Filtres Anti-virus/Anti-SPAM pour Sendmail

Matthieu Herrb

27 mars 2003

Introduction

L'internet qui n'est plus ce qu'il était...

Conséquences pour la messagerie :

- virus, vers,...
- spam, pollupostage,...

Les solutions au niveau utilisateur(s) ne sont pas suffisantes...

$\Longrightarrow$ **filtrage au niveau du transport**

Anti-virus

Approches :

- par signature : ce que font les anti-virus traditionnels
  Problèmes :
  « fenêtre de vulnérabilité » entre l'apparition d'un nouveau virus et la mise à jour des anti-virus.
  fausses alertes de plus en plus nombreuses (statistiquement inévitable).
  coût des solutions les plus performantes
- plus simple : bloquer l'exécution des contenus potentiellement dangereux.
  capable de bloquer de nouveaux vers/virus
  pas de faux positifs (par définition).
  problèmes :
  liste exhaustive des types MIME exécutables ?
  comment échanger des exécutables légitimes ?

Sur les contenus exécutables

Problème beaucoup plus vaste que la messagerie

C'est une technologie en plein boum : web services, mises à jour automatiques d'OS ou de logiciels, ...

Problème de sécurité majeur : vecteurs de diffusion de virus/chevaux de troie idéal (plus besoin de débordement de buffer pour exécuter du code malicieux

Les « bacs à sable » (Java, .NET, ...) n'arrivent pas à être imperméables

La signature numérique des contenus exécutables est mal partie

Palladium / TCPA / Liberty Alliance ??

Retour à la messagerie - le spam

SPAM = Messages électroniques non sollicités.

Nuisance de plus en plus importante.
(Contenus douteux...)

Diffusés par des relais ouverts à des listes d'adresses collectées sur le Web.

Pourquoi tant de SPAM ?

Envoyer un message ne coûte presque rien.

Un taux de retour très faible permet de faire des bénéfices.

Aucune chance de voir les spammeurs arrêter : activité lucrative et sans risques.
(Source : Wall Street Journal – http://online.wsj.com/article_email/0,,SB103713867922044714 8,00.html)

Exemple récent : spam vendant des outils anti-spam.

Seul point positif : les spammeurs continueront à utiliser les solutions les moins chères possible pour conserver leur marge.

Politiques de filtrage

Définir avec le conseil de laboratoire / CE une politique de filtrage claire

Basée sur la charte d'utilisation des moyens informatiques

Définir
- ce qui est du trafic légitime,
- ce qui sera bloqué (mis en quarantaine).

Tolérer une utilisation raisonnable à usage extra-professionel.

Respecter les conseils de la CNIL dans les fiches pratiques éditées avec le rapport sur la cyber-surveillance sur les lieux de travail :

http://www.cnil.fr/thematic/docs/entrep/cybersurveillance2.pdf
http://www.cnil.fr/thematic/docs/entrep/cyber_fiches.pdf

SPAM - aspect légaux

Le spam est illégal.

La loi sur l'économie numérique instaure le principe du :
« **consentement préalable** en matière de prospection directe opérée par des systèmes automatisés d'appel, télécopieurs ou courriers électroniques ».

Mais il y a des exceptions...

Cf. Délibération CNIL n° 02-093

LEN : http://www.droit-technologie.org/3_1.asp?legislation_id=138

commentaire : http://www.droit-technologie.org/1_2.asp?actu_id=714

Site général : http://www.spamlaws.com/

[Je ne suis pas juriste]

Lutte contre le SPAM - Recommandations de la CNIL

http://www.cnil.fr/thematic/internet/spam/lacnil_aide1.htm

- Faites toujours preuve de vigilance quand vous communiquez votre adresse électronique.
- Ne rendez pas visible les adresses méls de vos correspondants lorsque vous créez un groupe ou une liste de diffusion.
- Sensibilisez vos enfants sur l'utilisation qu'ils peuvent être amenés à faire de leur adresse électronique
- Ne répondez jamais à un « spam ».
- Ne communiquez pas à des tiers des adresses mél autres que la votre sans le consentement des intéressés.
- Utilisez un filtre de « spam ».
- Ne cliquez pas sur les liens hypertexte insérés dans le corps du « spam ».
- Ne jamais ouvrir un fichier joint figurant dans un « spam ».

