I INTRODUCTION

Dans ce nouveau chapitre nous exposons les différents types attaques utilisées en phishing, ainsi que les contre-mesures utilisées pour faire face à ces attaques qui ne cessent d'augmenter de jour en jour.

Ce qu'il faut retenir du chapitre précédent

« L'hameçonnage consiste en un envoi massif de courriels contrefaits, communément appelés courriels hameçon, utilisant l'identité d'une institution financière ou d'un site commercial connu de façon apparemment authentique.

Dans ces courriels contrefaits, on demande aux destinataires de mettre à jour leurs coordonnées bancaires ou personnelles en cliquant sur un lien menant vers un site Web illégitime qui est habituellement une copie conforme du site de l'institution ou de l'entreprise. Le pirate, ou criminel informatique qui a envoyé le courriel hameçon peut alors récupérer ces renseignements afin de les utiliser à son avantage.

Certaines techniques d'hameçonnage plus récentes, consistent en des logiciels malveillants, ou malicieux, qui sont développés dans le but de nuire à des systèmes informatiques.

Dans le cas d'une attaque d'hameçonnage, les logiciels malveillants peuvent varier d'un capteur de touches aux corrupteurs de données stockées.

Tandis qu'un capteur de touches est utilisé pour intercepter les touches frappées à l'aide du clavier et ainsi capter des mots de passe et autres informations personnelles, les corrupteurs de données de leur côté servent plutôt à corrompre les données des infrastructures de navigation afin de réorienter de façon automatique les utilisateurs vers des sites Web frauduleux à l'aide d'aiguilleurs (proxy) contrôlés par des pirates informatiques. »

II LES TYPES DES ATTAQUES [1]

1 EMPOISONNEMENT DU CACHE DNS (Cache poisoning)

Lorsqu'un serveur DNS est obligé d'interroger un autre serveur DNS pour obtenir l'adresse IP d'un nom de domaine faisant l'objet d'une requête, ce qui est le cas le plus général, il stocke temporairement (2 jours en moyenne) le résultat dans sa mémoire cache.

Ceci lui permet de pouvoir fournir immédiatement ce numéro en cas de nouvelle requête. Puisque ce cache est conçu pour recevoir des informations en provenance de l'extérieur, on conçoit que, s'il existe une faille de sécurité sur ce serveur, il soit possible à un pirate d'y

insérer un nom de domaine connu (par exemple www.google.fr) et lui faire correspondre le numéro IP d'un autre site (site piégé envoyant des programmes malveillants).
Un visiteur utilisant ce serveur DNS sera donc redirigé vers le site piégé au lieu d'atteindre le site demandé (Google dans l'exemple choisi).
C'est l'attaque par empoisonnement du cache.

2 INJECTION DE CODE

Les attaques de type **Cross-Site Scripting** (notée parfois *XSS* ou *CSS*) sont des attaques visant les sites web affichant dynamiquement du contenu utilisateur sans effectuer de contrôle et d'encodage des informations saisies par les utilisateurs. Les attaques Cross-Site Scripting consistent ainsi à forcer un site web à afficher du code HTML ou des scripts saisis par les utilisateurs. Le code ainsi inclus (le terme « injecté » est habituellement utilisé) dans un site web vulnérable est dit « malicieux ».
Il est courant que les sites affichent des messages d'information reprenant directement un paramètre entré par l'utilisateur. L'exemple le plus classique est celui des « pages d'erreur 404 ». Certains sites web modifient le comportement du site web, afin d'afficher un message d'erreur personnalisée lorsque la page demandée par le visiteur n'existe pas. Parfois la page générée dynamiquement affiche le nom de la page demandée. Appelons *http://site.vulnerable* un site possédant une telle faille. L'appel de l'URL *http://site.vulnerable/page-inexistante* correspondant à une page n'existant pas provoquera l'affichage d'un message d'erreur indiquant que la page « page-inexistante » n'existe pas. Il est ainsi possible de faire afficher ce que l'on souhaite au site web en remplaçant « page-inexistante » par toute autre chaîne de caractère.
Ainsi, si aucun contrôle n'est effectué sur le contenu fourni par l'utilisateur, il est possible d'afficher du code HTML arbitraire sur une page web, afin d'en changer l'aspect, le contenu ou bien le comportement.
De plus, la plupart des navigateurs sont dotés de la capacité d'interpréter des scripts contenus dans les pages web, écrits dans différents langages, tel que JavaScript, VB Script, Java, ActiveX ou Flash. Les balises HTML suivantes permettent ainsi d'incorporer des scripts exécutables dans une page web : <SCRIPT>, <OBJECT>, <APPLET>, and <EMBED>.
Il est ainsi possible à un pirate d'injecter du code arbitraire dans la page web, afin que celui-ci soit exécuté sur le poste de l'utilisateur dans le contexte de sécurité du site vulnérable. Pour ce

faire, il lui suffit de remplacer la valeur du texte destiné à être affiché par un script, afin que celui s'affiche dans la page web. Pour peu que le navigateur de l'utilisateur soit configuré pour exécuter de tels scripts, le code malicieux a accès à l'ensemble des données partagées par la page web de l'utilisateur et le serveur (cookies, champs de formulaires, etc.).

3 Attaque man in the middle

L'attaque « **man in the middle** » (littéralement « attaque de l'homme au milieu » ou « attaques de l'intercepteur »), parfois notée MITM, est un scénario d'attaque dans lequel un pirate écoute une communication entre deux interlocuteurs et falsifie les échanges afin de se faire passer pour l'une des parties.
La plupart des attaques de type « man in the middle » consistent à écouter le réseau à l'aide d'un outil appelé sniffer.

4 phishing avec pièces jointes

Une tendance lourde est actuellement en cours chez les phisheurs. De plus en plus d'arnaques mettent en jeu un message frauduleux accompagné d'une pièce jointe comprenant le formulaire de vol d'informations.

La SNCF a été victime récemment d'une attaque de phishing de ce type, avec un message potentiellement à même de tromper un certain nombre internautes.

Evidemment, le piège demeure plus aisément identifiable si vous n'avez pas commandé récemment de billets de train en ligne !

Des indices peuvent également permettre à un individu vigilant de déceler la tentative d'arnaque. Quelques fautes se sont en effet glissées dans les légères modifications textuelles effectuées par le pirate.

5 Phishing par virus (malwares)

Les keyloggers sont assez rare, les phisheurs installent un programme sur l'ordinateur de la victime qui enregistre tous ce qui est tapé au clavier y compris les mots de passes.
Pour lutter contre ce type de piratafe il faut avoir un antivirus de qualité, et surtout un PC mis à jours!

6 Phishing par fenêtres pop-up

Durant l'utilisation d'un site de confiance sécurisé, une fenêtre pop-up s'affiche invitant l'internaute à réinscrire son identifiant et son mot de passe. Une fois les informations validées, l'instigateur de l'attaque peut les réutiliser.

Ce type d'attaque utilise, généralement, un script javascript. Il est techniquement possible pour un script Javascript de déclencher une action si le site prédéterminé est visité en même temps que le site contenant le script. Si c'est le cas, le script javascript se déclenche et ouvre la pop-up. Cette attaque est notamment basée sur la faille de Cross Site Scripting.

1.7 Utilisation de bar d'adresses

De nombreux sites web d'hameçonnage désactivent la "barre d'adresse" du navigateur, ce qui signifie que vous ne pouvez pas voir l'adresse du site web que vous visitez. Ceci est délibéré, afin que vous ne remarquiez pas que le site que vous consultez est contrefait et n'a pas la bonne adresse.

III LES SOLUTIONS

1 Contrôle d'intégrité

Lorsqu'un serveur a été compromis, le pirate masque généralement son passage en supprimant les traces dans les journaux d'activités. Par ailleurs, il installe un certain nombre d'outils lui permettant de créer une porte dérobée, afin d'être à même de pouvoir revenir ultérieurement.
Le pirate pense généralement à corriger la vulnérabilité lui ayant permis de s'introduire afin d'éviter que d'autres pirates s'infiltrent.
Sa présence sur un serveur peut néanmoins être trahie par un certains nombre de commandes d'administration permettant d'afficher la liste des processus en cours ou bien tout simplement les utilisateurs connectés à la machine. Il existe ainsi des logiciels, appelés rootkits, chargés d'écraser la plupart des outils du système et de les remplacer par des commandes équivalentes masquant la présence du pirate.
Il est donc aisé de comprendre qu'en l'absence de détérioration il peut être très difficile pour un administrateur de s'apercevoir qu'une machine a été compromise. Une des premières actions lors de la découverte d'une compromission consiste à dater la compromission afin d'évaluer l'étendue potentielle sur les autres serveurs.
En effet, d'une manière générale les serveurs stockent dans des fichiers une trace de leur activité et en particulier des erreurs rencontrées.
Or, lors d'une attaque informatique il est rare que le pirate parvienne à compromettre un système du premier coup. Il agit la plupart du temps par tâtonnement, en essayant différentes requêtes.
Ainsi la surveillance des journaux permet de détecter une activité suspecte. Il est en particulier important de surveiller les journaux d'activité des dispositifs de protection car tout aussi bien configuré qu'il soit, il se peut qu'ils soient un jour la cible d'une attaque.

