IP SQUATTING APPLIQUÉ AU SPAM

Jérôme NICOLLE (@chiwawa_42) – Ceriz – jerome@ceriz.fr

Arnaud FENIOUX (@afenioux) – France-IX – afenioux@franceix.net

mots-clés : RÉSEAU / SPAM / BGP / HIJACKING / USURPATION

Les spammeurs sont des gens très créatifs. Progressant plus rapidement en réseau qu'en orthographe, certains ont industrialisé une pratique jusque-là connue des seuls bas-fonds de l'Internet : l'IP-Squatting.

1 L'attribution des adresses sur Internet

Internet permet le routage de paquets utilisant les protocoles IPv4 et IPv6 entre des adresses assignées à des réseaux, appelés « systèmes autonomes » (ou « AS ») identifiés par un *Autonomous System Number* (AS).

Les ressources devant être uniques pour que le réseau fonctionne, une organisation mondiale, l'*Internet Assigned Numbers Authority* (IANA), gère les assignations des numéros de réseau et des blocs d'adresses.

Avec la croissance du réseau, l'IETF a recommandé en 1992 que des entités continentales, les *Regional Internet Registries* (RIR) soient créées pour supplanter l'IANA. Elles sont aujourd'hui au nombre de cinq : AFRINIC, APNIC, ARIN, LACNIC et RIPE NCC. Les ressources jusque-là gérées par l'IANA sont déléguées aux RIR, et les assignations prédatant cette délégation sont maintenues avec un statut particulier, non contractuel. Toute nouvelle allocation désormais est gérée par un RIR et soumise à certaines règles et obligations. Pour la zone RIPE à laquelle est rattachée l'Europe, ces règles sont énumérées dans le document « RIPE-649 » **[0]**

Les relations entre un opérateur et un RIR peuvent prendre plusieurs formes, mais sont essentiellement contractuelles et requièrent le versement d'une cotisation annuelle ou de frais de gestion – fussent-ils d'un montant très faible. L'absence de paiement de ces frais ou cotisations a pour effet de révoquer les assignations d'identifiants numériques (ASN ou blocs d'adresses).

Les blocs et AS assignés avant la mise en place des RIR n'étant pas soumis au paiement de ces frais, la disparition de l'assignataire ne peut être constaté que par enquête et les ressources assignées sont susceptibles de rester dormantes faute d'initiative pour les récupérer.

Chaque réseau se voit assigner un numéro d'AS et un ou plusieurs blocs d'adresses. Les allocations sont consignées – plus ou moins rigoureusement – dans des registres (nommés IRR – *Internet Routing Registry*).

2 Le routage des blocs attribués

Le routage entre les réseaux constituant Internet se fait au moyen du protocole BGP (*Border Gateway Protocol*), qui permet à chaque réseau d'annoncer à ses pairs ses blocs d'adresses, et en retour de recevoir la liste des blocs que ces pairs savent joindre.

Chaque AS va établir des sessions BGP avec ses partenaires. Le protocole crée une relation de confiance implicite. Pour qu'un bloc d'adresse soit joignable, il doit être annoncé à tous les autres réseaux d'Internet. Ces annonces peuvent être directes (*peering*) ou indirectes (transit, par le biais d'un autre réseau).

Lorsqu'un réseau reçoit d'un de ses pairs une annonce pour un préfixe, il peut en contrôler la légitimité pour accepter ou ignorer cette annonce. Cela se fait au moyen de filtres configurés sur chaque routeur et pour chaque session BGP. Ces filtres peuvent être basés sur les informations contenues dans les IRR, nommément dans des objets de type « inetnum » et « inet6sum » désignant les blocs d'adresses, et des objets de type « route » et « route6 » consignant la légitimité d'une annonce telle que déclarée par le gestionnaire identifié de l'AS et des blocs d'adresses. La consultation des registres est toujours au moins possible par le protocole whois.

Un mécanisme plus récent, RPKI/ROA **[1]**, permet au routeur d'interroger un serveur, qui va à son tour interroger les registres, pour déterminer si une annonce est légitime ou pas, sans avoir eu à configurer de filtres spécifiques préalablement. Ce mécanisme est encore très rarement utilisé, car peu de registres sont suffisamment bien tenus et peu d'opérateurs acceptent la surcharge de travail que requiert le déploiement de ce mécanisme.

De fait, les interconnexions reposent encore majoritairement sur la confiance entre les opérateurs, qui configurent soit une limite au nombre de préfixes qu'un AS peut annoncer, soit une liste stricte de préfixes autorisés à être annoncés sur cette session, soit parfois aucun filtre. Le meilleur exemple est l'incident de juin

2015 affectant Level3 **[3]** (le plus gros opérateur IP au monde) après qu'un de ses clients, Telecom Malaysia, lui ait annoncé qu'il savait joindre tout Internet, à cause d'une erreur de configuration.

