Le phénomène du SPAM en 2003 !!!

Jacques.Eudes@ujf-grenoble.fr
CRIP / Grenoble 1

- Définition
- Législation
- Pourquoi combattre le SPAM ?
- Les différents types de SPAM
- Impact économique
- Le combat contre le SPAM
  - Evolution des méthodes
  - Eviter la collecte de nos adresses email.
  - Implémentations à l'UJF.

Qu'est ce qu'un SPAM ?

- Le "spamming" ou "spam" est **l'envoi massif**, et parfois répété, **de courriers électroniques non sollicités**, à des personnes avec lesquelles l'expéditeur n'a jamais eu de contact et dont il a capté l'adresse électronique de façon irrégulière.
- Constituent des spams **les messages adressés sur la base d'une collecte irrégulière de méls**, soit au moyen de moteurs de recherche dans les espaces publics de l'internet (sites web, forums de discussion, listes de diffusion, chat...), soit que les adresses aient été cédées sans que les personnes en aient été informées et sans qu'elles aient été mises en mesure de s'y opposer ou d'y consentir.

Le Spam et la législation

- Législation française (LEN : en cours)
  - SPAM interdit
  - Sauf vers les entreprises !
- Législation européenne : 12 juillet 2002
- Collecte d'adresses
  - réprimé article 25 de la loi du 6 Janvier 1978
- Utilisation de serveurs relais involontaires
  - Loi Godfrain 323-1 & 2

La loi

- Le mail est une correspondance privée
  - Son détournement est illégal
  - Son détournement par un fonctionnaire est considéré comme plus grave
    - affaire de l'ESPCI : art 432-9
- 2 conseils
  - Prudence
  - Prudence
- Le courrier électronique (Que sais-je ?)
  - Lucien Rapp

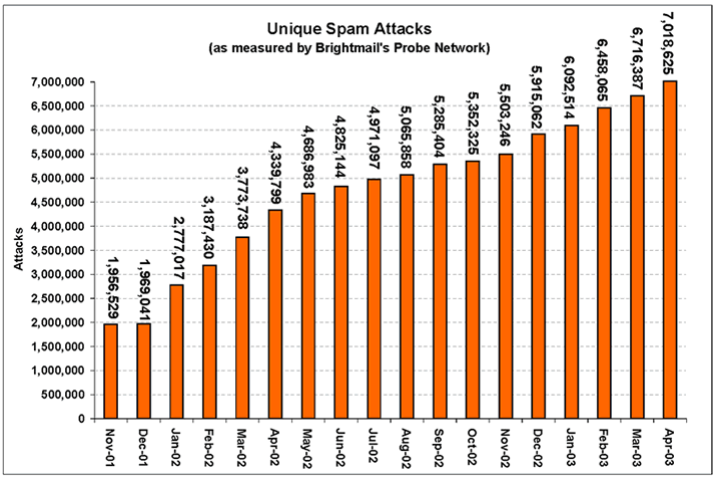
Définitions …

- HAM : courrier « correct »
- Faux négatif : SPAM non détecté
- Faux positif : HAM détecté comme SPAM
- Opt-in : Abonnement (pub)
- Opt-out : Désabonnement
- Double Opt-in : abonnement + confirmation

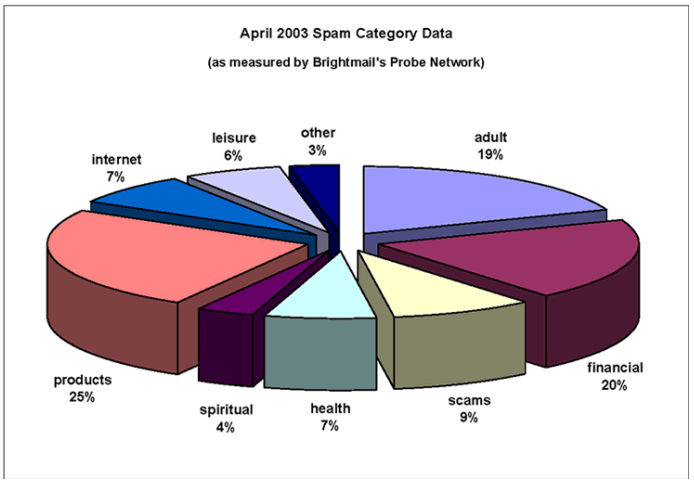
Pourquoi filtrer le SPAM ?

