Técnicas de control de spam

Ferramentas baseadas en software libre

Alberto García

<agarcia@igalia.com>

¿Qué é o spam?

- Basicamente: correo electrónico non solicitado.
- Definición máis precisa:
  - Correo enviado masivamente a moitos usuarios.
  - O correo non foi autorizado nin solicitado polo receptor.
  - O contido da mensaxe non é dependente do destinatario.

¿Por qué existe tanto?

- O protocolo SMTP é tradicionalmente “aberto”:
  - É trivial enviar un correo cun remite falso.
  - Un correo pode pasar por moitas máquinas antes de chegar ao destinatario: máis díficil de rastrexar.
  - Pódese enviar correo desde calquera máquina de Internet a calquera parte.
- En definitiva: é sinxelo, barato e difícil de controlar.

Relay *aberto de correo*

- É a máquina que permite a calquera enviar correo electrónico a todas partes.
- Este tipo de funcionamento era o habitual nos primeiros tempos de Internet.
- Desde mediados dos 90, o uso de servidores de terceiros para enviar spam foise facendo máis popular.
- Hoxe en día un *relay* aberto considérase problemático e acostúmase a bloquear o correo que veña del.

Controlando os relays

- Algunhas das técnicas máis comúns (todas implementables usando software libre):
  - Control por IP de orixe: só se permite o *relay* a certos IPs. É a máis sinxela de todas.
  - SMTP-AUTH: Permítese o *relay* aos usuarios que se autentiquen usando esta extensión de SMTP. Técnica vulnerable ao roubo de contrasinais.
  - Pop-Before-SMTP: O servidor permite o *relay* aos usuarios que se teñan autenticado previamente via POP3.

Detendo o spam

- A medida que o spam se converte nun problema máis e máis importante, comeza a aparecer software que o identifica e/ou detén automaticamente.
- Existen diversas técnicas para afrontar este problema, que poden funcionar de forma independiente.
- A continuación comentaranse algunhas destas técnicas e software disponible para implementalas.

DNS-based Blackhole Lists (DNSBL)

- Servidores DNS con IPs de spammers.
- Actualízanse constantemente.
- Existe unha grande cantidade disponible, tanto gratuitos como de pagamento.
- Os MTAs máis comúns permiten usalos.
- Débese complementar con outras técnicas para obter unha boa efectividade.

Identificación con heurísticas (1)

- O spam ten unha serie de características que no se ven normalmente nos correos convencionais:
  - Asunto da mensaxe en maiúsculas.
  - Links a páxinas externas.
  - Corpo só en HTML.
  - Referencias a medicamentos.
  - etc.
- Cantas máis características sospeitosas conteña unha mensaxe, máis susceptible de ser spam.

Identificación con heurísticas (2)

- Vantaxes deste método:
  - Funciona razoablemente ben cunha parte importante do spam.
- Problemas deste método:
  - É necesario manter unha base de datos de heurísticas constantemente actualizada.
  - O *spammer* tan só precisa adaptar as súas mensaxes para burlar as heurísticas.
  - Moitas mensaxes de spam non son correctamente identificadas, e en ocasións o sistema produce falsos positivos.

Filtrado bayesiano (1)

- O sistema antispam é entrenado previamente con un número de correos dos dous tipos (lexítimos e spam).
- Compútase unha estatística das palabras que aparecen en ambos tipos de correos.
- Cando chega un correo novo calcúlase se é spam ou non en función das palabras que contén.
- Existen outras variantes de métodos estatísiticos para a identificación de spam.

Filtrado bayesiano (2)

- Vantaxes deste método:
  - Grande precisión na identificación de spam.
  - Número moi reducido de falsos positivos.
  - Adáptase ben aos novos tipos de spam.
- Problemas deste método:
  - Requere que o usuario entrene o sistema.
  - Son necesarias bastantes mensaxes de ambos tipos antes de obter uns resultados razoables.

