

Pourquoi le filtre Bayésien est la technologie anti-spam la plus efficace

Obtention d'un taux de détection de spam de plus de 98% grâce à une approche mathématique

Ce livre blanc décrit le fonctionnement du filtre Bayésien et explique pourquoi il constitue le meilleur moyen de combattre le spam.

Introduction

Ce livre blanc décrit comment les mathématiques bayésiennes peuvent être appliquées aux problèmes de spam, donnant lieu à une technique flexible, 'd'intelligence statistique' qui offre un très fort taux de détection de spam.

Il explique également pourquoi l'approche bayésienne est la meilleure façon de lutter contre le spam une fois pour toute, vu qu'elle surmonte des obstacles rencontrés par des technologies plus statiques telles que la vérification de liste noire, comparant aux bases de données de vérifications de spam connus et de mots clés. Ces technologies ne sont pas obsolètes mais ne peuvent pas être fiables sans un filtre Bayésien.

Introduction.....	2
Techniques actuelles de détection de spam	2
Fonctionnement du filtre Bayésien anti-spam	2
Pourquoi le filtre Bayésien est le meilleur	5
A propos de GFI MailEssentials	6
A propos de GFI Software	8

Techniques actuelles de détection de spam

Le spam est devenu un problème grandissant. Le nombre de messages spam est quotidiennement en hausse – des études montrent que plus de 50% des emails sont du spam ; le Radicati Group prévoit que ce taux atteindra les 70% d'ici 2007. En plus de ça, les spammeurs deviennent de plus en plus malins et arrivent toujours à surpasser les méthodes anti-spam statiques.

Les techniques actuellement utilisées par la plupart des logiciels anti-spam sont statiques, ce qui signifie qu'il est assez facile de passer au travers en peaufinant un peu le message. Pour y arriver, les spammeurs examinent simplement les dernières techniques anti-spam et trouvent des moyens de les contourner.

Pour combattre le spam efficacement, une nouvelle technique adaptative est nécessaire. Cette méthode doit pouvoir s'habituer et se familiariser avec les tactiques changeantes des spammeurs. Elle doit aussi être capable de s'adapter à l'organisation qu'elle protège. La solution se trouve dans les mathématiques Bayésiennes.

Fonctionnement du filtre Bayésien anti-spam

Le filtre Bayésien est fondé sur le principe que la plupart des événements sont dépendants et que la probabilité qu'un événement se répète dans le futur peut être déduite des précédents de ce même événement. (De plus amples informations sur les bases mathématiques de filtres de

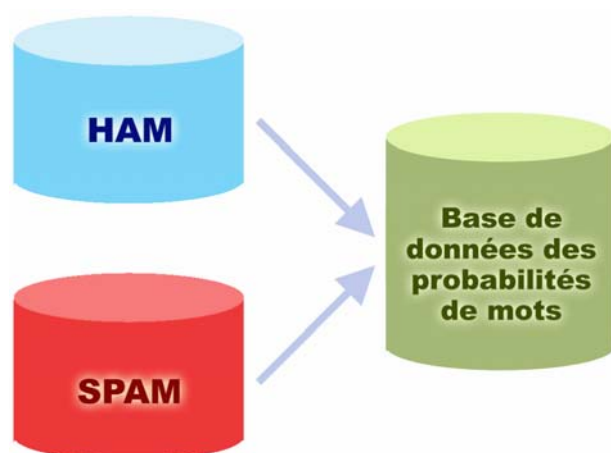
Bayes sont disponibles sur Estimation de Paramètres Bayésiens – http://www-ccrma.stanford.edu/~jos/bayes/Bayesian_Parameter_Estimation.html

et Une Introduction aux Réseaux Bayésiens et leurs Applications Contemporaines - <http://www.niedermayer.ca/papers/bayesian/bayes.html>).

Cette même technique peut être utilisée pour distinguer le spam. Si un morceau de texte apparaît souvent dans un spam et non dans un courrier légitime, il serait alors judicieux d'assumer que cet email est probablement un spam.