En pratique...

Pour l'utilisateur

- Le bouton 'Poubelle' (ou le raccourci clavier) est le moyen le plus rapide et le plus efficace jusqu'à environ 30 messages par jour.
- Au delà : filtrage (traité plus bas)

Pour l'administrateur système

- (In)former ses utilisateurs
  - Cf. recommandations de la CNIL
  - Ne pas devenir spammeurs (organisation de conférences, gestion de listes de mail, etc.)
- Mettre en place un dispositif anti-spam au niveau du serveur de messagerie
- Lutter contre les relais ouverts.
- Dénoncer à la justice les pratiques illégales (pédophilie, « chaînes », etc.)

Chapitre 1

Les black lists

Principe

Rejeter les connexions SMTP en provenance de relais ouverts connus.

Fonctionnement : Un serveur DNS gère un domaine spécial contenant une base de donnée de relais ouverts.

Exemple : Est-ce que `140.93.0.15` (`mail.laas.fr`) est un relais ouvert ?

$\rightarrow$ recherche `15.0.93.140.relays.ordb.org`.

Si réponse, alors c'est un relais ouvert...

Configuration : (sendmail/m4) Dans le fichier `.mc` ajouter :

```
FEATURE('dnsbl','relays.ordb.org','"550 rejected - see http://ordb.org/"')
```

Possibilité d'exceptions dans `accessdb` : `140.93.0.15 OK`

Quelques black lists :

rbl.net	payant
SpamCop.net	payant
relays.ordb.org	http://ordb.org/
spews.relays.osirusoft.com	http://spews.org/
inputs.orbz.org	http://www.orbz.org/

White Lists

Complémentaire extrême des blacks lists :

N'accepter que les messages de personnes identifiées

- despam : http://www.laas.fr/~felix/despam.html
- http://impressive.net/people/gerald/2000/12/spam-filtering.html
- n'accepter que des messages signés (S/MIME ou PGP) avec ou non une liste d'autorités de certification de confiance.

Black lists : bilan

Avantages

- Facile à configurer
- Consomme peu de ressources
- La mise à jour se fait toute seule

Inconvénients

- Pas de contrôle de la qualité des black lists
- Plus facile de rentrer dans une black list que d'en sortir
- Le blocage complet des messages rend la communication avec un site black-listé difficile

Le LAAS a utilisé des black lists de 1997 à janvier 2003.

Chapitre 2

Filtrage sur le contenu

Filtrage sur le contenu : principe

Classifier les messages en fonction du texte complet du message
- Par mots clés / patterns : SpamAssassin
- Par analyse statistique : Classification Bayesienne (bogofilter, bmf, etc.)
- Base de donnée de spams : Vipul's Razor.

SpamAssassin™

http://www.spamassassin.org/

Ensemble de tests sur le contenu
Chaque test attribue des points
Somme des points → score

Marque les messages qui dépassent un seuil.

Écrit en Perl.