2 Analyse de la présence de rootkits

Il existe certains logiciels (chkrootkit par exemple) permettant de vérifier la présence de rootkits sur le système. Néanmoins, afin de pouvoir utiliser ce type d'outils, il est essentiel d'être certain de l'intégrité de l'outil et de l'affichage qu'il délivre. Or, un système compromis ne peut pas être considéré comme fiable.

3 Analyse d'intégrité

Afin de s'assurer de l'intégrité d'un système, il est donc nécessaire de détecter les compromissions en amont. C'est ainsi l'objectif poursuivi par les contrôleurs d'intégrité tel que Tripwire.

Le logiciel Tripwire, développé à l'origine par Eugène Spafford et Gene Kim en 1992, permet d'assurer l'intégrité des systèmes en surveillant de façon permanente les modifications apportées à certains fichiers ou répertoires. Tripwire effectue en effet un contrôle d'intégrité et maintient à jour une base de signature. A intervalles réguliers il inspecte notamment les caractéristiques suivantes des fichiers afin d'identifier les modifications et les éventuelles compromissions :

- permissions
- date de dernière modification
- date d'accès
- taille du fichier
- signature du fichier

Les alertes sont envoyées par courrier électronique, de préférence sur un serveur distant, afin d'éviter tout effacement de la part du pirate.

* Limites du contrôle d'intégrité

Afin de pouvoir s'appuyer sur les résultats d'un contrôleur d'intégrité il est essentiel d'être sûr de l'intégrité de la machine lors de l'installation. Il est également très difficile de configurer ce type de logiciel tant le nombre potentiel de fichiers à surveiller peut être important. De plus, lors de l'installation de nouvelles applications il est indispensable de mettre leurs fichiers de configuration sous contrôle.

Par ailleurs, ce type de solution est susceptible d'envoyer un grand nombre de fausses alertes, notamment lorsque le système modifie seul des fichiers de configuration ou lors de mises à jour du système.

Enfin, si la machine est effectivement compromise, il est possible que le pirate tentera de compromettre le contrôleur d'intégrité avant la prochaine mise à jour, d'où l'importance de stocker les alertes sur une machine distante ou bien un support externe non réinscriptible.

4 Les liste blanche

3 Listes noires

Jusqu'à aujourd'hui, la « liste noire » est sous-doute la technique de filtrage la plus commune (que la solution soit positionnée ou non sur le poste client). La logique de cette technique consiste à « marquer » certaines certains domaines dont il est risqué d'accéder.

Le problème posé par cette approche est que les utilisateurs doivent maintenir manuellement et/ou mettre régulièrement à jour leur liste noire auprès d'une base de données centralisée afin de toujours posséder la dernière version de la liste des « sites illégitimes ». De plus, cette technique peine en réalité à bloquer toutes les adresses, parce que les phisheurs peuvent trop facilement ouvrir d'autres sites, et rendent ainsi la liste noire inutile.

Il existe aussi des listes noires, mais basée sur le contenu des messages : l'exercice consiste à classifier certains mots-clés comme étant illicites et bloquer les emails qui contiennent ces mots.

En quelques mots, nous pouvons définir les listes "blanches" et "noires" de la manière suivante : les expéditeurs en liste noire sont bloqués et les expéditeurs en liste blanche sont les bienvenus.

Les listes noires sont les listes ayant identifié des spams collectifs et sont listés afin de ne pas les délivrer.

Sur le même principe que les listes blanches, il y a des listes noires "locales", et des listes noires générales, communément appelées les RBL(RBL ou Realtime Blackhole)

La RBL est une liste noire de machines ou de domaines bannis, mise à jour en temps réel.

III.1 LE FILTRAGE

3. LES TECHNIQUES TRADITIONNELLES DE FILTRAGE

Avant de parler technologie, voyons en premier lieu les différents types d'architecture pour un système utilisant une messagerie protégée par un antiphishing.

3.1 Le positionnement du filtre : deux types d'approche

On observe classiquement deux types de scénarios :

1) Le filtrage antispam « à la demande ». Le filtre antispam est positionné sur un serveur proxy.

Le filtrage est alors automatique et transparent. Exemple : le produit SpamWeed.

2) Le filtrage « vérificateur ». La solution antispam s'interpose entre l'utilisateur et le serveur POP, qu'elle scrute périodiquement. Exemples : les produits SpamKiller ou Mailwasher.

Les deux types de scenarii et leurs points forts/faibles sont mis en valeur dans le tableau ci-dessous :

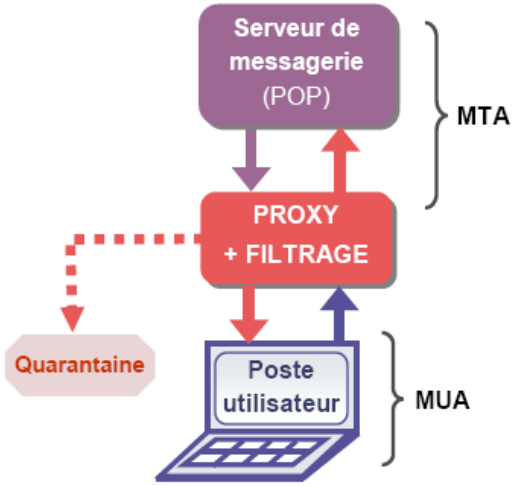
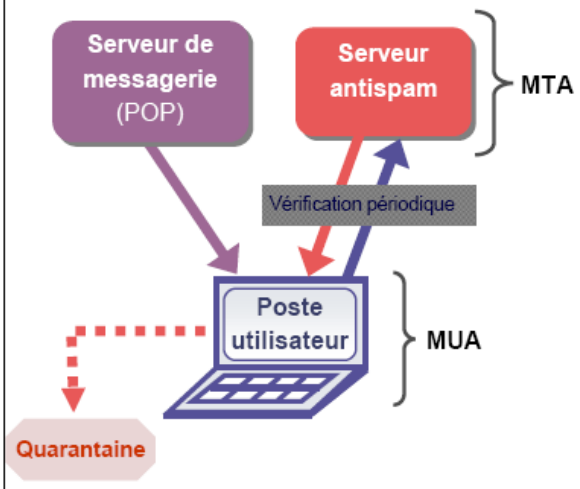
1) Filtrage « à la demande »	2) Filtrage « vérificateur »
Schéma architectural :  	*Schéma architectural* : 
Avantages : Facilité d'utilisation ; Toutes les opérations sont transparentes pour le client de messagerie.	*Avantages* : Le filtre est personnalisable et paramétrable par l'utilisateur.
Inconvénients : Le filtre n'est ni maîtrisable, ni personnalisable selon le contexte utilisateur ; ce dernier ne peut pas classifier manuellement ses mails en « spams / pas spams ».	*Inconvénients* : La partie MUA n'est pas toujours bien synchronisée avec le vérificateur anti-spam. De plus, la part de travail laissée à l'utilisateur l'utilisateur n'est plus négligeable.

