4 per essere riconosciuto. Dato che il punteggio è vicino alla soglia il messaggio non è stato utilizzato per il *training* del filtro bayesiano (autolearn=no)

Spamassassin 3/3

5 tra i molti test che effettua Spamassassin hanno dato risultato positivo, ma solo 2 di questi hanno provocato il superamento della soglia.

```
 pts rule name              description
---- ---------------------- --------------------------------------
 2.5 SUB_HELLO              Subject starts with "Hello"
 0.8 TO_ADDRESS_EQ_REAL     To: repeats address as real name
 2.1 BAYES_90               BODY: Bayesian spam probability is 90 to 99%
                            [score: 0.9274]
 0.1 HTML_MESSAGE           BODY: HTML included in message
 0.1 NORMAL_HTTP_TO_IP      URI: Uses a dotted-decimal IP address in URL
```

Il filtro bayesiano ha dato una probabilità di SPAM del 92% al messaggio, e il messaggio che inizia con 'Hello' ha contribuito per l'altra metà.
Però Razor e le blacklist hanno dato risultato negativo, malgrado il messaggio provenisse da una classe di indirizzi *dial up*.

Filtraggio lato client o lato server

Lato server:

- \+ Più messaggi, si filtra meglio
- \+ Risparmio di banda
- – Più facile scontentare gli utenti
- – Grande carico di elaborazione
- – Software più complesso da configurare

Filtraggio lato client o lato server

Lato server:

- \+ Più messaggi, si filtra meglio
- \+ Risparmio di banda
- − Più facile scontentare gli utenti
- − Grande carico di elaborazione
- − Software più complesso da configurare

Lato client:

- \+ L'utente può calibrare meglio i filtri
- \+ Sistema distribuito, più robusto
- \+ Filtri integrati nei client di posta (es. Mozilla)
- − Nessun risparmio di banda

Riferimenti

- http://www.paulgraham.com/spam.html
- http://www.spamassassin.org