Plusieurs modes de fonctionnement :
- filtre simple (utilisation avec procmail)
- filtre client d'un démon (meilleurs performances) (`spamc`)
- au travers de l'API milter (sendmail) → filtrage global

SpamAssassin - Exemple

```
SPAM: -------------------- Start SpamAssassin results ---------------------
SPAM: This mail is probably spam.  The original message has been altered
SPAM: so you can recognise or block similar unwanted mail in future.
SPAM: See http://spamassassin.org/tag/ for more details.
SPAM:
SPAM: Content analysis details:   (8.10 hits, 5 required)
SPAM: SUBJECT_MONTH      (-0.5 points) Subject contains a month name - probable newslet
SPAM: NO_REAL_NAME       (1.3 points)  From: does not include a real name
SPAM: LOSE_POUNDS        (0.5 points)  Subject talks about losing pounds
SPAM: DIET               (0.4 points)  BODY: Lose Weight Spam
SPAM: FULL_REFUND        (0.4 points)  BODY: Offers a full refund
SPAM: CLICK_BELOW        (0.3 points)  BODY: Asks you to click below
SPAM: SPAM_PHRASE_08_13  (1.4 points)  BODY: Spam phrases score is 08 to 13 (medium)
SPAM:                    [score: 8]
SPAM: DATE_IN_FUTURE_06_12 (1.1 points)  Date: is 6 to 12 hours after Received: date
SPAM: RCVD_IN_DSBL       (3.2 points)  RBL: Received via a relay in list.dsbl.org
SPAM:                    [RBL check: found 251.102.96.210.list.dsbl.org]
SPAM:
SPAM: -------------------- End of SpamAssassin results ---------------------
```

SpamAssassin - Installation

Dernière version : 2.51

- Téléchargement : http://www.spamassassin.org/downloads.html
- `perl Makefile.pl`
- `make`
- `make install`

Utilisation simple : dans `${HOME}/.procmailrc` :

```
#----------------------------------------
# Les SPAMs
:0fw
| /usr/local/bin/spamassassin
:0:
* ^X-Spam-Status: YES
spam
#----------------------------------------
```

Tests - SpamAssassin

Sur une base de messages triée à la main.

SpamAssassin :

nature	total	erreurs	pourcentage
HAM	1129	1	0.1%
SPAM	633	22	3.5%

Quelques secondes par message. (spamc)

SpamAssassin - 1er bilan

Utilisation(s) individuelle(s) pendant un an : (configurations par défaut - seuil de 5.0)

- filtrage très efficace
- peu de faux positifs
- nouvelles versions plus efficaces
- problème de lenteur/charge du serveur
  $\rightarrow$ utilisation de `spamd`/`spamc` évite l'initialisation de Perl pour chaque message

Classificateurs bayesiens

Classification bayesienne :
http://www.mathpages.com/home/kmath267.htm

Eric Horvitz & al. (Microsoft) 1998 :
http://research.microsoft.com/~horvitz/junkfilter.htm

Paul Graham 2002-2003 :
http://www.paulgraham.com/spam.html

Adaptive latent semantic analysis (Apple Mail.app) :
http://lsa.colorado.edu/papers/dp1.LSAintro.pdf

Implémentations libres :
http://bogofilter.sourceforge.net/ (Eric S. Raymond)
http://www.sourceforge.net/bmf/
http://spambayes.sourceforge.net/
http://www.mozilla.org/mailnews/spam.html (Mozilla Mail)
http://www.fourmilab.ch/annoyance-filter/

Rappels

$$P(C = c_k|X = x) = \frac{P(X = x|C = c_k)P(C = c_k)}{P(X = x)}$$

Ici deux classes : C_0 HAM
C_1 SPAM

Bayes naïf :

$$P(X = x|C = C_k) = \prod_i P(X_i = x_i|C = c_k)$$

Horvitz & al.

Brevet Microsoft (1998).

Classificateur Bayes naïf.

Utilise une liste de motifs discriminants déterminée par apprentissage. Supprime les mots trop peu fréquents pour être discriminants.

Prise en compte d'éléments spécifiques : proportion de caractères non alphanumériques, domaine de l'expéditeur.

Paul Graham

Décomposition en tokens du texte complet (y compris en-têtes)

Base d'apprentissage classée à la main. Pour chaque token calcule la probabilité d'être SPAM.