Figure 2 – Les deux types d'architecture classiques pour l'intégration d'une solution antispam

3.1.1 Installation d'un logiciel antiphishing

Qui veut offrir une solution antiphishing à son serveur de messagerie peut l'installer en tant que logiciel en local se comportant comme une « passerelle de messagerie ». Pour ce faire, il faut s'assurer que ce dernier est le premier à recevoir le courrier destiné au serveur de

messagerie (courrier entrant), ainsi que le dernier pour le courrier en partance (courrier sortant). Cette installation est aussi connue sous le nom de « Smart host » (Hôte Actif). Le schéma architectural suivant montre qu'une telle solution antispam s'apparente généralement à un serveur-relais de courrier :

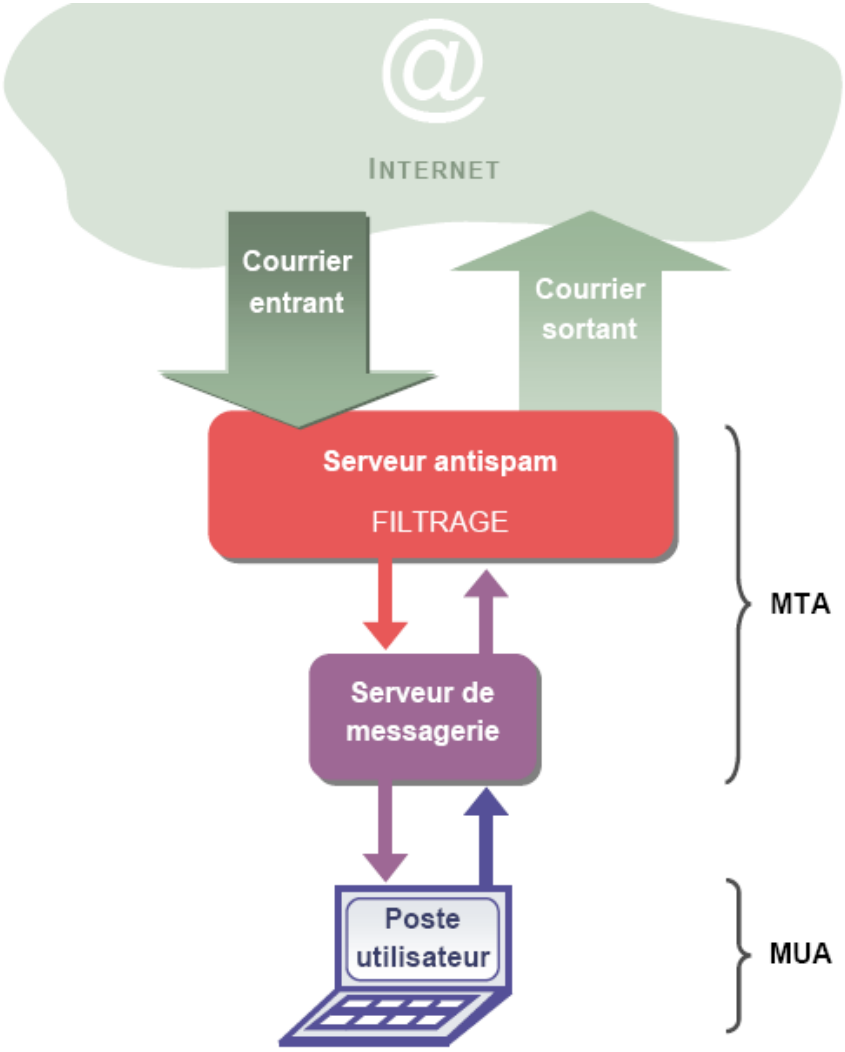


Figure 3 – Architecture d'un système de messagerie exploitant un serveur antispam en interne

Au passage, il faut noter qu'un bon nombre de logiciels antivirus s'interposent d'une façon équivalente entre le courrier entrant/sortant et le serveur de messagerie. On trouve même de nombreuses solutions intégrant à la fois la fonction d'antispam et d'antivirus. On peut citer en exemple l'outil ProtecMail

ProtecMail. Ce dernier, situé sur Internet, sait auprès de quel serveur de messagerie (niveau MTA) récolter les mails susceptibles de contenir des spams ou des virus.

L'avantage de ce type de service est double. D'une part, il permet à ses bénéficiaires d'éliminer une majorité de spam avant leur téléchargement, ce qui diminue leur temps de

connexion (ce qui n'est pas négligeable pour les utilisateurs d'un modem téléphonique 56Kb/s). D'autre part, il n'y a ni logiciel à installer, ni de mises à jour régulières à planifier : il suffit juste de paramétrer ce service une fois pour toutes. En revanche, un inconvénient des services de filtrage en ligne est de rendre difficile, voire impossible, la maîtrise du filtrage des courriels.

D'autres exemples de services de filtrages peuvent être Clinbox de Dolphian, ou Pop3Scan, ou encore Vade Retro ASP de Goto.

3.2 Les listes blanches/noires, base des procédés de filtrage traditionnels

La politique de liste blanche et de liste noire n'est pas propre au domaine de la lutte antiphishing : les serveurs proxy des entreprises en sont souvent pourvus afin de restreindre l'accès vers les sites web n'étant pas jugés en relation avec l'activité professionnelle exercée. Dans ce cas, l'expression « liste noire » désigne une liste contenant les URLs interdites. Au contraire, l'expression « liste blanche » contient les URLs autorisées. Ce concept impose souvent un compromis entre le coût d'entretien de la liste blanche par les administrateurs de l'entreprise et la perte engendrée18 par le fait que certains sites ne sont pas couverts par la liste noire. Il faut évaluer la taille d'une liste blanche, ainsi que son évolution dans le temps. Quant aux listes noires, elles sont mises à jour régulièrement, et, le plus souvent, par des sociétés spécialisées fournissant ce service.

Dans le cas de l'antisphishing, les listes blanches ou noires ne référencent plus les URLs autorisées ou interdites, mais plutôt les expéditeurs de mails (tantôt correspondants légitimes, tantôt spammeurs présumés).

3.2.2 Les listes blanches

Le filtrage de spams par liste blanche, comme par listes noires, nécessite l'usage d'un outil spécifique qui permet à la fois des mises à jour fréquentes et une personnalisation possible du contenu de ces listes.

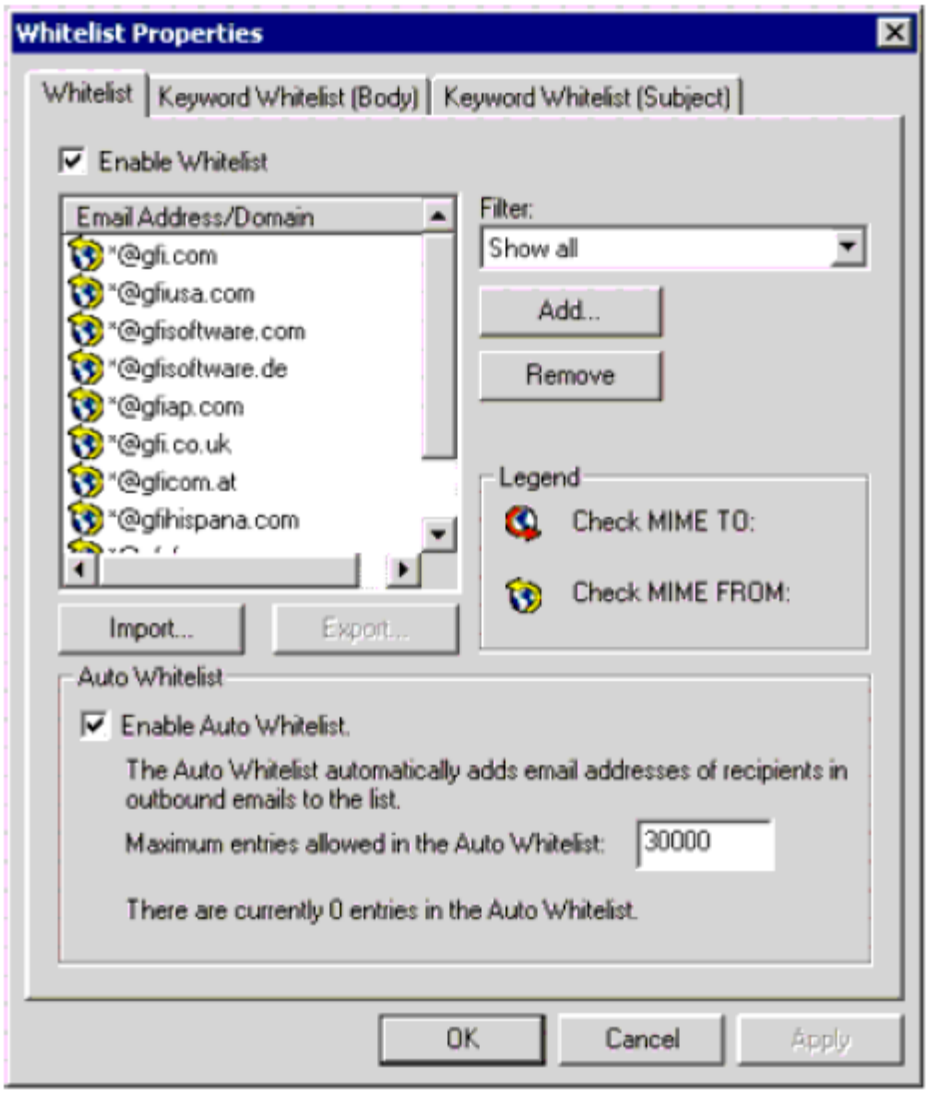


Figure 5 – Configuration d'une liste blanche dans le logiciel GFI MailEssentials

L'inconvénient des listes blanches est que le destinataire doit traiter manuellement les messages provenant d'expéditeurs inconnus.