Création sur mesure d'une base de données Bayésienne de mots

Avant que le courrier puisse être filtré en utilisant cette méthode, l'utilisateur a besoin de générer une base de données de mots et de signes (tels que \$, adresses et domaines IP, et ainsi de suite), rassemblés à partir d'exemples de courrier spam et 'ham' (courrier valide).



Création d'une base de données de mots pour le filtre

Une valeur de probabilité est assignée à chaque mot ou signe ; la probabilité est basée sur des calculations qui prennent en compte le nombre de fois qu'un mot apparaît dans un spam contrairement à un courrier légitime (ham). Cela se fait en analysant le courrier sortant des utilisateurs et en analysant les spams connus : tous les mots et signes des deux catégories de courrier sont analysés pour générer la probabilité qu'un mot en particulier signale qu'un courrier est un spam.

Cette probabilité par mot est calculée de la façon suivante : Si le mot « mortgage » (anglais pour Hypothèque) apparaît dans 400 courriers spam sur 3000 et dans 5 courriers légitimes sur 300, par exemple, alors sa probabilité spam devrait être 0,8889 ($[400/3000]$ divisé par $[5/300 + 400/3000]$).

Création d'une base de données ham (selon les besoins de votre compagnie)

Il est important de noter que l'analyse de courrier ham est effectuée sur les courriers de la compagnie, et est donc taillée pour les besoins de celle-ci. Par exemple, une institution financière qui emploie l'expression « mortgage » très souvent recevra beaucoup de faux positifs si elle utilise un jeu de règles anti-spam général. Par contre, le filtre Bayésien, si adapté à votre compagnie après une période d'apprentissage initiale, prend en compte le courrier légitime sortant (et reconnaîtra le terme « mortgage » comme étant fréquemment utilisé dans les messages valides), et a donc un meilleur taux de détection spam et un taux de faux positifs considérablement plus faible.

Notez que certains logiciels anti-spam aux capacités Bayésiennes basiques, tel que le filtre spam d'Outlook ou l'Internet Message Filter du serveur Exchange, ne créent pas de fichiers de base ham sur mesure pour votre compagnie, mais ajoute un fichier de données ham standard lors de son installation. Bien que cette méthode ne requière pas de période initiale d'apprentissage, elle a 2 inconvénients majeurs :

1. Le fichier de données ham est disponible à tous et donc peut être trafiqué par des spammeurs professionnels et donc contourné. Si le fichier de données ham est spécifique à votre compagnie, alors l'altération des fichiers de données ham est inutile. Par exemple, il y a des hacks disponibles pour contourner les filtres anti-spam d'Outlook 2003 ou du serveur Exchange de Microsoft.
2. De plus, le fichier de données ham est général, et non adapté à votre compagnie, il ne peut pas être aussi efficace et vous allez faire face à beaucoup plus de faux positifs.

Création d'une base de données spam

Mis à part le courrier ham, le filtre Bayésien dépend aussi d'un fichier de données anti-spam. Ce fichier doit inclure un large spécimen de spams connus et doit être constamment mis à jour sur les derniers spams grâce à un logiciel anti-spam. Cela garantit que le filtre Bayésien connaît les derniers tours des spammeurs, et en conséquent offre un fort taux de détection anti-spam (remarque : cela se produit une fois la période initiale d'apprentissage de deux semaines finie).

Comment se déroule le filtrage même

Une fois les bases de données ham et spam établies, les probabilités des mots peuvent être calculées et le filtre est prêt à l'emploi.

A l'arrivée d'un nouveau message, celui-ci est décomposé en mots et les mots les plus importants – par ex., ceux qui sont les plus significatifs lors de l'identification de la nature d'un message, spam ou non sont isolés. A partir de ces mots, le filtre Bayésien calcule la probabilité qu'un nouveau message soit un spam ou non. Si la probabilité est supérieure au palier

maximum, disons 0,9, alors le message est considéré comme spam.

Cette approche Bayésienne au spam est très efficace – un article de la BBC paru en Mai 2003 a rapporté que des taux de détection de spam de plus de 99,7% peuvent être atteints avec un faible nombre de faux positifs !