Nouveaux messages :
- assigne initialement 0.4 comme probabilité aux mots inconnus ,
- garde les 15 probabilités les plus significatives (les plus éloignées de 0.5),
- calcule la probabilité conditionnelle que le mail soit un SPAM,
- en fonction de la décision recalcule les probabilités des tokens.

Implémentations :
- bogofilter
- bmf

Tests

Base d'apprentissage : 1333 SPAMS 1534 HAM

(Même base de test que SpamAssassin)

bmf

nature	total	erreurs	pourcentage
HAM	1129	6	0.5%
SPAM	633	74	11.7%

bogofilter (Robison-Fisher) :

nature	total	erreurs	pourcentage
HAM	1129	3	0.3%
SPAM	633	133	21.0%

Quelques dizaines de mili-secondes par message.

Classification par score ou probabiliste ?

Qu'y a-t-il derrière un score ? Comment les attribuer ?

Scores $\rightarrow$ difficiles à personnaliser

Probabilités $\rightarrow$ valeurs absolues

Apprentissage personnalisé automatique (2 boutons : SPAM/HAM)

Permet de suivre l'évolution des contenus du SPAM.

Autres critères de filtrage

Rejeter les messages mal formés :
- respect des RFC
- entêtes incohérents (User-Agent forgé par ex.)
- caractères illégaux

$\rightarrow$ rejette SPAM et virus/vers naïfs.

Les pièges à SPAM

Principe : pénaliser les diffuseurs de SPAM en ralentissant jusqu'à l'insupportable la transaction SMTP s'il s'agit de SPAM.

- **OpenBSD : spamd**
  http://www.openbsd.org/cgi-bin/man.cgi?query=spamd&manpath=OpenBSD+Current&format=html
  Basé sur des black lists et sur le filtre de packets **pf** pour rediriger les connexions des spammeurs vers spamd.
- **Spam tarpit - tarproxy**
  http://www.martiansoftware.com/articles/spammerpain.html
  Utilise un filtre baysien pour analyser le message et ralentir la transaction si le score augmente.

