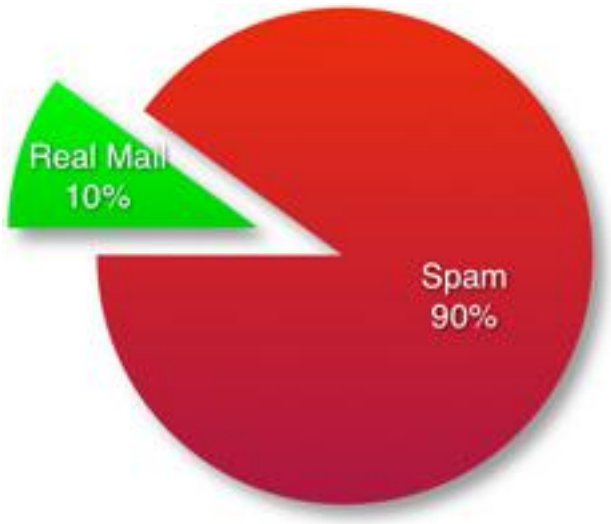
I messaggi di SPAM, possibili soluzioni

Antonio De Martino

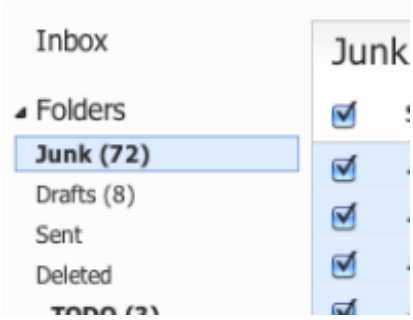
Una e-mail di spam è una comunicazione pubblicitaria, promozionale, di informazione commerciale o di vendita diretta che viene recapitata nella cassetta di posta elettronica di migliaia di utenti che non hanno la facoltà di revocare al mittente l'autorizzazione all'invio.
Spesso chiamate bulk e-mail oppure junk e-mail, le e-mail di spam sono proposte con una frequenza molto elevata ed in alcuni casi superano il numero di messaggi reali al punto tale da indurre utenti e associazioni per la tutela e i diritti dei consumatori a richiedere alle autorità competenti provvedimenti che ne regolino il corretto utilizzo. Alcuni dati confermano che ad oggi il 90% delle comunicazioni di posta elettronica è spam.



Fonte www.spamhaus.org

La risorsa fondamentale per un produttore di spam è rappresentata da indirizzi e-mail validi.
Gli indirizzi e-mail possono essere acquistati, rubati grazie alle intrusioni informatiche nei sistemi delle aziende, oppure cercati in rete grazie a programmi chiamati Spambot, programmi automatici che setacciano il web alla ricerca di parole che contengano il simbolo @. In altri casi sono presi di mira i provider di servizi più importanti (es. google.com, yahoo.com, virgilio.it ecc.) e con dizionari predefiniti si procede alla costruzione automatica degli indirizzi.
Completata la fase di raccolta, viene effettuato un primo invio di spam per ottenere una lista valida di indirizzi da contattare.