Pourquoi le filtre Bayésien est le meilleur

1. La méthode Bayésienne prend l'intégralité du message en compte. Elle reconnaît les mots clés qui identifient le spam, mais aussi elle reconnaît des mots qui dénotent un message valide. Par exemple : tous les emails qui contiennent les mots « free » et « cash » ne sont pas des spam. L'avantage de la méthode Bayésienne est qu'elle considère les mots les plus intéressants (comme définis par leur dérivé) et propose une probabilité qu'il s'agisse d'un spam. La méthode Bayésienne trouvera les mots « free » et « cash » intéressants mais elle reconnaîtra aussi le nom du contact qui a envoyé le message et de là considèrera le message comme légitime par exemple ; cela permet de créer une « balance ». En d'autres termes, le filtre Bayésien est une approche bien plus intelligente car elle examine tous les aspects d'un message, au contraire de la vérification de mots clés qui classe un message dans la catégorie spam en se basant sur un simple mot.
2. Un filtre Bayésien s'adapte constamment - En apprenant à partir des nouveaux spam et des nouveaux emails valides sortants, le filtre Bayésien évolue et s'adapte aux nouvelles techniques spam. Par exemple, lorsque des spammeurs ont commencé à utiliser « f-r-e-e » au lieu de « free », ils sont arrivés à évaser la vérification de mot clé jusqu'à ce que « f-r-e-e » soit aussi incorporé dans la base de données. D'un autre côté, le filtre Bayésien remarque automatiquement de telles tactiques. En fait, si le mot « f-r-e-e » est détecté, il constitue un indicateur de spam encore plus évident, vu qu'il est peu probable qu'il apparaisse dans un message ham. Un autre exemple serait d'utiliser le mot "5ex" au lieu de « Sex ». Vous n'aurez probablement pas un mot tel que 5ex dans un courrier ham, d'où plus de chance qu'il s'agisse d'un spam.
3. La technique Bayésienne est sensible à l'utilisateur – elle apprend les habitudes de messagerie de la compagnie et comprend que, par exemple, le mot « mortgage » peut indiquer un spam si la compagnie qui exécute les filtres est, disons, un revendeur de voitures, alors qu'il ne le classerait pas comme spam si la compagnie est une institution financière s'occupant d'hypothèques.
4. La méthode Bayésienne est multi linguale et internationale. Un filtre Bayésien anti-spam, étant adaptable, peut être utilisé pour n'importe quel langage requis. La plupart des listes ne sont disponibles seulement qu'en Anglais et sont donc inutiles aux régions ne parlant pas la langue. Le filtre Bayésien prend aussi en compte certains dérivés linguistiques ou les divers usages de certains mots dans différents domaines, même si le même langage

est parlé. Cette intelligence permet à un tel filtre d'attraper plus de spam.

5. Un filtre Bayésien est difficile à tromper, à l'opposé d'un filtre à mots clés. Un spammeur avancé qui voudrait tromper un filtre Bayésien doit soit utiliser moins de mots qui normalement indiquent un spam (tels que free, Viagra, etc.), ou plus de mots qui indiquent généralement un courrier valide (tel que le nom d'un contact, etc.). Le dernier est impossible car le spammeur doit connaître le profile email de chaque destinataire et un spammeur ne pourra jamais amasser ce genre d'informations pour chaque boîte visée. L'utilisation de mots neutres, par exemple le mot "public", ne marchera jamais du fait qu'ils sont ignorés lors de l'analyse finale. Décomposer des mots associés au spam tels que « m-o-r-t-g-a-g-e » au lieu de « mortgage », ne fera qu'augmenter les chances de découvrir le spam, vu qu'un utilisateur n'écrira jamais le mot « mortgage » « m-o-r-t-g-a-g-e ».

Filtres Bayésiens ou mise à jour des listes de mots clés ?

Certains types de logiciel anti-spam téléchargent régulièrement des nouveaux fichiers de mots clés. Bien que ceci est, évidemment, préférable à ne pas mettre à jour la liste de mots clés, cette approche est plutôt inconsistante et est facilement contournée. Télécharger des mises à jour rend les choses un peu plus difficiles, mais le système principal est endommagé par rapport à un filtre Bayésien.

Où est le piège ?