Il n'est jamais conseillé de compter uniquement sur les listes noires ou listes blanches. En effet, ces méthodes ne se sont pas gérables pour ceux qui reçoivent une grande quantité de messages provenant de personnes inconnues. Cependant, la technologie en elle-même n'est pas complètement inutile.

Elle peut être utilisée en complément à d'autres technologies pour améliorer les taux de détection et rendre les produits anti-phishing plus simples d'utilisation et plus personnalisables.

3.3 L'analyse par mots-clés

La correspondance par mots-clés est une très ancienne méthode. Elle marque les messages en tant que spam si certains « mots suspects » sont détectés, et les marque plus favorablement en courrier valide si certains « bon mots » sont trouvés.

Cette approche implique d'analyser le corps d'un email pour des mots-clés spécifiques et des expressions. Ces mots sont en effet peu susceptibles d'apparaître dans une correspondance professionnelle classique. Mais l'analyse de mot-clé en tant que solution antiphishing autonome est une technique très primitive, car :

- Elle produit un taux élevé de faux-positifs.
- Il est possible pour les spammeurs d'abandonner certains mots-clés et d'en utiliser d'autres pour exprimer la même chose.
- Elle ne peut rien contre les mots suspects incorporés dans des images… À moins, bien sûr, de disposer d'un produit antiphishing avec reconnaissance de caractères intégrée, mais ce n'est apparemment pas proposé sur le marché à ce jour.

3.4 L'analyse lexicologique

Une problématique se pose fréquemment dans la lutte antispam : la présence d'un mot ou d'une expression suspecte par elle-même ne signifie pas nécessairement que le message en question est un spam : il faut donc à tout prix éviter qu'il soit reconnu à tort comme un spam. À la différence de l'analyse par mots-clés, l'analyse lexicologique analyse le contexte de tous les mots et les expressions dans un message particulier. À chaque mot ou expression est assigné un poids qui dépend principalement du contexte dans lequel on le trouve.

3.5 Le filtrage bayésien (filtrage- solution)

Le filtrage bayésien est une technique dite adaptative, qui reflète notre propre définition de ce qu'est et n'est pas un spam. Il se situe parmi les techniques de détection du spam les plus efficaces, ce qui le rend très populaire dans le champ de bataille de l'anti-spam. Ce que l'on appelle la « classification naïve bayésienne » ne date pas d'aujourd'hui, mais de 1763, basée sur la théorie statistique du scientifique anglais T. Bayes.

Le principe de tout filtre bayésien est d'apprendre par les exemples. En effet, lors d'une première utilisation, il peut paraître insatisfaisant. Cependant, après quelques jours d'entraînement, il devient extrêmement précis. Bien que l'entraînement puisse être légèrement gênant, comparé aux efforts requis par les autres méthodes, cela reste trivial, et les avantages pèsent bien plus lourd que le coût exigé.

L'intérêt incontestable de cette méthode est qu'elle bénéficie d'un taux élevé de détection de spam, tout en garantissant un nombre très bas de faux-positifs. Dans l'article écrit par Paul

Graham19, "Un plan pour le spam", l'auteur affirme qu'après une petite période d'apprentissage, les filtres bayésiens bloquent plus de 99,5% des spams, avec 0,03% de faux-positif. Une solution comme 'GFI MailEssentials' à base de filtre bayésien, estime son taux de reconnaissance du spam à 98% après une période de seulement deux semaines20.

3.5.1 Fonctionnement d'un filtre bayésien

Imaginons un courrier candidat au filtrage bayésien. Supposons que ce courrier contienne certains mots qui avaient déjà été repérés auparavant dans des courriers considérés comme spam.

Supposons également que ces mots ne sont encore jamais apparus dans un courrier valide (encore appelé « ham » – en raison de « Paul Graham » ?). Alors, il est légitime de considérer que ce courrier électronique est un spam.

C'est ainsi qu'un filtre bayésien accumule sans cesse les résultats de ses analyses, qui sont euxmêmes utilisés par la suite pour aider à démasquer des courriers futurs, comme le montre le schéma suivant :

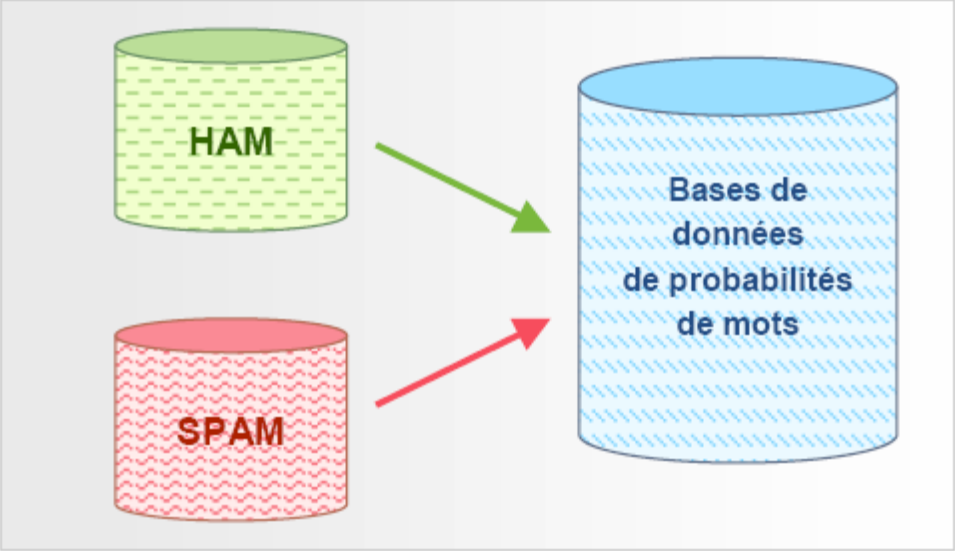


Figure 6 – Création d'une base de données de mots pour le filtre bayésien

C'est une valeur de probabilité qui est affectée à chaque mot ou unité lexicale. Celle-ci est basée sur des calculs qui tiennent compte du nombre de fois que ce mot se présente en tant que spam, par opposition au « ham » qui est le courrier valide. Cela se fait en analysant d'une part le courrier sortant des utilisateurs, et d'autre part les spams connus : tous les mots et

unités lexicales des deux regroupements de courrier sont analysés pour définir la probabilité pour qu'un courrier soit un spam.

Cette probabilité par mot est calculée de la façon suivante : si par exemple le mot « mortgage »21 apparaît dans 400 des 3.000 messages spam, et dans 5 des 300 messages légitimes, alors sa probabilité d'être un spam serait de 0,8889. (22)

Il est alors logique de concevoir qu'un filtre bayésien a besoin d'un minimum de temps pour devenir pleinement efficace. C'est ce que l'on appelle le « temps d'apprentissage » du filtre bayésien. Le produit GFI MailEssentials est un exemple de solution antispam utilisant le filtre bayésien comme technique de défense principale. En effet, la société GFI recommande de laisser au filtre le temps de s'adapter à sa messagerie pendant au moins une semaine, durant laquelle l'utilisateur est sollicité pour aider le système à classifier le courrier en « spam / pas spam ». Selon GFI, c'est seulement après cette période qu'il devient utile d'activer le filtrage.

2.5.2 L'endroit où s'opère le filtrage bayésien

Le serveur sur lequel est installé le filtre bayésien est généralement placé en amont du serveur de messagerie (MTA).

2.5.3 La mise à jour d'une base de données bayésienne

Beaucoup de logiciels à base de filtre bayésien rendent possible d'utiliser une mise à jour desdictionnaires de probabilités de mots à partir d'un serveur centralisé sur Internet, donc commun auxentreprises du monde entier qui utilise ce logiciel. Cela revient du partage d'informations.