Chapitre 3

Mimedefang - outil général de filtrage avec Sendmail

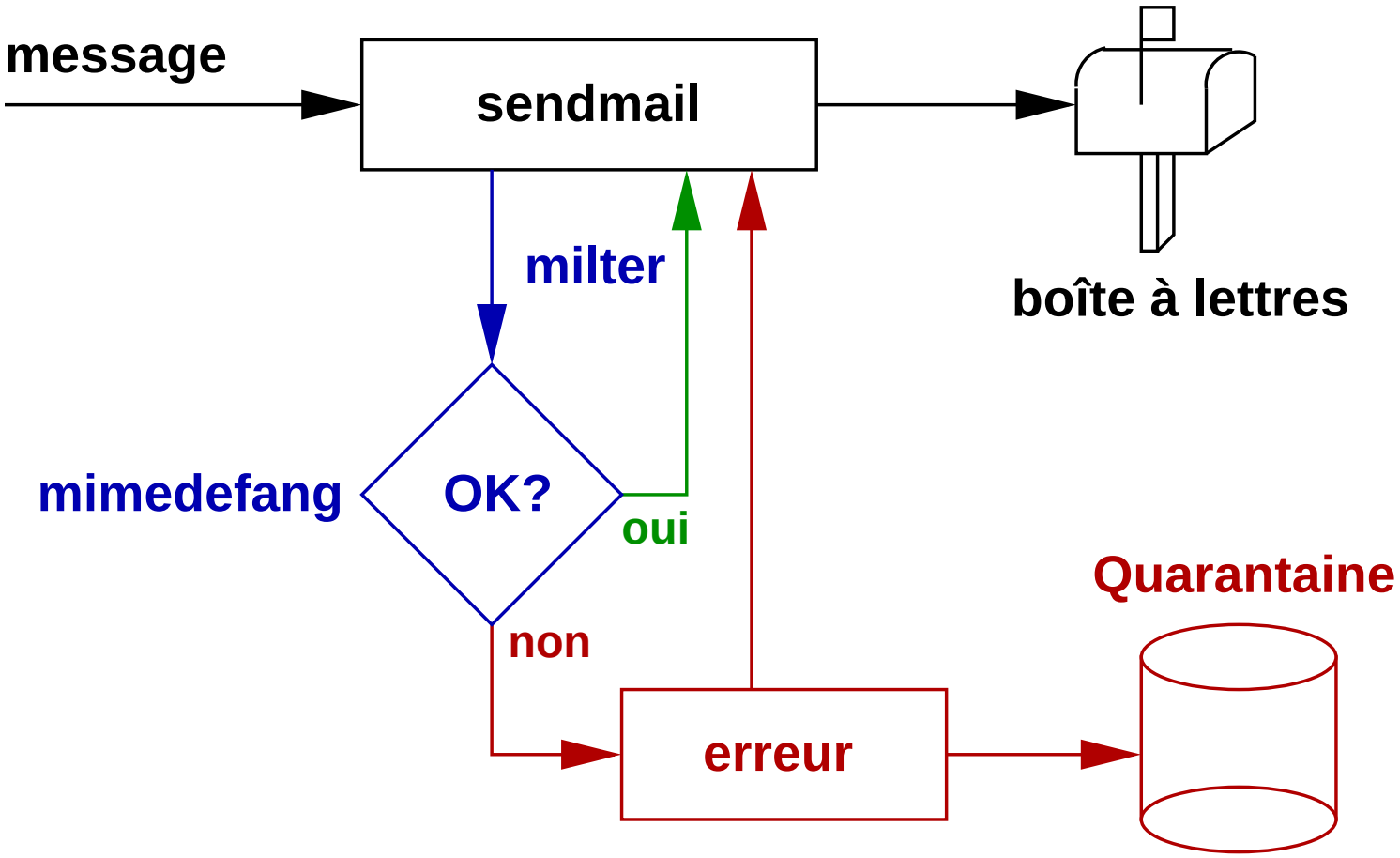
Mimedefang

http://www.roaringpenguin.com/mimedefang/

Programme général de filtrage de mails, utilisant l'API milter de Sendmail.
- destruction, modification ou mise en quarantaine de fichiers attachés « dangereux »
- interface avec Anti-virus existants, avec SpamAssassin
- ajout de notices aux messages
- actions configurables en fonction du domaine, de l'utilisateur, du relais utilisé, etc.
- utilise un « pool » de processus pour les serveurs chargés.
- ...

→ Permet de faire anti-virus et anti-spam en même temps.

Interaction sendmail/mimedefang



Mimedefang - installation (1/2)

Pré-requis :

- Sendmail 8.12.x avec support de milter
- Un certain nombre de modules perl
- Un utilisateur `defang` pour exécuter le démon

```
./configure
make
make install
```

Créer :

- `/etc/mail/mimedefang-filter` - règles de filtrage
- `/var/spool/MIMEDefang` (propriété de defang, mode 700).

Modification du fichier de configuration de sendmail :

```
INPUT_MAIL_FILTER(‘mimedefang’,
    ‘S=unix:/var/spool/mimedefang/mimedefang.sock,T=S:1m;R:1m’)
```

Mimedefang - installation (2/2)

- **Lancer les démons mimedefang avant Sendmail**
  Utiliser le script fournit dans `examples/init-script`
- **Configuration de SpamAssassin pour Mimedefang**
  fichier `/etc/mail/spamassassin/sa-mimedefang.cf`
- **Relancer sendmail**

- **Maintenance**
  Après modification de mimedefang-filter, envoyer SIGHUP au processus `mimedefang-multiplexor`

Mimedefang - filtres

`mimedefang-filter` est un script Perl.

