

MATEMATICA INTORNO A NOI

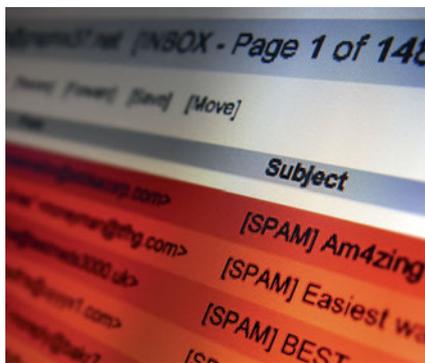
I filtri antispam

Le e-mail indesiderate, o spam, sono una seccatura. Oltre a intasare le caselle di posta elettronica, causano un notevole dispendio di tempo ed energie.

Come si possono bloccare le e-mail di spam, sfruttando il teorema di Bayes?

LA RISPOSTA

Si calcola che attualmente più del 50% delle e-mail che viaggiano nella rete sia costituito da messaggi di spam. Il fastidioso fenomeno, esploso alla fine degli anni Novanta, ha notevoli ripercussioni economiche: le energie complessivamente impiegate nel mondo per passare al vaglio e cancellare questi messaggi causano perdite per svariati miliardi di euro.



Di fronte alla fiumana di spam

Molti fornitori di servizi internet hanno adottato dei filtri antispam che aiutano l'utente a scartare i messaggi di posta elettronica indesiderata. Si tratta di programmi informatici che determinano, ogni qual volta giunge una nuova e-mail, se il messaggio è legittimo, e quindi va recapitato, oppure spam, e quindi va cestinato.

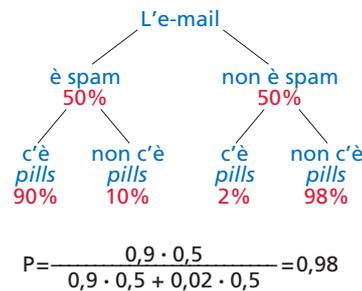
Il funzionamento di questi programmi del computer, alla cui scrittura i ricercatori hanno lavorato per diversi anni, si basa sulla teoria delle probabilità. I filtri antispam non fanno altro che stimare automaticamente la probabilità che un messaggio e-mail in arrivo sia di spam o meno.

Sulla base di una certa casistica, il programma informatico viene inizialmente istruito sulla probabilità che una parola sia contenuta in una e-mail buona o indesiderata. Per esempio, poniamo che il 90% delle e-mail di spam contenga la parola *pills*. Se il computer bloccasse qualunque messaggio e-mail contenente questa parola, si avrebbero molti «falsi positivi», ovvero finirebbero eliminati anche i messaggi assolutamente legittimi, contenenti la parola *pills* ed erroneamente classificati come spam. Ma non solo. Gli spammer imparerebbero ad aggirare il problema, alterando leggermente la grafia della parola e scrivendo, per esempio, *pills* o *pi1ls* al posto di *pills*.

Il programma antispam, quindi, non può cestinare un'e-mail solo perché contiene una parola altamente sospetta. Deve tener conto di tutte le altre parole presenti nell'e-mail.

Ecco come funziona il filtro

Il computer assume inizialmente che ciascun messaggio abbia una probabilità a priori del 50% di essere spam. Il filtro utilizza poi la formula di Bayes per calcolare la probabilità condizionata *P* che, data una certa parola, per esempio *pills*, il messaggio sia spam.



Combina poi queste probabilità in una probabilità complessiva che il messaggio sia spam, sulla base di tutti i vocaboli che compaiono.

L'indice di spamicity

Il risultato finale, ovvero la probabilità bayesiana a posteriori che il messaggio sia effettivamente spam, è un indice chiamato *spamicity*. Una volta calcolato il valore di *spamicity*, basta stabilirne la soglia (per esempio, superiore al 90% o 0,9), oltre la quale il computer smista il messaggio come spam e lo elimina.

Il criterio descritto, che passa al vaglio le parole, può essere integrato anche con altri fattori, per esempio l'eventualità che intere righe del messaggio siano scritte in lettere maiuscole (opzione spesso preferita dagli spammer).

Alcuni filtri molto avanzati sono addirittura in grado di eseguire calcoli personalizzati per ciascun utente, cioè di aggiornare i loro indici di probabilità sulla base delle e-mail ricevute presenti nella posta elettronica.

L'origine dello spam è in una scatoletta di carne

SPAM è una marca di carne in scatola americana realmente esistente. Negli anni Settanta, spinti dall'invasiva politica pubblicitaria della casa produttrice di SPAM, i Monty Python, famoso gruppo di attori comici, proposero uno sketch di denuncia in cui la cameriera nevrastenica di un fast food proponeva, con tono sempre più fastidioso e incalzante, piatti esclusivamente a base di carne SPAM: uova e SPAM, uova, bacon e SPAM, uova, bacon, salsicce e SPAM. Spam è diventato così sinonimo di messaggio pubblicitario insistente e non desiderato.

