SPAM al CILEA? No, grazie (se possibile)

Francesca Bonini
CILEA, Segrate

Abstract
Il crescente volume di posta elettronica indesiderata ha portato il CILEA a prendere delle contromisure a riguardo, installando un software per filtrare i messaggi in arrivo ed evitare di intasare le caselle di posta inutile, non richiesta, indesiderata e, molto spesso, dannosa. In breve lo *spam*.

Keywords: mail, *spam*, posta elettronica.

Visto l'altissimo tasso di crescita del traffico di posta non desiderata (comunemente: *spam*[1]) che invade ormai tutte le caselle di posta elettronica, il CILEA ha provvisto il sistema di posta elettronica (mail.cilea.it[2]) di un apposito software anti-spam: *PreciseMail AntiSpam gateway* (PMAS di Process Software [3]).

PMAS è un software che analizza la posta elettronica e può individuare messaggi di *spam* applicando sia regole euristiche, sia algoritmi propri dei metodi d'intelligenza artificiale. Un approccio evidentemente complesso ma indispensabile per una classe di problemi altrettanto complessa: cercare segnali che diano indicazione sul contenuto di un messaggio, ma più in termini di semantica ("è una pubblicità o la battuta di un amico?") che non di dimensioni (grande, medio o piccolo?) di provenienza (Dall'interno? Da amici? Da sconosciuti?) o pericolosità ("Ha un virus?","È pulito?"). E discriminare i messaggi di *spam* da quelli legittimi senza l'intervento "umano" è una cosa realmente molto difficile da fare.
Il sistema, come la maggior parte di quelli oggi disponibili sul mercato, si basa su un insieme di regole, costantemente aggiornate dal fornitore, ma modificabili ed integrabili dall'amministratore di sistema. Esse vengono applicate sempre sia al *subject* del messaggio sia al contenuto. In parallelo a questa analisi viene utilizzato anche un motore Bayesiano[4] di intelligenza artificiale, in grado di "imparare" e perfino di essere addestrato a discriminare tra posta legittima e *spam*. La combinazione di questi due metodi rende il prodotto particolarmente efficiente con un tasso di riconoscibilità (diciamo `buono/cattivo`) molto vicino al 100% dei messaggi esaminati.

L'analisi di PMAS su di un generico messaggio si concretizza con l'attribuzione di un punteggio numerico, che cerca di misurare il livello di spam (con un bruttissimo termine italiano spammicità o inglese *spamicity*). Sulla base di tale valore, l'azione di PMAS può quindi essere una combinazione delle seguenti:

- scartare il messaggio;
- porre il messaggio in una quarantena: un "parcheggio" temporaneo che consente all'utente comunque, se interessato, di recuperarlo;
- variare gli *header* del messaggio aggiungendo un campo (`X-PMAS:`) che riporta il suo punteggio come *spam* e consegnarlo;
- modificare il *Subject* del messaggio e consegnarlo.

Ognuna di queste azioni viene intrapresa in base al punteggio di *spam* attribuito da PMAS al messaggio ed a tre valori di soglia (A, B e C qui

[1] Per saperne di più:
http://www.collinelli.net/antispam
http://www.spamlaws.com/
http://spam.abuse.net

[2] o anche, icil64.cilea.it

[3] http://www.process.com

[4] http://www.process.com/precisemail/bayesian_filtering.htm

di seguito) definiti dalla sua configurazione corrente.
Ad esempio:

- al di sotto di un valore A il messaggio è considerato legittimo e consegnato senza alcuna ulteriore azione,
- per valori compresi tra A e B il messaggio è considerato sospetto e viene cambiato il *subject* prima di consegnarlo,
- per valori compresi tra B e C il messaggio è quasi sicuramente *spam*, viene quindi posto in quarantena e ne viene mandata notifica al ricevente originale,
- al di sopra di un valore C il messaggio è sicuramente *spam* e viene scartato.

Nella configurazione del CILEA il valore C è sostanzialmente infinito: in questo modo nessun messaggio viene scartato "a priori" ma si lascia al ricevente la facoltà di recuperarlo o meno dalla quarantena. Il valore di A è normalmente 3, quello di B attualmente è 8.