- C'est entre 30 et 60% du trafic email journalier.
- Consommateur
  - en bande passante
  - en CPU, mémoire, espace disque
- Les usagers ont une messagerie polluée
  - Inefficacité du traitement du courriel => donc perte d'argent pour les usagers/l'entreprise.
- Pas de possibilités de se désinscrire
  - Les adresses « Unsubscribe » fonctionnent à 37% (Statistique U.S.)
- Pas d'action législative efficace pour le moment

Le volume du Spam explose !!



Les différentes catégories de SPAMS 1/2



Les différentes catégories de SPAM 2/2

- Adult : Porno, Rencontres, …
- Health : Santé, Herbe, Médecine, …
- Products : Vente de produits divers ..
- Financial : Finances, Banque, …
- Scams : Chaîne d'argent, Nigéria,Escroquerie…
- Internet : Hébergement, Ventes liste email, ..
- Leasure : Casino, Jeu, …
- Spiritual : Astrologie, Org.Religieuses, …
- Autre : le reste.

Les impacts économiques du SPAM

- Il ne coûte rien au spammeur
- Coûte 10 millards d'Euros par an en Europe
- => un marché est apparu : combattre les SPAMS. De nombreuses sociétés sont présentes sur ce secteur, et sont actuellement rejointes par les éditeurs d'antivirus.

Le combat contre le Spam

- La préhistoire :
  - Fabrication de filtre « interne » à chaque entité qui sont maintenu par le staff local.
  - Effectue la destruction des messages des spammeurs notoires identifiés localement
  - Ensuite qqs filtres sur des « mauvais mots »
- Très très dur à configurer, fonctionnement centralisé, beaucoup de travail pour le maintien à jour

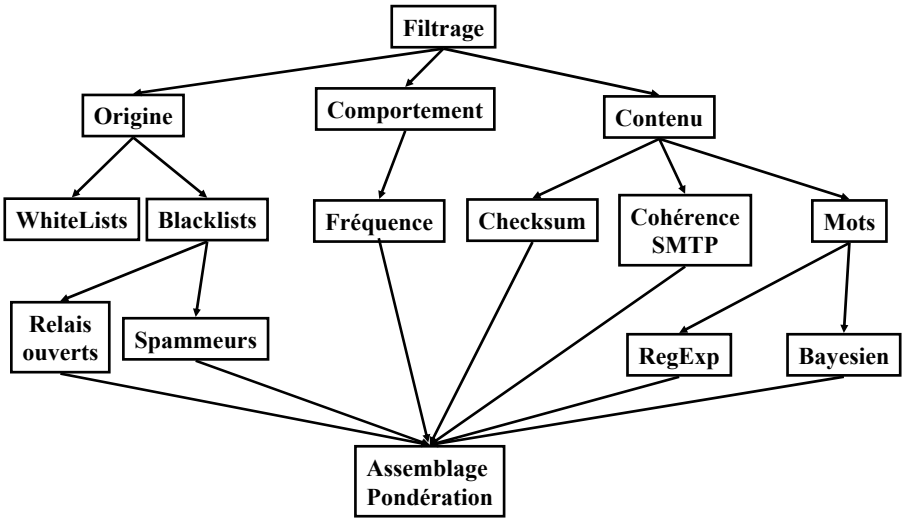
Le combat contre le Spam

- Le passé très proche : DNS Blacklist
  - Identifier les sources de spam par leur adresse IP
  - Autoriser le système mail de regarder dans une base sur Internet lorsque le courriel arrive
  - Bloquer le courriel, si l'adresse de l'émetteur est « blacklisté ».
  - Existence de près de 20 Serveurs « DNS Blacklist »
  - Attention au faux « positifs » sur des serveurs « DNS Blacklist » mal géré.

