

Les spams : la pollution informatique

Xavier D'Hondt, Vincent Stradiot, Jérémy Vion

« Encore un spam ! » Combien de fois ne vous êtes-vous pas fait cette remarque en ouvrant votre boîte mail ? En effet, vous êtes certainement familiers de ces messages publicitaires qui nous envahissent littéralement. Mais vous êtes vous déjà demandé la raison de la réception de ces messages indésirables ? Nous allons tenter de vous faire comprendre la problématique ainsi que les enjeux de ces courriers si envahissants.



En premier lieu, il faut savoir qu'il existe 2 grandes familles de spams : les spams mails et les spams web. Pour cette exposition, nous nous sommes attaqués au type le plus connu et le plus répandu, le spam mail sous forme de texte. Toutefois, il est bon de connaître l'existence des autres types.

Les spams sont devenus au fil des années un gros problème dans le monde de l'informatique. À un tel point qu'on estime qu'ils représentent aujourd'hui entre 90 et 97% du nombre total d'e-mails qui circulent chaque jour sur internet.

À première vue, il est difficile de comprendre pourquoi tant de ces messages indésirables se retrouvent chaque jour dans nos boîtes de réception, mais une fois qu'on saisit la portée de cette publicité peu coûteuse, notre esprit s'éclaire. En effet, le spamming de masse peut rapporter entre 5000 et 100 000 € par jour à un expéditeur de masse. Cependant, il y a des risques d'amendes faramineuses ainsi que de lourdes peines de prison.

D'un autre côté, les désagréments sont légion : perte de temps, perte d'argent pour les entreprises, source d'ennui, pollution, insécurité, etc. Le manque à gagner pour les entreprises peut se chiffrer en millions d'euros.

Le moyen le plus simple pour « arrêter » les spams sont les filtres. Il existe différents types de filtres qui classifient les mails entre « bons » et « mauvais » suivant le contenu, les adresses des émetteurs ou encore le serveur mail d'où provient le message. Les filtres de contenus peuvent être basiques en refusant simplement certains mots ou plus complexes en se basant sur des probabilités qu'un mail soit un spam en analysant chaque terme du message.

À notre stand, nous expliquerons un peu plus en détail le filtre bayésien, qui est le filtre le plus utilisé dans le monde ainsi que le plus efficace. Cette explication sera accompagnée d'un exemple afin de se rendre compte du fonctionnement. Les élèves pourront également tester les filtres que nous avons implémentés, séparément ou en les combinant, en simulant l'envoi de mail vers une adresse électronique. Ils pourront créer leur propres « spams », utiliser les spams qu'ils ont reçu sur leur adresse personnelle ou bien se servir dans une série de messages des mails récupérés un peu partout.

Nous donnerons également quelques conseils simples et efficaces afin d'éviter que les spams n'envahissent leurs boîtes mails. « Le meilleur moyen de gagner la bataille est de ne pas avoir d'ennemi à combattre ! »