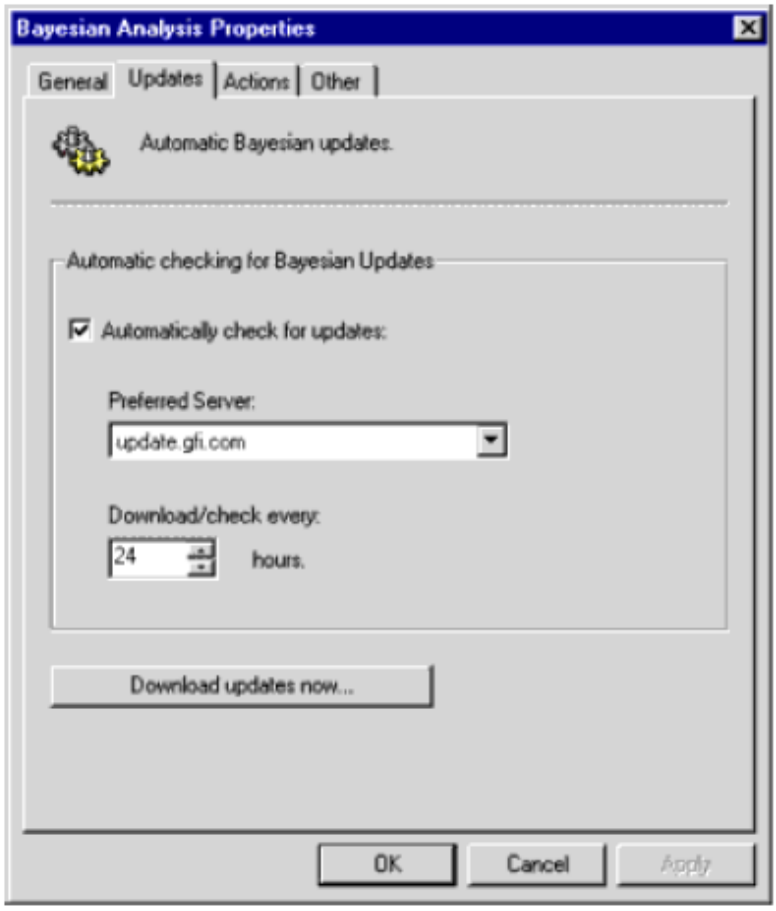


Figure 7 – Paramétrage des mises à jour d'un filtre bayésien dans GFI MailEssentials

3.5.4 Les limites du filtrage bayésien

Si les filtres bayésiens sont actuellement réputés pour leur efficacité, cela ne durera peut-être pas :

- La phase d'apprentissage d'un filtre bayésien dure souvent plus longtemps que ce que l'on lit dans les argumentaires commerciaux des différents produits : il faudrait en réalité compter plusieurs mois, et non 2 semaines comme le prétend GFI.
- Ils sont facilement coutournables par les spammeurs et ne peuvent pas analyser les spams dans toutes les langues, et particulièrement ceux à base d'images
- Ces filtres sont inefficaces contre la fraude. Par exemple, un spam connu est celui d'une fausse banque (CitiBank) nous conviant à saisir sur leur site notre numéro de carte bleue ainsi que notre code confidentiel23… De plus, ce spam est à base d'image, ce qui nous renvoie de toute façon à l'inconvénient détaillé dans le point précédent. Ceux qui sont en mode texte contiennent un vocabulaire commercial non suspect, et n'ont donc pas de raison d'être interceptés par un filtre bayésien.

- Les filtres bayésiens sont souvent utilisés avec une mise à jour régulière auprès d'une base de données centralisée commune à tous les utilisateurs d'un même produit (voir paragraphe précédent). C'est un tort, car là se situe justement l'intérêt d'un filtre bayésien : pouvoir s'adapter à un contexte unique.

Au regard de ces inconvénients, la remarque suivante de la part d'Arabella Hallawell24, aussi étonnante qu'elle puisse paraître pour les inconditionnels du filtrage bayésien, serait donc pleinementjustifiée : « Les filtres bayésiens ont montré qu'ils n'étaient pas fiables pour filtrer les spams enentreprise ».

3.6 Le contrôle d'en-tête

Dans certains logiciels antispam, un module permet de contrôler l'en-tête de chaque email pour renforcer les chances de détecter les spams. Un tel module analyse chaque champ individuel d'une en-tête, soient les champs « SMTP » et « MIME ». Les champs SMTP sont spécifiés par le serveur de messagerie, alors que les champs MIME sont spécifiés par le client de messagerie (qui crypte le mail en MIME).

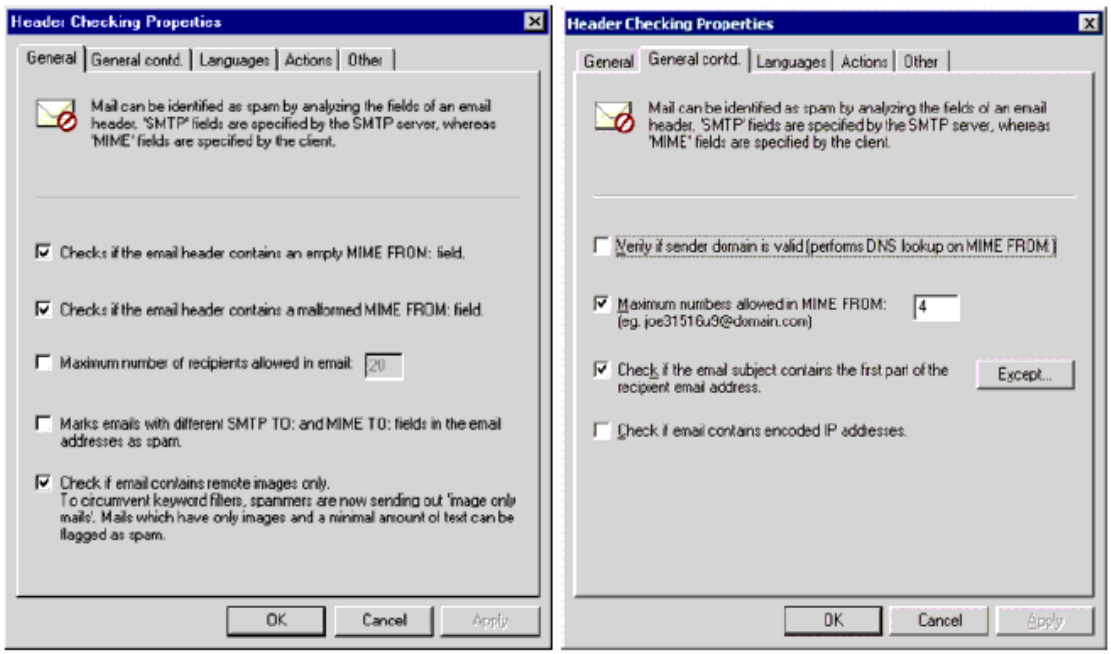


Figure 8 – Configuration du contrôle d'en-tête dans GFI MailEssentials

La copie d'écran ci-dessus montre l'onglet général de la fenêtre des propriétés du contrôle d'en-tête de la solution GFI MailEssentials. Cette fenêtre propose les options suivantes, qui sont communes à la plupart des filtres d'entêtes :

1. ’Checks if the email header contains empty MIME “From:” field'. Cette fonctionnalité vérifie si l'expéditeur s'est identifié dans le champ `FROM'. .Si ce champ est vide, alors ce mail provient très certainement d'un spammeur ;
2. ‘Checks if the email header contains a malformed MIME “From:"'. Cette fonction vérifie si l'en-tête
MIME du champ expéditeur est correcte, c'est-à-dire s'il correspond aux RFC (les spammeurs y insèrent souvent une adresse erronée).
3. 'Marks emails with recipient lists of more then X recipients as spam'. Cette fonctionnalité marque les courriers comme spam s’ils sont adressés à un grand nombre de destinataires. Cela arrive lorsque le spammeur est « juniors » ou négligent. Mais attention, il peut aussi s’agir de mails d’amis au contenu humoristique, ou encore, par exemple, de voeux de fin d’année. Il est donc essentiel de notifier dans une liste d’exceptions, si on utilise ce type d’option, les expéditeurs légitimes susceptibles d’envoyer des mails à grand nombre de destinataires.
4. ‘Marks email with different SMTP “to:” et MIME “to:” fields in the email addresses as spam’.
Vérifie si les SMTP « to: » et MIME « to: » sont les mêmes. Le serveur de messagerie d’un spammeurs doit forcément inclure une adresse SMTP « to: ». Cependant, l'adresse email MIME
to: est souvent omise ou est différente. Cette fonction bloque beaucoup de spams, mais certains serveurs de listes n'incluent pas de MIME « to: » non plus. Donc, pour utiliser cette fonction, il faut mettre sur liste blanche les adresses des expéditeurs de newsletters si elles sont marquées comme spam par cette fonction.
5. L'email contient-il principalement des images localisées à distance ? En effet, pour éviter les filtres de mots-clés tout en faisant en sorte d’alléger leurs mails, les spammeurs envoient souvent des emails contenant des images localisées sur Internet, et pas dans la source du mail.
GFI MailEssentials peut marquer comme spam les courriers présentant à la fois cette caractéristique et une quantité minimale de texte.
6. ‘Verify if sender domain is valid'. Cette fonction met le point sur une autre forme de filtrage du spam souvent utilisée : la « résolution DNS ». Une vérification DNS du domaine spécifié dans le champ MIME du destinataire est réalisée, afin de savoir si ce domaine existe bien. Si le domaine n'est pas valide, c'est un indice supplémentaire pour identifier du spam.
Remarque : Cette fonction requiert que le serveur DNS soit correctement configuré. Si le serveur