Le combat contre le Spam

- Les solutions émergeantes
  - **Tendre où l'on peut au Zéro-configuration**
  - **Utilisation de plusieurs règles pour déterminer si un message est un SPAM ou non.**
    - Utilisation de logique « floue » pour les règles
    - Le résultat de ces règles est combiné pour produire un « score »
    - En fonction d'un seuil défini, le message est considéré comme un SPAM.
  - **Une règle seule ne peut plus déterminer si un message doit être considéré comme un SPAM**

Filtrage : De nombreux critères



Un outil anti-spam : SPAMASSASSIN 1/4

- Merci à Denis Ducamp
- http://www.spamassassin.org
- Filtre à pondération
- Version 2.5x (intègre du bayesien)
- En perl
- Disponible avec un daemon et un client évitant le rechargement de perl

Un outil anti-spam : SPAMASSASSIN 2/4

- Ensemble très important de tests (923)
  - Analyse des en-têtes
  - Analyse du contenu
  - Listes noires (MAPS, ordb, spamhaus, etc.)
  - Listes collaboratives de checksum (RAZOR 1 & 2) (plus pour très longtemps)
  - Bayesien (depuis la 2.50)

Un outil anti-spam : SPAMASSASSIN 3/4

- Pondération obtenue par algorithme génétique
- Exemples de tests
  - Habeas -6.4
  - Razor2 0.001
  - Faked Undisclosed-recipients 4.3
  - Référence au Senate Bill 1618 2.8

Un outil anti-spam : SPAMASSASSIN 4/4

- Résultat obtenu sous forme d'étoile
  - X-Spam-Status : ******
  - X-Spam-Status : **********
- Facile à utiliser pour les clients
- Permet à l'utilisateur de régler le niveau
- Très volubile si nécessaire

Eviter d'être dans les bases des SPAMMEURS – Comment ?

- Eviter de donner une vraie adresse email
  - Utiliser une autre adresse pour poster sur des forums et newsgroups
- Cocher systématiquement les opt-out
- Modifier les liens
  - Crypter ses @ emails sur les sites Web pour éviter la collecte des robots.
- Ne pas se désabonner (confirmation)

Combattre le SPAM entrant à l'UJF ?

- Hypothèses
  - La législation interdit la destruction de correspondance
  - Les bayesiens nécessitent 1000 mail
  - Les filtres efficaces sont lourds en traitement
- 2 Solutions sont possibles :
  - Evaluation au niveau du site et décision individuelle
    - header de statut ou de probabilité de SPAM positionné
    - filtre basé sur la présence du header sur le MUA
  - Evaluation et décision individuelle (popfile, ifile)

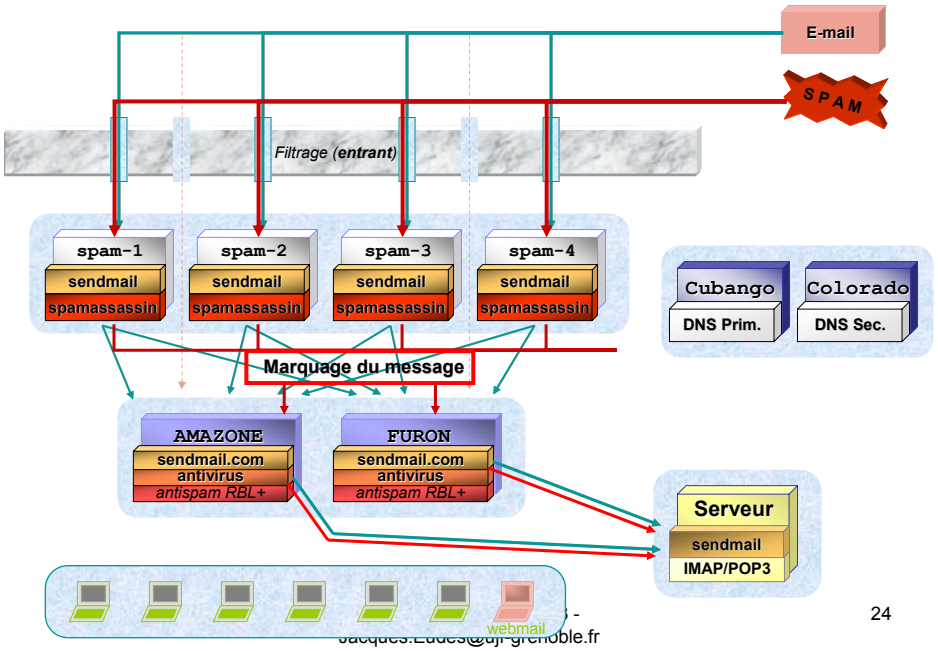
Combattre le SPAM entrant à l'UJF ?