Quarantena dei messaggi

Quando un messaggio è identificato come *spam*, viene posto in quarantena finché il ricevente non specifica le azioni da prendere a riguardo. Ad intervalli di tempo regolari (solitamente una o due volte al giorno), viene automaticamente notificato agli utenti, tramite mail, un unico mail di sommario per tutti i mail ricevuti e posti in quarantena per l'utente stesso. A questo punto, il ricevente può scegliere di rilasciare i mail che considera legittimi e desidera ricevere. In ogni caso, la durata della permanenza in quarantena è a termine e regolata da un parametro di configurazione, scaduto il quale il messaggio viene eliminato automaticamente. L'attuale tempo di permanenza in quarantena è di 14 giorni.

Per ricevere il messaggio l'utente deve:

- aprire il messaggio di notifica, analogo a questo:

```
From:    PreciseMail@mail.cilea.it
To:      pippo@mail.cilea.it"
CC:
Subj:    PreciseMail Quarantined Messages
PreciseMail Anti-Spam ha messo in quarantena i seguenti messaggi diretti a te.
Per recuperare uno o piu' di questi messaggi, fai un REPLY a questo mail
cancellando tutto i testo TRANNE la riga che inizia con "Message:" relativa
ai messaggi quarantinati che intendi recuperare.

Per un aiuto via mail sull'uso di PreciseMail Anti-Spam, manda un messaggio a
"PreciseMail" che contenga nel testo solo la parola HELP.

Dal momento di questa notifica, tutti i messaggi qui citati sono recuperabili
per 14 giorni, cioe' fino a 25-Nov-2003. Dopo tale data verranno
cancellati automaticamente, senza necessita' di un tuo intervento.
================================================================================
Message: SPAM$20031111194830923AEA7BD3.SPAM
Date:    Tue, 11 Nov 2003 19:48:24 +0100
From:    fuffo@somewhere.com
To:      pippo@mail.cilea.it
Subject: Fwd: Visa application #: 4587094

================================================================================
PreciseMail Anti-Spam has quarantined those incoming messages for you.
To retrieve a message from the quarantine area, REPLY to this message,
deleting all of the text except the "Message:" lines you want
retrieved.  For help on using PreciseMail Anti-Spam, send the command HELP
in the body of a mail message to "PreciseMail".

At the time of this notification, these messages will be retrievable
for 14 days (until 25-Nov-2003).
```

- fare un "*reply*" a tale messaggio
- modificare il testo del messaggio originale lasciando solo i riferimenti relativi ai messaggi che si vogliono ricevere, nel caso dell'esempio lasciare solo la riga:

```
Message:
SPAM$20031111194830923AEA7BD3.SPAM
```

- spedire il messaggio.

A questo punto l'utente riceverà prima una notifica riguardo la richiesta fatta e quindi il messaggio o messaggi richiesti, cui è stato aggiunto lo *header* "X-PMAS-QUARANTINE: ..." a ricordo dell'azione di rilascio da parte dell'utente stesso. Se il messaggio originale è già stato rimosso perché in quarantena da troppo tempo, il tentativo d rilascio riporterà un errore e la descrizione di tale possibilità ("...a "file not found" error may mean the message was deleted as a result of routine maintenance.").

Configurazione personale di *Whitelist* e *Blacklist*

- Una ***whitelist*** è un elenco di tutti gli indirizzi che devono venire accettati, indipendentemente dal contenuto del messaggio, che quindi non verranno analizzati da PMAS ma semplicemente inoltrati al ricevente;
- una ***blacklist*** è un elenco di tutti gli indirizzi di posta elettronica che si considerano *spammer*, i cui messaggi verranno automaticamente cancellati dal sistema.

Un messaggio il cui campo *From*: appartiene alla *whitelist* verrà automaticamente accettato, quello di un *From*: appartenente alla *blacklist*, verrà automaticamente scartato.
Esistono due livelli di *white/blacklist*: di sistema, che vengono configurate dai gestori della macchina e hanno validità generale e personali, la cui configurazione è lasciata al singolo titolare di account.
La definizione di tali liste personali può essere fatta tramite mail, inviati al processore di PMAS (quindi nel caso di utenti CILEA all'indirizzo PreciseMail@mail.cilea.it) seguendo alcune regole, qui indicate.

