

# Antispam: un approccio centrato sull'utente

Diego Fantoma – Centro Servizi Informatici Facoltà di Psicologia  
*diego@fantoma.it*

In collaborazione con:

Stefano Catani – Centro Servizi Informatici d'Ateneo

Dott. Riccardo Fattorini – Dipartimento di Psicologia - L.I.A.

Presentazione rilasciata sotto licenza GPL



Università degli Studi di Trieste

Facoltà di Psicologia

[www.psico.units.it/fac](http://www.psico.units.it/fac) Tel. +39-040.558.2786 Via S. Anastasio, 12 I-34134 Trieste, ITALIA

[www.psico.units.it](http://www.psico.units.it)



Completare la seguente frase:

**Lo spam è difficile da.....**



# Lo spam è difficile da Identificare da un punto di vista computazionale



## C'era una volta un antispam

Inizialmente si è scelto di attivare un server sperimentale da anteporre a quello di posta, che svolga funzioni di antivirus e antispam per i domini della sede periferica della Facoltà di Psicologia.

La configurazione scelta è la classica linux-oriented postfix + amavis (*sophos, clamav, sophie*) + spamassassin + razor

L'idea era di ridurre il traffico dovuto allo spam sulla linea di connessione, con traffico a pagamento, tra la sede centrale (C.S.I.A.) e quella periferica, oltre a ridurre il carico del server e le lamentele degli utenti.

La configurazione iniziale prevedeva il blocco di messaggi infetti con l'invio a mittente e destinatario di un alert ed il pass-through dei messaggi riconosciuti come spam con anteposizione della stringa "\*\*\* SPAM \*\*\*" nel subject poi sostituita con "\* PROBABILE SPAM \*".

Per poter avere dei dati statistici si scelto di far passare tutto il sistema attraverso uno script di content filtering di secondo tipo (sulle queue) che memorizzasse i dati e consentisse di operare molto granularmente sulle decisioni in base ai risultati dell'analisi di ciascun messaggio

## Alcuni risultati sullo spam

La sperimentazione si è svolta in due fasi, diverse per l'uso delle RBL e per il posizionamento dello script nel processo di trattamento del messaggio.

### *Prima fase*

Su 18.047 messaggi analizzati solo 2.854 venivano riconosciuti come spam (~16%) su un totale di circa 330MB di messaggi trasferiti: si sarebbero risparmiati 13MB. La durata media di processo è stata di circa 1 minuto a messaggio ma con un totale di 15.923 messaggi elaborati in meno di 10 secondi (~88,2%).

### *Seconda fase*

Su 14097 messaggi analizzati vengono riconosciuti come spam 3101 (~22%). Ma da questo totale bisogna togliere tutti i rejected (4370, ~31%) ottenendo una base di 9724 messaggi da discriminare: su questi la percentuale di spam va ad essere pari al 30%, fino ad ottenere un insieme di messaggi giudicati corretti pari a 6626, ossia circa il 47% del totale dei messaggi ricevuti. Su complessivi 660 MB si sarebbero quindi risparmiati (attenzione: non si considerino i rejected che non generano traffico) circa 35MB.

In entrambi i casi la frequenza di superamento di soglia è più alta all'inizio del periodo di osservazione e più bassa al termine.

# La prima sensazione degli utenti - antispam

Curiosamente, nonostante sia stato inviato un messaggio per segnalare agli utenti l'introduzione del servizio e la spiegazione di come filtrare i messaggi di spam...

pochissimi hanno letto il messaggio in quanto lo reputavano spam e quindi:

gli utenti continuavano a lamentarsi dello spam e a chiedere una soluzione;

nonostante le istruzioni con le immagini su come configurare un filtro sul proprio client (outlook express), nessuno l'ha fatto;

nessuno si è reso conto che il messaggio inviato aveva la riduzione dello spam come oggetto.

Molti mi chiedevano come eliminare lo spam per paura di venire infettati

dopo iterate spiegazioni ad personam ed aver attuato la configurazione dei client gli utenti non capivano "perché avere due caselle di posta"

Evidentemente era l'approccio sbagliato nel rapporto antispam - utenti



## La prima sensazione degli utenti - antivirus

Si è scelto di eliminare i messaggi contenenti virus e di inviarne la segnalazione a mittente e destinatario

minor (teorico visto il posizionamento) consumo di banda dovuto all'assenza di traferimento di file eseguibili (un risparmio complessivo di circa il 24%);

maggiore traffico di email (per ciascun virus ricevuto, due messaggi generati);

abnorme incremento di messaggi di errore relativi a indirizzi non esistenti, dovuti al fatto che i virus possono utilizzare indirizzi casuali come mittenti: generare l>alert verso un indirizzo non esistente creava questo problema;

sensibile disagio degli utenti che pensavano fosse arrivato un virus anziché l>alert;

sensibile disagio di utenti di altre strutture che si vedevano recapitare l>alert come mittenti e cercavano di capire quando poteva essere stato spedito il virus, chiedendosi (-mi) se fossero infetti.