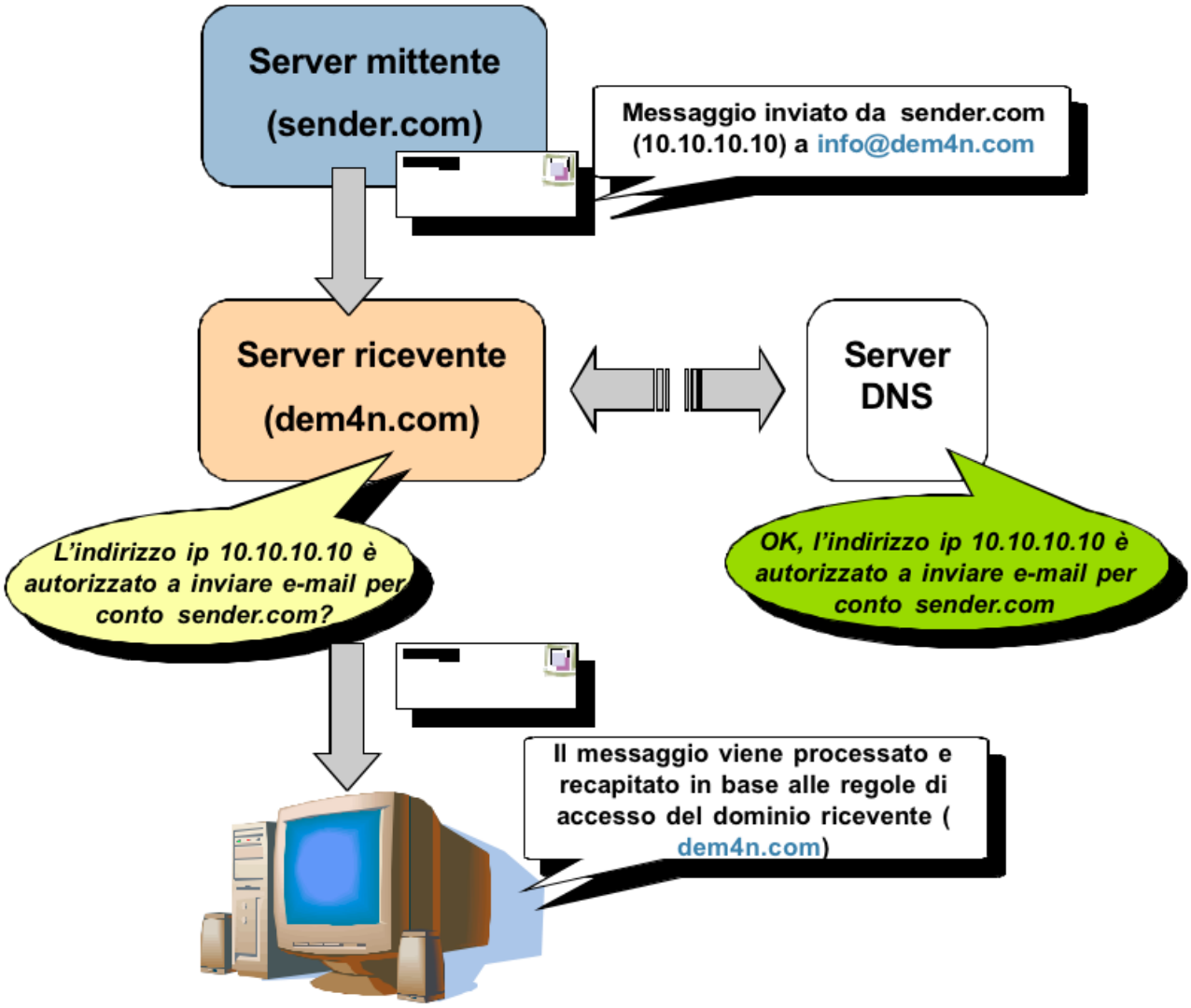
Semplificando, le tipiche architetture di posta elettronica prevedono nella catena di ricezione un sistema anti-spam esposto direttamente sulla Internet, a seguire un sistema anti-virus e gli altri elementi necessari.
Il compito del meccanismo anti-spam è quello di filtrare la posta in arrivo cercando di individuare i messaggi junk e-mail. Questa operazione viene eseguita principalmente attraverso block list basate sui DNS e black list di indirizzi ip o e-mail definite direttamente sui sistemi di posta e dalle quali non sono accettate le comunicazioni.



A questo si aggiunge una funzione di filtro probabilistico che può essere raffinato con l'azione dell'utente ha la facoltà di marcare un messaggio come spam: i messaggi vengono quindi classificati e segnalati come probabile spam in base alle parole contenute e sono consegnati in una cartella dedicata. L'azione applicata dall'utente contribuisce ad innalzare la probabilità che messaggi simili possano essere indesiderati.
Questo tipo di filtro è adattivo, ovvero cambia a seconda delle preferenze dell'utente e diventa molto preciso nel tempo. Si parla di filtri bayesiani, dal nome dell'omonimo matematico Thomas Bayes.