DNS n'est pas correctement configuré (ce serait souvent le cas), il y a un délai et le courrier sera traité lentement, en plus beaucoup d'emails valides seront marqués comme spams.

7. S'il y a plus de 3 chiffres dans le champ MIME « from », l'expéditeur est la plupart du temps un spammeur. La raison est que les spammeurs utilisent souvent des outils de création automatique de l'adresse de réponse (« reply to ») : adresses sur Hotmail et autres services de messagerie gratuite. Ils emploient fréquemment au moins trois caractères dans ce champ pour s'assurer que cette zone soit unique.

8. Une dernière vérification est possible : est-ce que l'objet de l'email correspond au début de l'adresse électronique du destinataire ? En effet, afin de personnaliser le spam, les spammeurs incluent fréquemment la première partie de l'adresse électronique du destinataire dans l'objet du spam. Ils utilisent cette particularité avec des adresses génériques telles sales@company.com.

La vigilance est alors nécessaire : un client qui répond à une réponse automatique avec un objet tel que « votre mail à notre service commercial » serait alors qualifié de spam. Pour éviter ceci, il est possible d'indiquer une liste d'exceptions comprenant les adresses email pour lesquelles ce contrôle ne devrait pas être réalisé.

9. Voir si le message contient des adresses IP encodées - cette vérification recherche un URL avec des encodages octaux/hexadécimaux (http://0072389472/hello.com) ou avec une combinaison de nom d'utilisateur/mot de passe (exemple www.citibank.com@scammer.com).

Ces tours sont souvent utilisés par les spammeurs ainsi que par les pirates informatique. Serait alors marqué comme spam : http://12312www.microsoft.com:hello%01@123123.

3.6.1 Détection de la langue

Il faut citer en dernier lieu, l'adaptation multi-langues d'un filtre antispam, ce qui est de plus en plus requis étant donné que le spam est de moins en moins anglophone en proportion par rapport aux quantités de spams envoyés dans le monde.

Reprenons à titre d'exemple le même logiciel (GFI MailEssentials) et, plus précisément, l'onglet 'langages' dans la boîte de dialogue de ses propriétés du contrôle d'en-tête. Celle-ci contient les options de détection de langue :

Figure

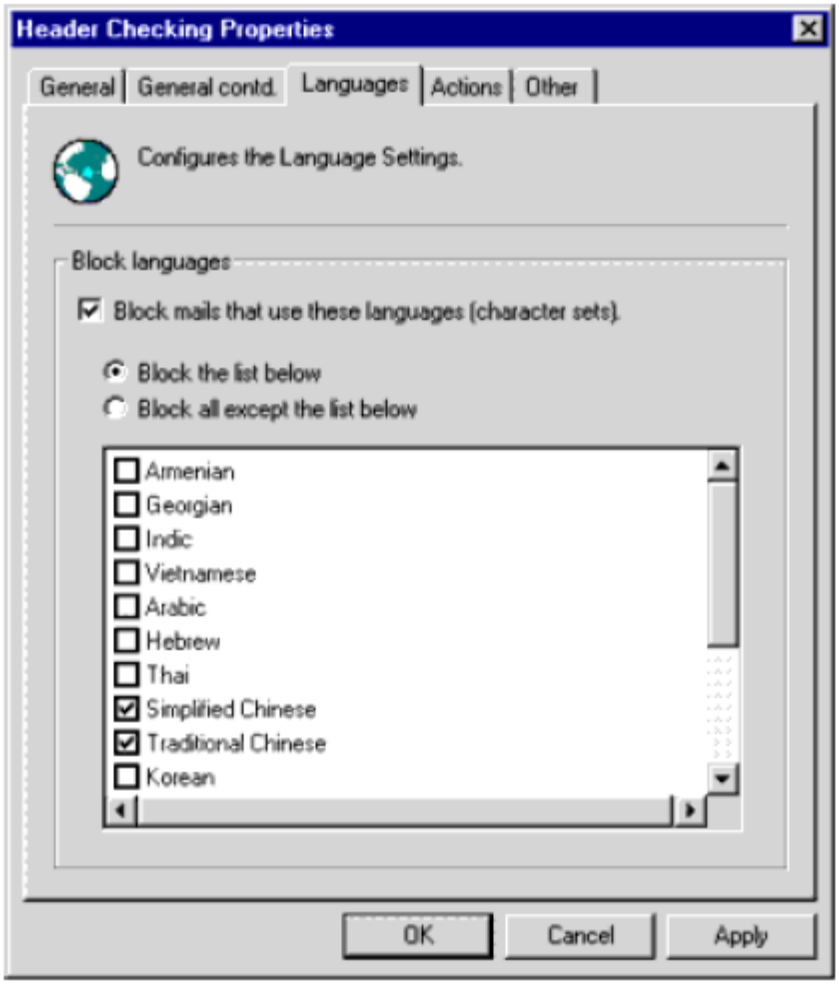


Figure 9 – Paramétrage de la langue dans GFI MailEssentials

Beaucoup de ces courriers indésirables ne sont même pas dans la langue de l'utilisateur, ce qui signifie qu'ils peuvent très simplement réduire le spam en bloquant le courrier écrit en chinois ou en vietnamien, par exemple. En fait, avec ce critère on peut bloquer le courrier selon qu'il utilise tel ou tel jeux de caractères : GFI MailEssentials, par exemple, ne peut pas distinguer l'italien du français parce que ces deux langues utilisent le même jeu de caractères… ce qui, d'un certain point de vue, constitue une limite.

4. DES APPROCHES COMPLÉMENTAIRES AUX FILTRES TRADITIONNELS

À l'image des tentatives de traque des spammeurs (qui, au demeurant, se révèlent inefficaces) il existe des moyens de lutte contre le spam complémentaires aux techniques traditionnelles de filtrage exposées précédemment dans ce rapport. Cela peut aller de simples

recommandations sur les habitudes à prendre en tant qu'utilisateur, jusqu'aux méthodes récentes (brevetées ou non) de filtrage du spam.

4.1 Confirmation de l'expéditeur (Challenge Response)

C'est une astuce devenue relativement populaire : l'idée est d'envoyer un "challenge" à l'expéditeur du message, lequel devra y répondre. Si une réponse est reçue, l'expéditeur est ajouté à la liste blanche et ne sera plus questionné. Dans le cas contraire, les messages de cet expéditeur ne seront plus affichés dans la boîte de réception, car mis en liste noire.

Cependant, la recommandation est de ne jamais utiliser cette méthode, car, quoi qu'en disent les partisans de cette technologie ou l'impression positive qu'elle peut laisser, ce n'est pas un moyen sûr. Cette technique a des conséquences fortement indésirables ; voici une liste non exhaustive des inconvénients de la technique challenge-response :

- Elle est inamicale et impolie pour les expéditeurs légitimes ;
- Les expéditeurs légitimes peuvent, à l'occasion, utiliser une autre adresse email, qui elle n'a pas été ajoutée à la liste blanche. Elle aussi devra être confirmée, ce qui est encore plus frustrant ;
- Les expéditeurs légitimes peuvent oublier de répondre ou recevoir le message seulement quelques jours plus tard, lorsqu'ils sont en déplacement ou n'ont pas accès à leurs ordinateurs. Dans ce cas, le message original (éventuellement très important) peut être significativement retardé ;
- Si l'expéditeur et le destinataire installent tous deux un logiciel basé sur le principe de confirmation, ils pourraient obtenir une boucle sans fin de confirmations, laquelle paralyserait le système ;
- Les messages provenant de services automatiques, telles les confirmations d'enregistrement ou de transaction (utilisées par des sites comme Amazon ou e-Bay), ne parviendront jamais à atteindre la boîte de réception de l'utilisateur : les messages de ce type sont envoyés par des robots… qui ne répondent jamais aux confirmations ;
- Les souscripteurs de bulletins d'informations seront bombardés de messages de confirmation
- Pour les gros FAI, les nombreux messages de confirmation doubleront leur trafic et pénaliseront significativement leurs systèmes. C'est certain que ces FAI ne seraient pas très heureux de voir leurs systèmes, déjà bien lésés, doublement alourdis à cause des spammeurs.