Le filtre Bayésien, si implanté correctement et adapté aux besoins de la compagnie est de loin la technologie la plus efficace contre le spam. Y a-t-il un piège ? Et bien, d'une certaine façon oui, il y a un piège, mais il peut être facilement surmonté : Avant d'utiliser et de juger le filtre de Bayes, vous devez attendre qu'il fasse sa formation pendant deux semaines- ça ou vous pouvez créer une base de données spam et ham vous-même. Cette tâche peut être très complexe, il est donc recommandé de laisser le filtre faire son éducation tout seul. Avec le temps, le filtre de Bayes devient de plus en plus efficace alors qu'il en apprend plus sur les habitudes email de votre organisation. Comme on dit, il faut savoir être patient.

Il est important donc de garder cela à l'esprit lors de l'évaluation du logiciel anti-spam. Si le produit a avancé, personnalisé l'analyse Bayésienne, ensuite il pourra être évalué après quelques semaines. Il est possible que le logiciel anti-spam de base soit initialement plus performant, mais après quelques semaines, le filtre de Bayes le rattrape et le dépasse une fois pour toutes.

A propos de GFI MailEssentials

GFI MailEssentials for Exchange/SMTP offre une protection contre le spam au niveau du serveur et élimine le besoin d'installer et de mettre à jour les logiciels anti-spam sur chaque poste de travail. GFI MailEssentials offre une installation rapide et un fort taux de détection du spam, en utilisant une analyse bayésienne et d'autres méthodes – Aucune configuration n'est

requis, très peu de faux positifs au travers de la liste blanche automatique, et la possibilité de s'adapter à l'environnement de messagerie de vos utilisateurs afin de constamment régler et améliorer la détection de spam. Il vous permet aussi d'envoyer les messages spam vers les fichiers 'junk mail' (courrier indésirable) des utilisateurs. GFI MailEssentials ajoute aussi des utilitaires de messagerie clés à votre serveur de messagerie : disclaimers, rapport d'utilisation des emails, archivage et vérification des emails, réponses automatiques basées sur le serveur, et téléchargement POP3. Plus d'informations, et une version d'évaluation gratuite sont disponibles sur <http://www.gfsfrance.com/fr/mes/>.

A propos de GFI Software

GFI est l'un des leaders dans le domaine de la réalisation de logiciels qui fournit une seule source intégrée permettant aux administrateurs de résoudre les problèmes de sécurisation de réseaux, de contenu et de messageries. Grâce à sa technologie innovatrice, une stratégie agressive de commercialisation et sa concentration sur le marché de petites et moyennes entreprises, GFI répond aux besoins de continuité d'affaires et de productivité des entreprises et d'autres organisations sur une grande échelle. Fondée en 1992, GFI est une entreprise internationale qui possède des bureaux à Malte, à Londres, Raleigh, Hong Kong, Adelaïde et à Hambourg avec plus de 200.000 installations de ses logiciels à travers le monde. GFI est une entreprise spécialisée et possède un réseau de plus de 10.000 partenaires à travers le monde. Partenaire stratégique de Microsoft, GFI est membre certifié du partenariat Microsoft Gold Certified Partner. Pour plus d'informations à propos de GFI, visitez le site <http://www.gfsfrance.com>.

© 2007 GFI Software. Tous droits réservés. L'information contenue dans ce document représente le point de vue actuel de GFI sur les questions abordées à la date de sa publication. Etant donné que GFI doit répondre aux conditions dynamiques du marché, il ne devrait pas être interprété comme un engagement de la part de GFI, et GFI ne peut pas garantir l'exactitude d'aucune information présentée après la date de la publication. Ce livre blanc est seulement à titre informationnel. GFI NE FAIT AUCUNE GARANTIE, EXPRESSE OU IMPLICITE, DANS CE DOCUMENT. GFI, GFI EndPointSecurity, GFI EventsManager, GFI FAXmaker, GFI MailEssentials, GFI MailSecurity, GFI MailArchiver, GFI LANguard, GFI Network Server Monitor, GFI WebMonitor et leurs logos de produit sont des marques déposées ou des brevets commerciaux de GFI Software aux Etats-Unis et/ou dans d'autres pays. Tous les noms de produit ou d'entreprises mentionnés ci-dessus peuvent être les marques déposées de leurs propriétaires respectifs.