Evidentemente era l'approccio sbagliato nel rapporto antivirus - utenti



## La prima sensazione del system admin

Gli utenti:

non conoscono la differenza tra spam e virus

non leggono i messaggi se contengono nel subject le parole "spam" o "virus"

ma, soprattutto...

non leggono i miei messaggi!!

*e tutto ciò è molto triste.*

## Ipotesi per l'antivirus

Perché inviare un alert al mittente e al destinatario?

Raramente il mittente è quello "vero": non si invia più il messaggio al mittente.

Viceversa, se il mittente fosse vero, sarebbe gradito che il destinatario, che magari attendeva un file con ansia, sappia che non l'ha potuto ricevere a causa di un virus e possa quindi richiedere la trasmissione di un nuovo file, avvisando del problema il mittente.

Ma poiché gli utenti non leggono questi messaggi scambiandoli per virus si decide di non inviare più il messaggio nemmeno ai destinatari.



## Ipotesi per l'antispam

Il problema è decisamente più complesso per vari motivi:

Innanzitutto lo spam è più facilmente confondibile producendo falsi positivi, poi c'è il problema di definirlo in modo computabile.

Ad esempio si potrebbe affermare che il tipico messaggio "Enlarge your penis" è spam ma se fossi un americano con oggettivi problemi di questo tipo?

C'è anche un aspetto legale da considerare: bloccare la posta è vietato; nel caso dell'antivirus non lo è poiché sono dichiaratamente prodotti atti al danneggiamento di un sistema informatico (legge n. 547 dd, 23 dicembre 1993 e succ. mod.).

Bisogna quindi verificare quali sono le esigenze di un complesso di utenti



## Analisi delle problematiche sullo spam

Gli utenti non sono in grado di discernere lo spam se non dopo averlo letto

particolari avvisi non servono

non serve inserire parole chiave e filtrare i messaggi sul client o duplicare le caselle di posta

abbiamo come parametri dello script mittente e destinatario

abbiamo un file con una copia del messaggio

dobbiamo risparmiare traffico

A questo punto si interroga lo psicologo a portata di mano.



## Un primo passo verso la soluzione

Si preparano alcuni strumenti per effettuare dei test psicologici di interazione uomo-macchina ipotizzando alcuni scenari:

messaggi con contenuti in html che evidenzino l'eventuale spam e con opportune istruzioni

messaggi che rimandino ad un webmail funzionante sul server antispam per la lettura dei messaggi marcati

l'invio ad ogni fine giornata di un messaggio contenente in attachment tutti quelli marcati come spam

una procedura che ogni tot tempo invii un messaggio a ciascun utente pregandolo di verificare sul webmail detto prima se tra i messaggi ci fossero alcuni interessanti, comprensivo di un elenco dei mittenti.



# Un secondo passo verso la soluzione

I migliori risultati si ottengono con le ultime due soluzioni, in cui i soggetti vengono ad interessarsi del contenuto. I messaggi di avviso, naturalmente, devono essere molto semplici.

L'ultima è preferibile per evitare il traffico indesiderato.

Il problema è a questo punto definire dopo quanto tempo inviare il messaggio e verificare se sia possibile tenere traccia di quanto ciascun utente sfrutti il sistema e quanto interagisca con i messaggi che lui stesso giudica spam (rapporto tra messaggi elencati e messaggi di cui è stato visualizzato il contenuto).

In questo caso si deve tenere presente che se si mantiene costante l'indirizzo web cui rivolgersi dopo un po' si ha un apprendimento operativo da parte dell'utente.



## Finalmente la soluzione

Si è deciso di operare nel seguente modo:

Per ciascun utente si memorizza data e ora dell'ultima visita al webmail e il conteggio dei messaggi di cui si è presa visione.

Devono passare almeno 48 ore dall'ultima visita prima di inviare il messaggio ma se l'utente non ha messaggi nuovi o ha guardato almeno due contenuti completi si ritarda di un giorno l'invio.

Se l'utente visita il webmail da almeno 18 ore dall'ultimo messaggio inviato, si può pensare che abbia acquisito una certa indipendenza nel verificare i messaggi per cui il tempo minimo diventa di 72 ore e tale permane fino ad un massimo di tre segnalazioni non soddisfatte.

Viene tenuta traccia dei messaggi cancellati per scopi futuri.

## Conclusioni

Attualmente il sistema è nella sua seconda fase di test.

L'utenza, intervistata, ha gradito la soluzione proposta e il numero di visite sui webmail conferma l'utilizzo.

Attualmente non si hanno ancora dati analizzati per trarre le conclusioni definitive ma vi sono i presupposti per pensare al successo del sistema.

Per il futuro si pensa di implementare un algoritmo statistico per effettuare una seconda cernita dei messaggi sulla base delle modalità di interazione degli utenti.



**Grazie**

---



Università degli Studi di Trieste

**Facoltà di Psicologia**

www.psico.units.it/fac Tel. +39-040.558.2786 Via S. Anastasio, 12 I-34134 Trieste, ITALIA

www.psico.units.it