Steve Atkins, un consultant antispam à Redwood City, Californie, affirme ce qui suit (traduit de l'anglais) : « Cette technologie est suffisamment tentante pour que les gens l'utilisent et ne réalisent pas toutes les mauvaises choses qui commenceront à se produire ». Ainsi, si l'on est

très heureux de voir sa boîte de réception propre après avoir utilisé cette technique, il est très souhaitable d'examiner
la liste ci-dessus afin de voir si on n'est pas affecté d'une quelconque façon par un de ces inconvénients.
Réfléchir à deux fois, donc, avant d'utiliser cette trouvaille aux effets souvent indésirables.

4.2 Se désabonner ou invalider les messages (Bounce back)

Certains produits antispam offrent la possibilité de renvoyer au spammeur les messages non désirés, spécifiant que l'adresse email est fausse ou n'existe pas, et dans l'espoir d'être retiré de sa liste de diffusion. Mais cette méthode peine à stopper le spam, car d'une part les spammeurs ne vérifient pas les emails retournés, et d'autre part cela confirme que l'adresse destinataire existe, cela peut donc empirer le phénomène. De plus, dans les cas « d'usurpation d'adresse », cela risque de polluer des innocents, puisque l'adresse de l'expéditeur peut être celle d'un ami qui, lui, n'a pas envoyé ce message. Il n'est donc pas particulièrement conseillé d'utiliser cette méthode.

4.3 Réseaux anti-phishing collaboratifs

Les partisans de la technologie anti-spam collaborative affirment que les spammeurs envoient généralement le même message à des millions de personnes. Ainsi, si un utilisateur trouve un message de spam, il ou elle peut utiliser un « réseau communautaire » afin d'envoyer une "signature" du message à tous les utilisateurs ayant souscrit au même service. L'intérêt c'est que l'action d'un utilisateur empêchera les autres utilisateurs d'être dérangés par le même message.
Cette technologie fonctionne. Elle comporte cependant quelques inconvénients :
1) Le taux de détection des réseaux communautaires n'est pas toujours aussi élevé qu'il est sensé l'être ;
2) Des tests effectués par la société SpamWeed ont démontré que, d'un point de vue général, tous les produits fondés sur des "bases de données de définitions de spam mis à jour en ligne" causeront certaines gênes à l'usage. Selon eux, la vitesse de détection est lente et sujette aux erreurs réseau. Mais il faut préciser que SpamWeed est lui aussi une solution antispam De plus, l'expérience concrète des utilisateurs dépend de la conception et de la qualité de chaque produit.

4.4 La technologie ne suffit pas : précautions de base

On ne sera pas sans rappeler l'importance de « prévenir plutôt que guérir », et ce autant que possible. Dans le domaine de la lutte antispam, certaines précautions de bases devraient être en effet le réflexe de tout utilisateur de courrier électronique. Elles sont les suivantes :

- La première des recommandations aux utilisateurs est de ne pas choisir de se désinscrire d'un spam. En effet, tout lien de désinscription dans un spam est trompeur : il sert en fait à vérifier que l'adresse email est valide. Ainsi, celui qui se désinscrit recevra à coup sûr encore plus de spam.
- Éviter de communiquer son adresse e-mail sur un site dont on n'est pas sûr. Sinon, c'est prendre inévitablement le risque qu'elle finisse un jour dans les mains des spammeurs. Ainsi, si on est amené à communiquer son adresse électronique pour bénéficier d'un service, recevoir des bulletins d'information, effectuer un achat en ligne ou accéder à une partie à accès restreint d'un site Web, il faut à tout prix être prudent, et donner le moins possible son adresse.
- Utiliser une adresse jetable. Pour éviter toute pollution de son adresse principale, il est bon d'en avoir une qui soit annexe, ou temporaire (qui s'autodétruit ou que l'on supprime manuellement).

4.5 Laisser la main aux utilisateurs

Malgré une automatisation certaine du filtrage du spam rendu possible par des outils comme les filtres bayésiens, il est toutefois utile pour les administrateurs de laisser le contrôle aux utilisateurs, en leur permettant de faire savoir au système ce qu'ils considèrent, eux, comme spam. C'est en particulier lors de la phase d'apprentissage d'un filtre bayésien que cette étape est hautement recommandée. De la même façon, si un système laisse les utilisateurs ajouter eux-même leurs destinataires valides à des listes blanches, alors le risque de faux-positifs sera diminué.

4.6 SPF, un référentiel communautaire de serveurs attestés

Sous ce principe, tout serveur de messagerie est invité à se faire « certifier conforme » pour se distinguer officiellement des « robots-spammeurs ».

4.6.1 Introduction à SPF

SFP (pour Sender Policy Framework) est un effort communautaire qui gagne rapidement du terrain.

Ce référentiel requière que l'entreprise d'un expéditeur ait publié son serveur de messagerie dans un « enregistrement SPF ». Par exemple, si un email est envoyé à partir de xyz@societeABC.com, alors

societeABC.com doit publier un enregistrement SPF pour que le référentiel SPF puisse déterminer si le message était réellement envoyé à partir du réseau societeABC.com, ou s'il a été forgé. Si un enregistrement SPF n'est pas publié par societeABC.com, le résultat SPF sera "inconnu".

Les domaines enregistrés à SPF ont la garantie que chacun de leurs courriers sortants seront obligatoirement admis par les filtres antispam des autres sociétés, à condition qu'ils utilisent un filtre tirant partie de SPF.

4.6.2 Comment SPF fonctionne-t-il ?

Les domaines utilisent des enregistrements publics (DNS) pour diriger les requêtes pour différents services (web, email, etc.) vers les machines qui offrent ces services. Tout domaine publie déjà un enregistrement email (MX) pour indiquer au monde quelles machines reçoivent les emails pour ce domaine.

La copie d'écran suivante montre un exemple d'exploitation de SPF dans un logiciel (ici, GFI MailEssentials). Cet exemple montre que l'on peut définir un niveau de blocage, ainsi qu'une configuration des exceptions :

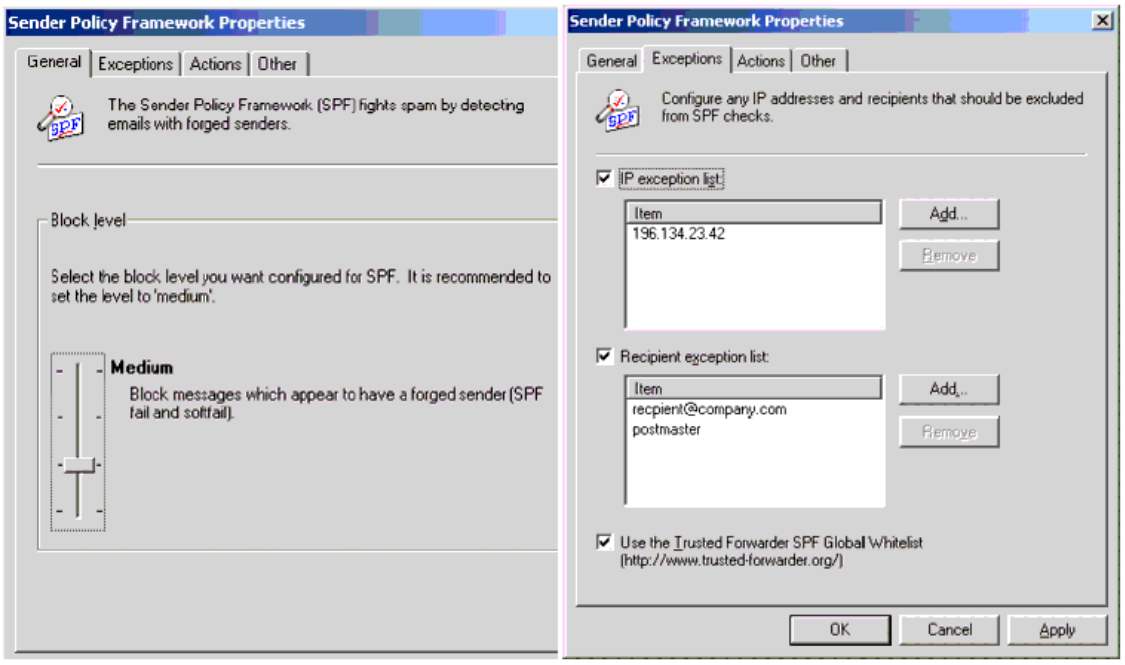


Figure 10 – Exemple de paramétrage du SPF dans GFI MailEssentials

SPF fonctionne par domaine, publiant un enregistrement texte dans le DNS de ces domaines afin d'indiquer au monde quelles machines envoient le message à partir du domaine. Quand on reçoit un message à partir d'un domaine, un logiciel exploitant SPF (comme par exemple