3 Changement de priorité

Revenons-en au SPAM. La particularité de cette activité est qu'elle est très peu chère à pratiquer, car la part de travail coûteuse, le traitement et le stockage des messages, est à la charge du destinataire. C'est ce qui fait la rentabilité du SPAM depuis des décennies.

Une autre particularité du SPAM est qu'il requiert une implémentation (à peu près) correcte du protocole SMTP. Celui-ci fonctionnant sur le port TCP/25, une communication bidirectionnelle est donc indispensable.

La chasse aux open-relays

Dans les premiers réseaux de messagerie (UUCPNET, BITNET, FidoNet...), chaque serveur de messagerie acceptait de recevoir, stocker et retransmettre des messages pour pallier aux limites imposées par les connectivités disponibles à l'époque. Au milieu des années 90, les spammeurs ont commencé à utiliser ces « Open-Relays » afin d'amplifier leur capacité d'envoi de messages aux frais de tiers, et de mélanger les messages illégitimes au flux de messages traités par des serveurs légitimes.

Afin d'endiguer cette pratique, des recommandations officielles sont apparues à la fin des années 90 et le protocole SMTP a été enrichi de fonctionnalités permettant de limiter ces pratiques. Les distributions UNIX ont modifié les configurations par défaut des logiciels de transport de messages, et la proportion de serveurs de messagerie « Open-Relays » est passée de 90% à moins de 1% en moins d'une décennie [2].

La disparition des Open Relays a permis de faciliter la distinction des serveurs légitimes de ceux dédiés au SPAM, et a favorisé l'apparition des Blacklists recensant ces derniers pour permettre les filtrages de niveau 3.

La lutte contre le SPAM a utilisé plusieurs techniques au fil des années. Le filtrage par analyse du contenu (Spam Assassin) a vite montré ses limites (consommation de ressources CPU côté destinataire) face à la croissance du trafic, et le risque de faux-positifs a toujours été un facteur limitant dans la précision de son fonctionnement.

Le filtrage par contenu est donc plus souvent couplé d'autres méthodes. Le *Grey Listing* par exemple repose sur le fait qu'un spammeur ne tentera pas la retransmission du mail si celui si est temporairement rejeté. Il s'agit alors d'une règle de filtrage applicative. Mais la technique la moins coûteuse reste le filtrage par l'adresse IP de l'émetteur (souvent lié à l'utilisation de *DNS-based Blackhole List* – DNSBL – pour une mise à jour en temps réel).

Un spammeur va généralement utiliser une machine dédiée à son activité. L'adresse IP de cette machine n'est pas supposée servir à des envois de messages légitimes depuis que les campagnes de chasse aux open-relays ont fonctionné. C'est ainsi que les adresses IP « propres » (non listées dans une Blacklist) sont devenues la matière première du SPAM.

4 Chercher la ressource

Pour « travailler », un spammeur a donc besoin d'une machine disposant d'une connectivité performante et surtout d'une adresse IP répondant aux caractéristiques suivantes :

- libre (ou presque) ;
- inconnue des listes de blocage ;
- correctement routée sur l'ensemble d'Internet.

Le dernier point fait débat, car il n'est pas si indispensable que ça. Lorsque la majorité des adresses mail cibles est gérée par un nombre réduit d'hébergeurs, il suffit d'obtenir un routage vers ces quelques hébergeurs pour que les messages soient transmis. Les plus gros hébergeurs étant présents sur des points d'échanges avec des politiques d'interconnexion ouvertes, il est relativement aisé de cibler le point d'échange le plus laxiste possible en terme de filtrage pour que ces hébergeurs reçoivent et acceptent les blocs annoncés.

Voici quelques typologies d'adresses qui pourraient servir de matière première pour l'envoi de SPAM.

4.1 Les serveurs dédiés

Les spammeurs se sont tournés vers des hébergeurs afin de louer une machine connectée et son adresse. Une fois la campagne de messages envoyée, l'adresse était repérée comme ayant servi à l'envoi de SPAM et ajoutée dans des listes de blocage, la rendant impropre à l'envoi de mails légitimes par la suite, et ce pour une durée parfois longue.