Filtrado baseado en checksums

- Unha mesma mensaxe de spam chega a milleiros de destinatarios.
- É posible calcular un *hash* dunha mensaxe que permita identificalo a outros usuarios.
- Existen redes que manteñen “catálogos” de spam enviados polos usuarios.
- *Vipul’s Razor*: software libre baseado nesta tecnoloxía.

Hashcash

- En cada correo lexítimo envíase unha cabeceira *única*, por exemplo un *hash* SHA-1 baseado en:
  - O enderezo do destinatario.
  - Un *timestamp*.
  - Un número grande de bits a 0.
- Canto maior sexa o número de bits, o cálculo da suma será computacionalmente máis complexo.
- Introduce un pequeno retarto para alguén que envía un correo lexítimo, pero supón un grande custo para alguén que envía milleiros de correos á vez.

Greylisting *(1)*

- A maior parte do spam envíase con software específico para este propósito.
- A eficacia do *greylisting* baséase na implementación do protocolo SMTP que acostuman ter estes programas.
- Esta técnica é totalmente independente do contido das mensaxes.
- Está pensada para ser complementaria ás outras vistas anteriormente.

Greylisting *(2)*

- Funcionamento:
  - Considéranse tres datos de cada mensaxe que chega ao servidor: IP de orixe, remetente e destinatario.
  - A primeira vez que se recibe unha terna, rexéitase *temporalmente* co código SMTP 451 e colócase nunha “lista gris”.
  - Se se repite a mesma terna pasado un tempo breve (15 minutos, 1 hora, ...) acéptase e colócase nunha “lista branca”.
  - Se non hai reintentos dun elemento da lista gris en varias horas, elimínase da lista.

Greylisting *(3)*

- ¿Por qué é efectivo este método?:
  - O software SMTP convencional (Sendmail, Qmail, Exim, Postfix, ...) reacciona correctamente ante erros temporais.
  - O spam envíase normalmente con software específico, que ignora este tipo de erros.
  - Sobrepasar o *greylisting* obriga ao spammer a manter o msmo IP durante un tempo considerable.

Greylisting *(4)*

- Vantaxes deste método:
  - Fácil de configurar, non requere case mantemento.
  - Consume poucos recursos.
  - Grande efectividade.
- Problemas deste método:
  - Deberíase combinar con outra técnica para eliminar o spam que pasa este filtro.
  - Introduce un retardo nos correos recibidos.
  - Se o servidor que envía non funciona conforme ao estándar é posible que se perdan correos.

SpamAssassin

- Unha das ferramentas máis populares actualmente.
- Programado en Perl usando a licenza libre de Apache.
- Dispón dunha base de datos de heurísiticas para identificar spam.
- Desde fai bastante tempo tamén incorpora filtrado bayesiano.
- Integrable con procmail, sendmail, postfix, ...

DSPAM

- Software antispam con licenza GNU GPL.
- Programado en C e pensado para ser eficiente e escalable.
- Baséase exclusivamente en métodos estatísticos.
- Promete unha efectividade superior ao 99 %.
- Integrable coa maior parte dos MTAs.
- Incorpora unha interface web de administración.
- Fácil de entrenar: mediante un *forward* ou via web.
- Para Debian: `http://people.igalia.com/berto`

Greylistd

- Implementación de greylist para Exim.
- Programada en python para o sistema Debian GNU/Linux.
- Licenza GNU GPL.
- Instalación trivial
- Requere pouquísima configuración e mantemento.
- Existen implementacións similares para outros MTAs (ej: postgrey para Postfix).

Ligazóns e máis información

- Teoría sobre o filtrado bayesiano:
  `http://www.paulgraham.com/spam.html`
- Teoría sobre *greylisting*:
  `http://projects.puremagic.com/greylisting/`
- Páxina da Wikipedia:
  `http://en.wikipedia.org/wiki/E-mail_spam`
- Proxecto SpamAssassin:
  `http://spamassassin.apache.org/`
- Proxecto DSPAM:
  `http://nuclearelephant.com/projects/dspam/`
  `http://people.igalia.org/berto/`

Técnicas de control de spam

Ferramentas baseadas en software libre

Alberto García

<agarcia@igalia.com>