# Le filtre anti-spam MailCleaner

Martin.Ouwehand@epfl.ch, PL-DIT

Les membres de l'EPFL submergés par les messages publicitaires peuvent désormais s'abonner au filtre anti-spam MailCleaner. Ce filtre est basé sur le logiciel SpamAssassin (***http://eu.spamassassin.org/***) qui fait passer à chaque message toute une batterie de tests heuristiques détectant par exemple les tics de langage et les tournures qu'affectionnent les spammers. Chaque test passé octroie un nombre variable de points et le message est considéré comme étant du spam si son score dépasse un certain niveau. Selon le choix de l'abonné, les spams probables sont soit retenus dans une *quarantaine*, soit marqués (le champ **Subject** débute par la chaîne **{Spam?}**). On consultera le mode d'emploi ci-dessous pour l'inscription au filtre, sa configuration et plus de détails sur la gestion de la quarantaine et le traitement des *faux positifs*.
On appelle ainsi un message qui vous est vraiment destiné et qu'il est sans doute important que vous lisiez mais que le filtre marque à tort comme étant du spam. Le but des mainteneurs du filtre est bien sûr que ce cas se produise le plus rarement possible, mais il est impossible de garantir qu'il n'arrivera jamais. **Nous ne saurions donc assez souligner la nécessité que l'utilisateur du filtre prenne ses dispositions pour que ces faux positifs ne lui échappent pas**. Le guide ci-dessous explique comment le faire.

Nous avons confié à l'entreprise régionale (elle est basée à Saint-Sulpice) MailCleaner (***http://www.mailcleaner.net/***) l'installation de ce logiciel sur deux serveurs hébergés dans les locaux du PL-DIT ainsi que sa maintenance (mise à jour des tests). La proximité géographique a permis un dialogue dans la phase de préparation du projet que nous n'aurions sans doute pas eu avec des entreprises basées à l'autre côté de la planète, l'entreprise MailCleaner ayant répondu rapidement à nos suggestions de modification pour adapter son produit à la situation spécifique et souvent exotique de notre site.

Bien que le filtre MailCleaner ait affiché un taux de succès remarquable de l'avis de la vaste majorité des testeurs qui l'ont utilisé durant l'été, je dois attirer l'attention des personnes intéressées sur quelques inconvénients de cette solution et leur rappeler qu'il en existe d'autres. Tout d'abord, MailCleaner ralentit l'acheminement des messages. Alors qu'en temps normal un message parviendra dans votre boîte en quelques secondes tout au plus, le travail de MailCleaner peut prendre une à deux minutes parce qu'il est organisé en *batches* destinés à répartir la charge très importante engendrée par les centaines de tests que subit chaque message. Le détour par un serveur de filtrage crée aussi un risque de panne supplémentaire empêchant les messages de vous parvenir (pour y pallier, le filtre tourne sur deux serveurs redondants). Enfin sur les serveurs filtre il n'y a pas d'autre sauvegarde qu'un back-up disque (configuration RAID). Ceci vous fera peut-être préférer une solution entièrement locale à votre poste de travail, tel que MailShield pour Windows (***http://winsec.epfl.ch/core/index.asp?newcateg=14***) ou les filtres intégrés aux clients mail de Netscape (***http://channels.netscape.com/ns/browsers/mail.jsp#two***) et Mozilla (***http://www.mozilla.org/mailnews/spam.html***). Quant aux as d'Unix/Linux, ils pourront s'essayer à une installation locale de SpamAssassin ou Bogofilter (***http://bogofilter.sourceforge.net/***).

## Mode d'emploi

Vous trouverez en tout temps une version à jour de ce mode d'emploi à l'adresse:
***http//mailwww.epfl.ch/MailCleaner/top.html***.

### Abonnement au filtre

On s'abonne au filtre en en visitant la page ***https://mailwww.epfl.ch/MailCleaner/subscribe.cgi*** (la résiliation de l'abonnement se fait au même endroit). Note: ceci nécessite une authentification auprès de Gaspar.

### Configuration personnelle du filtre

- allez sur la page d'accueil de MailCleaner: ***http://mailcleaner.epfl.ch/*** (note: ceci nécessite une authentification auprès de Gaspar);
- choisissez à gauche l'onglet **paramètres**;
- modifiez les paramètres:
  - **l'action du filtre**: vous avez le choix entre
    - *mise en quarantaine*: les spams sont retenus
    - *marquage du sujet*: les spams parviennent dans votre boîte aux lettres mais sont marqués dans le **Subject**, comme suit:

      **Subject: {Spam?} Make money fast on the Internet !!!**

      Ceci vous permet de les trier dans un *folder* à part dans votre outil de mail.
  - **la fréquence des rapports**: si vous avez choisi la mise en quarantaine ci-dessus, il est indispensable de recevoir régulièrement un rapport des messages retenus, afin de détecter les faux positifs. Cette rubrique permet de spécifier la fréquence désirée.

Les valeurs par défaut sont: *mise en quarantaine avec l'envoi d'un rapport quotidien*:

### Les faux positifs

S'il existait un filtre détectant à coup sûr les spams, ceci réglerait définitivement le problème. Dans la réalité, bien que le taux de succès soit élevé (plus de 90 %), il faut prendre les dispositions pour traiter les erreurs du filtre. Le plus délicat est le cas des **faux positifs**, soit les messages marqués à tort

comme spam alors qu'il vous sont vraiment destinés.

Il est donc important de viser tous les messages que MailCleaner marque comme spam pour détecter ces faux positifs, en se basant sur l'expéditeur et le sujet.