GFI MailEssentials) peut vérifier ces enregistrements afin de s'assurer que les messages proviennent d'expéditeurs sûrs.

Mais il n'est pas obligatoire, bien sûr, de publier un enregistrement SPF25.

4.6.3 Un exemple d'utilisation de SPF

1) Supposons qu'un spammeur forge societeABC.com et essaye de nous envoyer un spam.

2) Il se connecte à partir d'un endroit autre que SociétéABC.

3) Une fois le message envoyé, on voit « FROM: <adresse_forgee@SocieteABC.com>" », mais on n'est pas obligé de le croire. On peut demander à SociétéABC si l'adresse IP provient de leur réseau.

4) Dans cet exemple, SociétéABC publie un enregistrement SPF. Cet enregistrement indique à GFI MailEssentials comment déterminer si la machine expéditrice est autorisée à envoyer des emails au nom de SociétéABC.

5) Si SociétéABC indique qu'il reconnaît la machine expéditrice, celui-ci passe, et on peut penser que l'expéditeur est effectivement celui qu'il prétend être. Si le message ne passe pas les tests SPF, alors c'est un message forgé. C'est de cette manière que vous pouvez penser qu'il s'agit vraisemblablement d'un spammeur.

Pour plus d'informations sur SPF et comment il fonctionne, le site Internet du Sender Policy Framework est http://spf.pobox.com.

6. CONCLUSION

Ce groupe possède un système de traçage et d'enregistrement d'alerte causée par des attaques utilisant l'hameçonnage ainsi que des sites illégitimes dont les pirates se servent. Ils ont tout récemment amélioré leur système de traçage de rapport de courriels hameçons uniques en plus des sites utilisés afin de commettre des attaques d'hameçonnage. Nous entendons par rapports de courriels hameçons uniques un courriel unique qui est envoyé à plusieurs utilisateurs cibles, qui les dirigent vers un site illégitime servant à l'hameçonnage. L'APWG compte chacun de ces rapports de courriels uniques comme ceux qui dans un mois donné possèdent la même ligne de sujet dans le courriel. L'APWG répertorie et compte le nombre de sites servant à des buts d'hameçonnage.

Ceci est maintenant déterminé à l'aide des adresses uniques des sites illégitimes identifiées. Finalement, l'APWG répertorie aussi les logiciels malveillants en calculant un hachage MD5. Selon l'APWG, le nombre de sites illégitimes servant à l'hameçonnage détectés par le groupe

a augmenté à 55,643 en avril dernier, un bond majestueux par rapport au mois de mars qui affichait 35,000 sites illégitimes.

Éventuellement, les filtres anti pourriels et les autres méthodes de détection d'attaques basées sur l'hameçonnage vont s'améliorer. Le nombre et la qualité des courriels augmenteront malheureusement au même rythme que les techniques de détection.

Cependant, même avec toutes les technologies disponibles, demeurer

vigilant face aux messages que nous recevons et traitons semble être la solution la plus efficace contre l'hameçonnage présentement.

****

Comme c'est déjà le cas avec les virus informatiques, la proportion de courrier indésirable dans nos boîtes aux lettres est en progression continuelle. C'est pourquoi des solutions antiphishing sont plus que jamais nécessaires pour protéger à la fois les ressources réseaux et la tranquillité de chacun. La conception et le choix d'une solution antiphishing sont cependant plus litigieux que pour un antivirus car, à la différence des virus, tout le monde n'a pas systématiquement la même définition de ce qu'est un message indésirable.

Un même piège est récurrent, quelque soit les filtres utilisés : un produit antiphishing qui fonctionne bien peut aussi fonctionner « trop bien ». En d'autres termes, le risque de faux positifs n'étant pas négligeable pour certains produits, il est important que les utilisateurs de messageries aient accès aux messages mis en quarantaine par leur solution antispam. Et pourtant, l'idéal serait que l'utilisateur ne se rende même pas compte qu'un filtre antiphishing existe : cela lui faciliterait la vie, et éviterait aussi qu'il configure mal son filtre (car un paramétrage trop agressif augmente classiquement le nombre de faux-positifs).

Les techniques évoluent, car les procédés traditionnels de filtrage du antiphishing (comme les listes noires ou les recherches par mots-clés) s'avèrent très rigides à l'utilisation. Même les filtres bayésiens (de type statistique), les plus efficaces des procédés classiques, deviendront de plus en plus inefficaces contre les nouvelles techniques « antiphishing ». En effet, les courriers indésirables comportent de plus en plus d'images ou de vocabulaire dit « anti-heuristique » (à base de mots valides ou transformés). De plus, les messages sont de moins en

moins anglophones, ce qui révèle la faiblesse de nombreux filtres, pauvres en gestion multi-langues. De plus, les filtres bayésiens ne sont pas toujours bien utilisés, par exemple lorsque leurs bases de données de mots sont uniquement mises à jours à partir de serveurs centralisés (ce qui n'est pas approprié à leur nature, qui consiste à s'adapter).

En définitive, les tendances des procédés antiphishing sont diverses. Il faut dire que chacun y va de ses trouvailles à caractère propriétaire, sans compter que ces nouveaux procédés sont rarement autonomes, puisque souvent polyvalents : la majorité des produits antiphishing utilisent conjointement plusieurs techniques de filtrage traditionnel, le tout étant parfois associé à des méthodes brevetées inventées par l'éditeur.

La pensée commune est que, dans un système utilisant plusieurs filtres, les avantages des uns compensent les inconvénients des autres. D'une part, de grands comptes lancent depuis peu de nouvelles techniques de filtrage par authentification de l'expéditeur (selon son domaine ou encore son IP). Par exemple, le produit Symantec BrightMail 6.0 implémente l'idée de référencer les adresses IP autorisées ou interdites. D'autre part, il faut noter qu'une autre tendance, plus générale, est d'utilisation des filtres polyvalents, de plus en plus multi-langues, mariés à des techniques nouvelles. En outre, on parle de plus en plus de techniques dites « proactives », dans le sens où elles s'évertuent à stopper les messages indésirables avant qu'ils ne parviennent aux boîtes aux lettres. On peut enfin se demander si des logiciels d'OCR (reconnaissance de caractères) seront dans l'avenir intégrés à des solutions antiphishing, car les messages sont de plus en plus constitués d'images.

Autre réflexion : utiliser une base de données centralisée, partagée entre des utilisateurs du monde entier, peut représenter un avantage comme un inconvénient : tout dépend du type d'information que l'on veut centraliser. En effet, si l'on met à jour un dictionnaire pour filtre bayésien à partir d'une base unique en ligne, alors ce filtre perdra de son intérêt pour deux raisons. D'une part, les filtres bayésiens, qui sont en l'occurrence parmi les plus performants, sont faits pour s'adapter à un contexte professionnel ou amical particulier : ils doivent correspondre à un usage attentif et personnel, et pas une base de mots interdits ou autorisés qui soit universelle. D'autre part, les phishers connaissent les dictionnaires bayésiens mis à disposition de tous, et trouvent donc les manières détournées pour duper ces filtres. Par contre, l'utilisation d'informations centralisées peut représenter un réel avantage lorsque ce sont les messages eux-même, et pas leurs caractéristiques, qui sont spécifiquement référencés en temps réel dans une base de données partagée : de cette manière, il est possible de stopper très vite la propagation d'un pollupostage.