Certaines listes indélicates ont pris l'habitude de bloquer tout l'hébergeur d'un coup, perturbant l'envoi de messages légitimes d'autres clients de cet hébergeur. Ce dernier n'avait alors parfois pas d'autre choix que de payer l'éditeur de la liste de blocage pour dé-lister ses réseaux, sans garantie qu'ils ne soient pas rapidement listés à nouveau par la faute d'un client indélicat.

Les hébergeurs sont ainsi devenus plus regardants quant aux activités de leur clientèle. Certains sont allés jusqu'à filtrer tout les flux de mails sortant de leur réseau pour annuler l'intérêt de leur offre aux yeux des spammeurs et éviter d'avoir à payer régulièrement les éditeurs de blacklist – et dédommager et/ou perdre les clients légitimes impactés.

D'autres hébergeurs sont par contre relativement tolérants, permissifs, voire enthousiastes – moyennant finances – à accueillir ce type d'activités. On les dit « Bulletproof ».

4.2 Les connexions résidentielles

Au début des années 2000, la méthode prévalant pour l'envoi massif de SPAM était l'utilisation de botnets de milliers ou millions de machines résidentielles infectées. Ces machines pouvaient tenter de joindre directement des serveurs SMTP destinataires de messages.

Les premières blacklists ont rapidement intégré les blocs d'adresses résidentielles comme des sources illégitimes de mails. Des fournisseurs d'accès ont commencé à bloquer le trafic à destination du port TCP/25 d'autres machines que le relais SMTP mis à disposition de son abonné. Ce relais implémente traditionnellement des mécanismes de filtrage par analyse de contenu.

De nouvelles utilisations des botnets ont favorisé leur développement et en font néanmoins un vecteur toujours d'actualité pour la diffusion du SPAM.

4.3 Des blocs assignés au spammeur

Se voyant refusés par les hébergeurs, certains spammeurs ont créé leurs propres réseaux en obtenant des assignations de blocs d'adresses. Qui ont naturellement été très rapidement blacklistés - ce qui déclenche parfois des DDoS spectaculaires en représailles, comme l'affaire Spamhaus en 2013. Un spammer avait alors lancé une attaque massive paralysant les services du principal éditeur de listes de blocage **[4]**.

La raréfaction du nombre d'adresses IPv4 disponibles auprès des registres a créé un marché d'achat et location de blocs d'adresses. Malgré la moins-value qu'un spammeur crée pour un bloc lors de son utilisation, certains loueurs d'adresses acceptent des marchés à tarifs élevés pour mettre à disposition ces adresses de façon temporaire.

L'offre de blocs IPv4 disponibles diminuant et la demande étant à peu près constante, la rentabilité des campagnes de spam a poussé leurs auteurs à se tourner vers une autre source de matière première (adresses IP non blacklistées).

4.4 Des blocs inutilisés (quoi que...)

Le comptage de grands nombres d'adresses se fait par multiple de 2^24 – soit environ 16,7 millions – qu'on appelle « /8s » (prononcé « slash eight »). Sur les 256 /8s, 168,7 sont annoncés et routés sur Internet ; 35,3 sont réservés ; 3,6 ne sont pas encore assignés. Il y a donc 48,3 /8s qui sont assignés, mais pas utilisés dans le routage public.

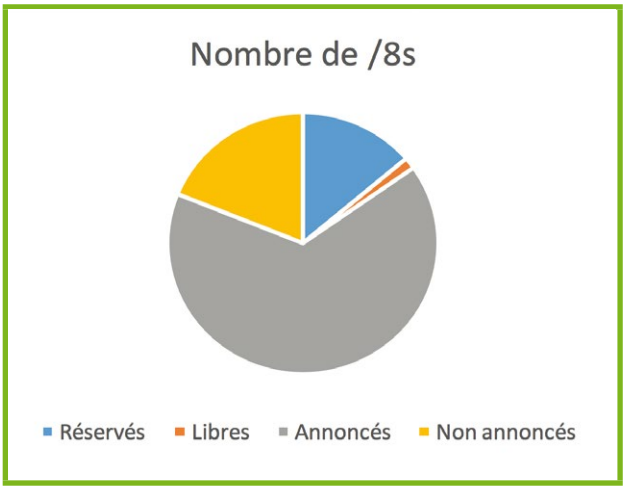

Figure 1 : Répartition de l'utilisation des blocs d'adresses IP.

La majeure partie de ces adresses correspondent à des assignations historiques, désignées *Legacy* ou *Early Registration Transfer* (ERX). Des compagnies ou administrations se sont vues alloués des /8s entiers et continuent parfois de les utiliser pour leurs réseaux internes. D'autres blocs ont été assignés à des entités qui ont disparu, par fusion ou faillite, et dont personne n'a réclamé les assignations. Ces blocs ne peuvent pas être réassignés faute de cadre contractuel concernant l'assignation initiale.