Utiliser SpamAssassin :

Dans `filter_end` :

```
# Spam checks if SpamAssassin is installed
if ($Features{"SpamAssassin"}) {
    if (-s "./INPUTMSG" < 100*1024) {
        # Only scan messages smaller than 100kB.
        my($hits, $req, $names, $report) = spam_assassin_check();
        if ($hits >= $req) {
            action_change_header("X-Spam-Score", "$hits $names");
            action_change_header("X-Spam-Status", "Yes");
        }
    }
}
```

Mimedefang + SpamAssassin pour les utilisateurs

Chaque message soupçonné par par SpamAssassin d'être un SPAM est marqué par des entêtes :

```
X-Spam-Score: 7.9 (*******) ALL_CAP_PORN,CLICK_BELOW,CLICK_BELOW_CAPS,
  CTYPE_JUST_HTML,DATE_MISSING,FROM_ENDS_IN_NUMS,LINES_OF_YELLING,
  LINES_OF_YELLING_2,LINES_OF_YELLING_3,PORN_4,SPAM_PHRASE_05_08,
  SUPERLONG_LINE,UPPERCASE_50_75
X-Spam-Status: Yes
X-Scanned-By: MIMEDefang 2.28 (www . roaringpenguin . com / mimedefang)
```

Filtrage par chaque utilisateur avec :

- procmail (cf plus haut)
- logiciel de messagerie : Eudora, Netscape messenger, ...

Mimedefang + SpamAssassin - bilan

En place au LAAS depuis octobre 2002

Configuration Sun Netra X1 (UltraSparc IIe - 400MHz, 512 Mo RAM)

Mimedefang 2.31 + File::Scan (antivirus) + SpamAssassin

Environ 10000 messages/jour

Pas de problème de charge CPU (Ouf)

Taux de faux-positifs $\leq 1\%$

Besoin d'aide aux utilisateurs pour configurer les MUA !

Dans cette configuration de SpamAssassin :
- pas de personnalisation des filtres
- pas d'analyse des entêtes (messages passés par un relais ouvert)
- seuil fixé à 6.0

Mimedefang - Autres exemples (1)

Alias interne seul

```
sub filter_recipient {
        my ($recipient, $sender, $ip, $hostname, $first, $helo) = @_;
        if ($recipient =~ /^<?all\@laas\.fr>?$/i) {
                if ($sender !~ /\@laas\.fr>?$/i) {
                        return ('REJECT', 'User unknown');
                }
                return ('CONTINUE', 'ok');
        }
}
```

Rejette les messages à l'adresse `all@laas.fr` si l'expéditeur n'est pas dans le domaine `laas.fr`.

Mimedefang - Autres exemples (2)

Anti-spoofing sur HELO/EHLO

```
sub filter_relay {
       my($ip, $name, $helo) = @_;
       if ($helo =~ /laas\.fr/i) {
         if ($ip ne "127.0.0.1" and
             $ip !~ "^140\.93\." and
             $ip !~ "^195\.83\.132\.") {
               return('REJECT', "Go away... $ip is not in laas.fr");
         }
       }
       return ('CONTINUE', "ok");
}
```

Rejette les connexions qui essayent de passer `laas.fr` en paramètre de la requête SMTP **HELO** ou **EHLO**.

Interface SpamAssassin pour les autres MTA

Postfix :
http://www.geocities.com/scottlhenderson/spamfilter.html

Qmail :
http://qmail-scanner.sourceforge.net/

Exim :
http://dman.ddts.net/~dman/config_docs/exim4_spamassassin.html

Conclusion

Conclusion

Filtrage de la messagerie $\rightarrow$ outil de lutte contre les nuisances (Virus, SPAM)

Étapes pour la mise en place :
- définition d'une politique
- information/accord du personnel
- mise en place de la solution technique

Solution technique proposée : **sendmail + mimedefang**
avec les filtres : **SpamAssassin** et **File::Scan**

Futur : filtrage bayesien.