- Choix de la solution qui effectue l'analyse de manière globale. Pourquoi ?
  - 331 Machines déclarées pour le courrier. En fait 22 opèrent un véritable service de boîtes aux lettres.
  - Décharger les serveurs de messagerie de ce traitement
  - Les traitements anti-spamming évoluent « rapidement »
  - Demande une ressource relativement importante pour effectuer le traitement de détection qu'il convient de mutualiser
  - Mutualisation de la gestion de cet outil.
  - Accord écrit de notre présidence pour la mise en place d'outil d'inspection des messages (obtenu lors de la mise en place de la solution antivirus)

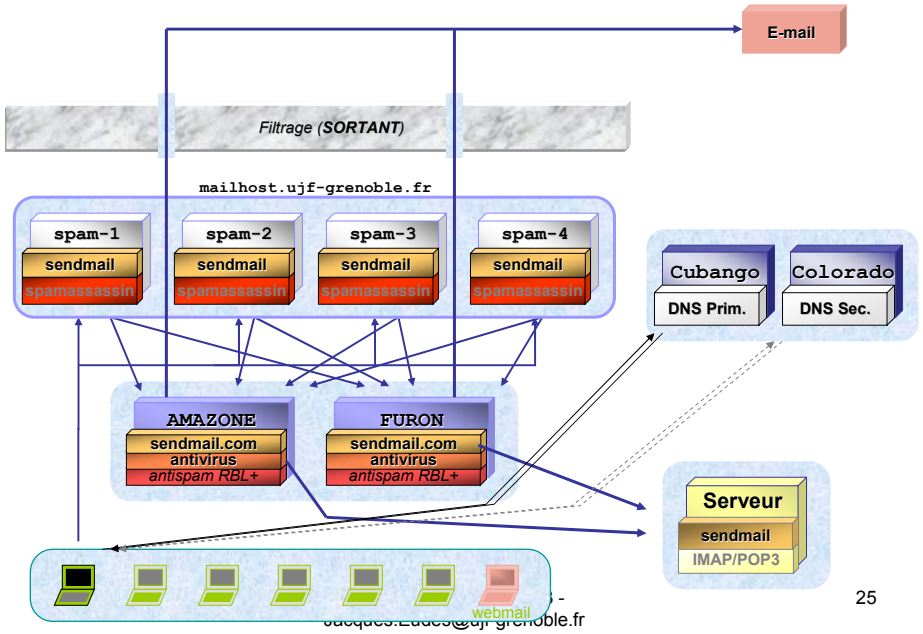
Combattre le SPAM entrant à l'UJF ?

- Le message sera seulement « marqué » et non détruit de manière automatique
- Ce marquage sera exploitable par les agents de lecture de messagerie capable d'effectuer les filtrages sur le sujet par exemple.
- Le marquage sera effectué dans le sujet : entête {SPAM?} ajoutée.
- Ce sera l'utilisateur qui décidera de détruire automatiquement ou non ces messages marqués avec son MUA. Changement d'outils de lecture des emails : choix entre Netscape, Mozilla, Outlook, MailMan

Relais de messagerie : fonctionnement/SpamAssassin (entrée)



Relais de messagerie : fonctionnement/SpamAssassin (sortie)



Outils utilisés

- Spamassassin
- Sendmail (Sendmail.org et sendmail.com)
- RBL+ acheté pour l'accès aux bases DNSBL gérée de manière fiable.
- Antivirus TrendMicro
- MailScanner