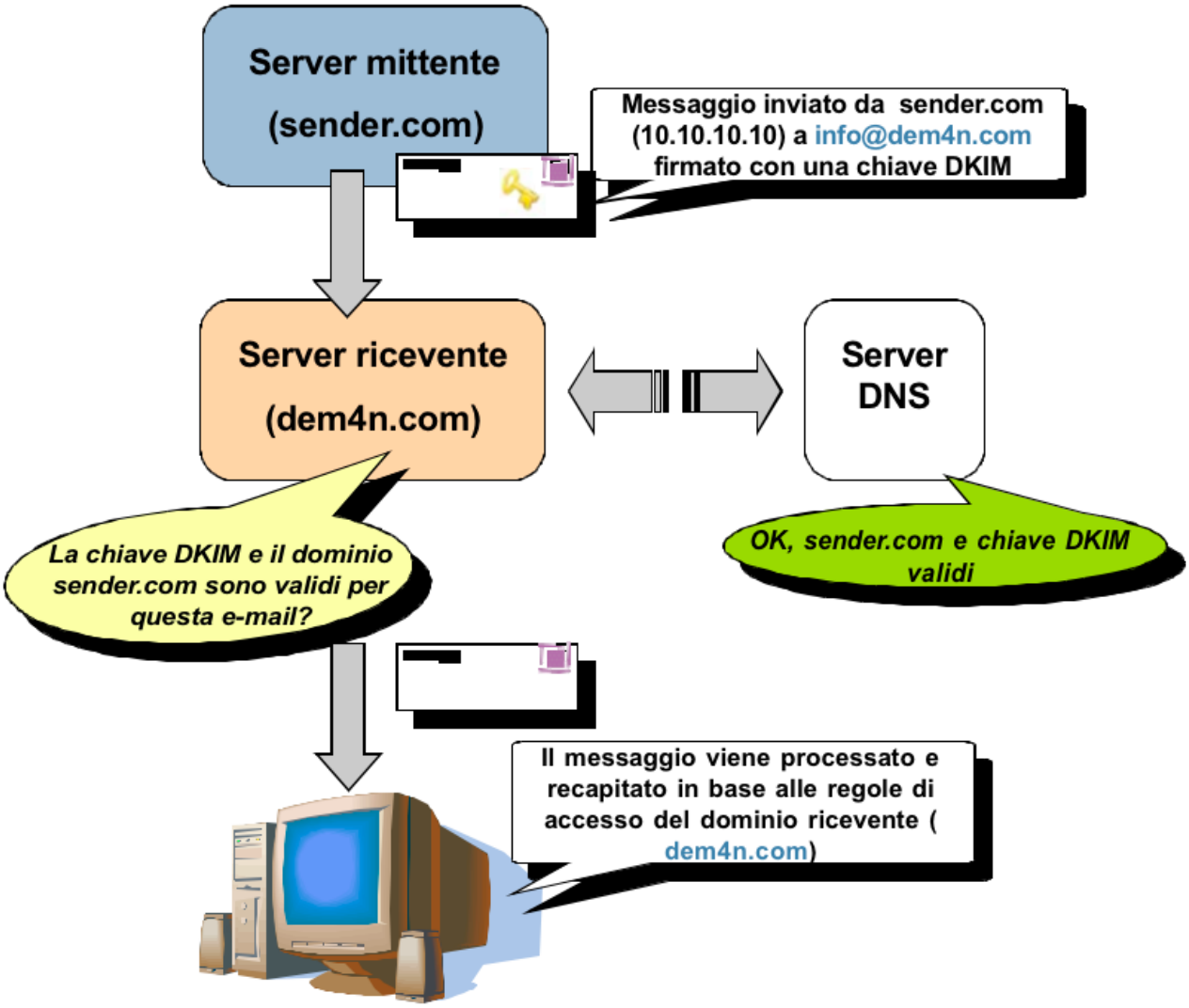
Il fenomeno dello spam causa "una lesione ingiustificata dei diritti dei destinatari", che sono costretti a sprecare del tempo per selezionare i messaggi di interesse tra tutti quelli ricevuti, con un aggravio dei costi per il collegamento telefonico che possono crescere in proporzione alla quantità di e-mail da verificare. Non sono trascurabili i costi in termini di tecnologie e processi organizzativi che le aziende devono sostenere per contrastare tale problematica.
Inoltre gli indirizzi di posta elettronica recano dati di carattere personale e sono da trattare nel rispetto della normativa in materia (art. 130, legge n. 196/03). "Il loro utilizzo per scopi promozionali e pubblicitari è possibile solo se il soggetto cui riferiscono i dati ha manifestato in precedenza un consenso libero, specifico e informato."
Il quadro normativo Italiano basato sulla scelta dell'interessato di opt-in, seguito dalla direttiva Europea n. 2002/58/CE, tutela in questo senso gli utenti che, in caso di violazione, possono in alcuni casi rendere ap- plicabile anche una sanzione penale. La prima condanna in Italia per lo spam è stata emessa il 16/12/2010 ai danni di un'azienda che produce e vende contenuti multimediali.
Le block list possono costituire una soluzione allo spam, ma sono difficili da manutenere anche perché gli spammer utilizzano per il loro scopo computer compromessi da virus appartenenti ad aziende o persone molto spesso ignare. E' possibile agire a livello di rete bloccando tutto il traffico in uscita verso la porta 25 fino al ripristino del normale funzionamento dei sistemi, ma questa soluzione comporta in ogni caso il blocco anche della posta legittima.
Un'idea per creare una ulteriore soluzione al problema, da aggiungere ai filtri probabilistici, potrebbe essere quella di agire sulla validità di un indirizzo e-mail e ingannare i sistemi degli spammer. Sfruttando opportunamente i messaggi di errore del protocollo SMTP è possibile generare per il sistema mittente l'errore "511 Sorry, Recipient address has invalid format" oppure "550 Requested action not taken: mailbox unavailable". Il sistema che produce spam, non solo si