Si vous avez choisi ci-dessus l'option *marquage du sujet*, il s'agit juste de vérifier les messages ayant la chaîne **{Spam?}** dans le **Subject** et de prendre le cas échéant les dispositions nécessaires, comme par exemple de reclasser le message dans le bon *folder*.

Dans le cas de l'option *mise en quarantaine* les messages détectés comme spam sont retenus et il existe deux possibilités pour traiter les faux positifs:

- **Consulter la quarantaine**: aller sur la page d'accueil de MailCleaner et choisir l'onglet **quarantaine**. Si d'après l'expéditeur et le sujet vous pensez que tel ou tel message pourrait être un faux positif, vous pouvez le libérer (c'est-à-dire qu'il est envoyé à votre boîte aux lettres) en cliquant sur l'icône ➜ correspondante.
- **Consulter le rapport**: comme indiqué ci-dessus, vous pouvez recevoir à intervalle régulier par e-mail un rapport des messages retenus se présentant sous cette forme:

```
-------------------------------------------
Identificateur: 19Vo1l-0003ur-00
        De: "John Spammer" <JohnS@bidon.com>
        Date: 27-6-2003 09:52:50
        Sujet: Make money fast selling guinea
                        pigs on the Internet !!!
        Libérer: http://mailcleaner.epfl.ch/
                        fm.php?id=19Wo1l-000rdr-01

-------------------------------------------
Identificateur: 19Vo34-0003wR-00
        De: Chef Sympa <chef.sympa@epfl.ch>
        Date: 27-6-2003 09:53:50
        Sujet: Votre augmentation de salaire
        Libérer: http://mailcleaner.epfl.ch/
                        fm.php?id=19Wo34-0003Rw-00
```

Si d'après l'expéditeur et le sujet vous pensez que tel ou tel message pourrait être un faux positif (par exemple le deuxième message ci-dessus :-), vous pouvez le libérer en visitant l'URL indiqué dans le champ **Libérer** correspondant. Il parviendra quelques instants plus tard dans votre boîte aux lettres.

### Autres fonctionnalités de la quarantaine

- L'icône ➪✉ en haut de page permet d'obtenir à tout moment l'envoi d'un rapport sur une période désirée .
- L'icône ⓘ permet d'obtenir une explication succinte des raisons qui ont conduit MailCleaner à marquer comme spam le message correspondant.
- L'icône ➖○ permet d'envoyer automatiquement le message correspondant à l'équipe MailCleaner pour qu'elle puisse analyser pourquoi le filtre s'est trompé et ainsi de pouvoir l'améliorer (ceux qui ont choisi le marquage **{Spam?}** peuvent le faire en l'envoyant à ***error@mailcleaner.net*** de la manière documentée ci-dessous pour les faux négatifs).
- L'icône 🗑 permet de faire de l'ordre dans la quarantaine en effaçant **définitivement** tous les messages affichés sur la page. **Il ne faut donc le faire que si on est sûr et certain qu'aucun *faux positif* n'y subsiste !**

### Attention ! MailCleaner utilise votre adresse physique

Voici une subtilité à laquelle il faudra prendre garde: MailCleaner indexe les quarantaines des utilisateurs non pas selon leur adresse *logique* ***prenom.nom@epfl.ch*** mais selon leur adresse *physique* ***user@machine.epfl.ch*** (p.ex. ***user@mailbox.epfl.ch***). Si vous avez l'intention de changer d'adresse physique (votre boîte aux lettres passe sur un autre serveur POP/IMAP), cela signifie qu'il faut faire de l'ordre dans sa quarantaine et vérifier qu'elle ne contient plus de faux positifs avant de le changement.

### Les faux négatifs

C'est le cas où un spam déjoue les filtres de MailCleaner et parvient dans votre boîte aux lettres sans être marqué. Vous pouvez bien sûr simplement l'effacer, comme vous le faisiez auparavant, mais vous voudrez peut-être avertir MailCleaner pour qu'ils analysent les raisons de cet échec et améliorent leur filtre.

Pour cela il faut transmettre le message à l'adresse ***analyse@mailcleaner.net***, mais en prenant les précautions nécessaires pour que le message parvienne tel quel, avec tous les en-têtes, etc., car un simple copier/coller ou un *forward* risque de reformatter le message en le rendant inutilisable par l'équipe d'analyse de MailCleaner. Il faut donc transmettre le message en *attachment MIME* de type **message/rfc822** ou par un *resend*:

#### Sous Netscape,

Sélectionner le message et choisir l'entrée **Attachment** dans le sous-menu **Forward as...** du menu **Message**.

#### Sous Outlook Express,

Sélectionner le message et choisir l'entrée **Forward as Attachment** dans le menu **Message**.

#### Sous Outlook,

Ouvrir le message puis dans le menu **Actions** choisir l'entrée **Resend This Message**, acquiescer à l'avertissement apparaissant ensuite (comme quoi vous ne serez pas l'expéditeur apparent du message que vous allez envoyer) et modifier le champ de destinataire **To:**à ***analyse@mailcleaner.net***.

#### Sous Entourage,

Sélectionner le message et choisir l'entrée **Redirect** dans le menu **Message**.

#### Depuis le Web

Sous l'une des interfaces *Web* à votre boîte sur le serveur **mailbox.epfl.ch** (***http://mailbox.epfl.ch***), ouvrir le message et choisir en haut de page:

| | |
|---|---|
| pour IMP 2.2 : | **Bounce** |
| pour IMP 3.2: | **Redirect** |
| pour iPlanet WebMail: | **Forward** |

Malheureusement, Eudora ne permet pas de transmettre les messages de cette manière. ■