I comandi disponibili sono:

HELP	per ottenere l'help a riguardo
REVIEW	per avere il contenuto delle propria *whitelist* e *blacklist*,
MESSAGE:	per rilasciare un messaggio in quarantena,
WHITELIST:	per aggiungere una regola alla *whitelist*,
BLACKLIST:	per aggiungere una regola alla *blacklis,t*
UNWHITELIST:	per rimuovere una regola dalla *whitelist*,
UNBLACKLIST:	per rimuovere una regola dalla *blacklis,t*

N.B. ove indicati, i ":" fanno parte dei comandi.

Esempi:

- aggiungo indirizzi alla *whitelist* e alla *blacklist*, inviando un messaggio con testo:

```
WHITELIST: user@domain
BLACKLIST: spammer@example.com
```

- è possibile specificare degli indirizzi più generici utilizzando il metacarattere "*":

```
WHITELIST:    *@domain
BLACKLIST:    *offer*@*
```

Per aggiungere (o togliere) regole alla *whitelist* o alla *blacklist* si deve inviare un messaggio contenente le regole desiderate, attendere un messaggio di conferma da PMAS e quindi rispondere come conferma per abilitare effettivamente le nuove regole, a questo punto PMAS verifica la validità delle regola da inserire e, se tutto va bene, le inserisce.

Nota: le regole della *whitelist* vengono applicate prima di quelle della *blacklist*, quindi hanno la precedenza su queste ultime. Dunque, se una regola è specificata sia come *whitelist* che come *blacklist*, viene considerata solo la *whitelist*.

L'utilizzo delle *whitelist* è consigliato per assicurarsi di ricevere tutti i mail inviali da un particolare indirizzo (persone note, liste di posta). L'utilizzo delle *blacklist* è consigliato nel caso si ricevano ripetutamente mail di *spam* da uno stesso mittente, che si ritiene di poter cestinare.
Dato che l'uso di queste procedure può avere influenze importanti sulla ricezione della propria posta, soprattutto in caso di errori di

definizione, si consiglia di utilizzare le *white/blacklist* con estrema cura ed attenzione, iniziando con prove semplici, magari su di un singolo indirizzo ed arricchendole solo quando ciò sia realmente necessario.

Importante: quando si inviano mail a PMAS, sia per definire regole di *white/blacklist*, che per ricevere mail posti in quarantena, è bene assicurarsi di inviare solo messaggi testuali (*plain text*), senza l'eventuale copia in formato HTML, né *attchment*.

Interpretazione degli header X-PMAS

PMAS aggiunge agli *header* dei mail che vengono inoltrati, delle righe che indicano il passaggio attraverso la sua analisi. Tali righe iniziano sempre con il testo "X-PMAS-" e vengono aggiunte per facilitare la configurazione di filtri sui client di posta, in modo da catalogare i mail in base a tali campi.

In particolare i campi utili a tali filtri sono:

`X-PMAS-Final-Score: .....`
in cui viene indicato il valore del punteggio ottenuto tramite PMAS (più alto è il punteggio maggiore è la probabilità che il mail sia *spam*)

`X-PMAS-Spam-Level: +++...`
in cui il livello di *spam* viene indicato con un carattere (in questo caso +) per ogni punto

`X-PMAS-Not-Positive: -2.0`
nel caso il punteggio ottenuto sia negativo (il mail con tutta probabilità è legittimo)

`X-PMAS-Quarantined: PreciseMail`
nel caso il mail sia stato rilasciato da quarantena.

Infine, il *Subject* del mail può essere modificato, in base al punteggio ottenuto, per facilitare all'utente l'individuazione dei messaggi di *spam*:

```
Subject: [POSSIBILE Spam = 4.5]
Orig_subject
```

Attualmente la frase che viene aggiunta è proprio:

`[POSSIBILE Spam = NUM]`

dove *NUM* è il valore numerico calcolato dello *spam*.

La corretta configurazione del client comunemente utilizzato per leggere la posta in base a tali campi e al contenuto del *subject*, unitamente alla gestione dei mail posti in quarantena, dovrebbe permettere di eliminare (o per lo meno ridurre significativamente) l'ormai quotidiana fase di eliminazione manuale di tutti i messaggi reclamizzanti prodotti di dubbia utilità e gusto di cui purtroppo tutte le caselle postali sono infestate.

Ove necessario, ulteriori informazioni sull'argomento potranno essere richieste via mail (senza *spam*, grazie) all'indirizzo:

`staff_posta@cilea.it`