troverebbe inondato di risposte non valide, ma dopo i vari tentativi di invio predefiniti cesserebbe di mandare e-mail agli indirizzi che hanno generato gli errori 511 e 550.
Per combattere il fenomeno i service provider più bersagliati dallo spam (Google, Microsoft, AOL e Yahoo!) hanno proposto all'IETF (Internet Engineering Task Force) l'introduzione di una metodologia chiamata "Domain-based Message Authentication, Reporting & Conformance" in breve DMARC. Semplificando, il DMARC si basa su meccanismi già noti ovvero SPF (Sender Policy Framework) e DKIM (Domainkeys Identified Mail) che forniscono una sorta di autenticazione del server mittente cercando di accertarne l'identità.



Attraverso il Sender Policy Framework, un messaggio di posta elettronica inviato da un dominio mittente (es. sender.com) viene accettato dal server di posta del dominio destinatario (es. dem4n.com) solo se l'indirizzo IP del mittente è autorizzato a inviare email per conto del dominio con cui si presenta, nell'esempio sender.com.
Il server di posta destinatario ottiene la lista degli indirizzi IP autorizzati, a inviare posta elettronica per uno specifico dominio, attraverso un'opportuna interrogazione di un server DNS (Domain Name Server).
In altri termini, il ruolo del server DNS è quello di fornire in tempo reale la lista degli indirizzi autorizzati all'invio di posta verso una destinazione, una sorta di block list dinamica. In questo

scenario il computer destinatario confronta la lista di indirizzi ottenuta dal DNS con l'indirizzo IP del mittente, se esiste una corrispondenza allora il messaggio è regolarmente accettato e processato altrimenti è considerato un tentativo di invio di spam e viene scartato.



L'idea dell'architettura Domainkeys Identified Mail è leggermente più complessa e si basa sull'utilizzo dei certificati. Analogamente a SPF, esiste un server DNS al quale il ricevente richiede informazioni per autenticare il mittente prima di accettare il messaggio, ma nel caso DKIM le informazioni richieste sono una chiave di cifratura utilizzata per verificare l'autenticità del messaggio. In questo schema solo il mittente, possessore della chiave privata, può cifrare il messaggio da inviare garantendo univocamente la propria identità.
Il destinatario potrà essere certo dell'identità del mittente se riesce a decifrare il messaggio con il certificato fornito dal DNS. E' importante notare che sia nel caso dell'SPF che in quello del DKIM, ci si concentra sull'identità di chi invia il messaggio, ma non sul contenuto. Un'altra proposta inserita nel DMARC è l'estensione dell'ARF, Abuse Reporting Format, ovvero uno standard per l'interscambio di informazioni tra i gruppi Abuse di diversi Internet Service Provider. L'obiettivo è automatizzare il processo di segnalazione delle mail di spam/phishing.

È curioso conoscere l'origine del termine SPAM. Esso deriva da "SPiced hAM", un cibo in scatola commercializzato da una società americana e introdotto per la prima volta nel 1937. Nel 1972 un gruppo di comici inglesi i "Monty Python" ironizzarono su questo prodotto in uno dei loro sketch (disponibile su youtube.com) e successivamente il termine fu adottato nell'ambiente universitario in informatica per indicare i messaggi di posta elettronica indesiderata.

L'ing. Antonio DE MARTINO è laureato presso l'Università degli studi di Napoli Federico II in Ingegneria Informatica con tesi sull'Enterprise Application Integration. Si è occupato di progettazione di infrastrutture ICT sicure, hardening e sicurezza di apparati di rete, di piattaforme per l'identity management e controllo accessi. Oggi si occupa di sicurezza delle informazioni e project management di architetture di sicurezza informatica.