Les RIR ont mené plusieurs campagnes afin de formaliser ces assignations historiques, mais la disparition de l'assignataire est un cas de figure qui empêche tout mouvement. Afin de remettre la main sur des blocs abandonnés, certains spammeurs ont organisé l'usurpation d'identités d'entreprises et/ou de leurs dirigeants afin d'obtenir des mises à jour des registres offrant la légitimité apparente des annonces de tels blocs **[5]**.

Dans d'autres cas, plus nombreux, on voit simplement des préfixes apparaître dans la table de routage globale, derrière des AS n'ayant aucune légitimité apparente à les annoncer ou relayer. Une recherche ciblée à la suite de quelques incidents mineurs a par exemple révélé une grande quantité d'annonces illégitimes de l'AS43239 au cours de l'été 2014 **[6]**.

Comment les spammeurs parviennent-ils à router ces blocs qui ne leur ont pas été assignés ?

5 (Complicité d') Abus de confiance

Reprenons l'exposé depuis le début. Un spammeur a besoin d'adresses IP « propres » pour travailler. S'il obtient une assignation ou achète un bloc, celui-ci sera rapidement identifié par les services d'anti-spam. Sachant que le bloc sera blacklisté, peu d'opérateurs acceptent d'en louer aux spammeurs. Ces derniers n'ont

donc qu'une solution : utiliser n'importe quel bloc – de préférence « disponible » (i.e. pas actuellement annoncé sur Internet) – sans en avoir la légitimité.

Pour pouvoir router ce bloc sur Internet, il faut qu'au moins un opérateur accepte de fournir un transit au spammeur, et que les annonces soient relayées et acceptées par les pairs et transits de cet opérateur. De deux choses l'une :

- soit le spammeur parvient à abuser la confiance d'un registre afin de prendre le contrôle des enregistrements représentant le routage légitime du bloc d'adresses (via récupération d'un nom de domaine abandonné, pointé dans les enregistrements) ;
- soit les opérateurs acceptent de fournir un transit au spammeur sans aucun contrôle de la légitimité des annonces.

Il est assez facile de berner les opérateurs faisant des contrôles de légitimité sur des IRR peu sûrs (par exemple avec RADb **[7]** qui ne fait pas de vérification avant d'accepter un nouvel enregistrement **[8]**).

Alors que la première approche repose essentiellement sur le social-engineering, la seconde consiste finalement à trouver un complice plus ou moins volontaire. Pour toucher un plus grand nombre de destinataires, il faut réunir plusieurs composants :

- un hébergeur bulletproof, qui acceptera d'héberger un spammeur ;
- cet hébergeur doit avoir de nombreux transitaires et peerings, configurés de façon aussi laxiste que possible, pour que les annonces remontent jusqu'à au moins un Tier1 ;
- dès qu'un Tier1 reçoit et accepte l'annonce, alors la visibilité totale est quasiment garantie.

Notez que la plupart des contrôles se font en associant le numéro d'AS au bloc d'adresses annoncé. Si le spammeur usurpe les deux, et qu'un opérateur accepte une session avec un AS usurpé sans en avoir contrôlé la légitimité, alors la propagation est encore plus facile, car il n'y a pas de modification à effectuer dans les registres. Cela peut arriver dans deux cas : soit le transitaire est complice, soit le spammeur usurpe l'identité de l'assignataire de l'AS et des blocs d'adresses avec succès **[9]**.

De même, si un réseau accepte de fournir un transit (Tier1 ou Tier2) à un réseau lui-même transitaire pour d'autres réseaux (Tier2), alors il pourrait vouloir s'assurer que les clients de son client (Tier3) sont réels et légitimes, car le Tier3 pourrait tout à fait configurer un routeur pour qu'il s'annonce comme étant l'AS d'origine légitime d'un préfixe usurpé. La vérification transitive (à travers plusieurs AS) est plus compliquée que si elle est effectuée dès le début (en edge), car elle implique la construction de filtres plus complexes et basés sur les déclarations des intermédiaires aux registres.

Une fois le bloc annoncé, accepté et globalement joignable, alors l'envoi de SPAM peut commencer. Celui-ci peut durer de quelques minutes à quelques heures, en fonction de la réactivité des éditeurs de listes de blocage.

Le langage RPSL

C'est à un besoin de transparence que répond le langage RPSL (*Routing Policy Specification Language* - RFC2622) utilisé pour décrire les relations inter-AS au sein de certains registres. Deux problèmes se posent alors :

- l'information doit être la plus exhaustive possible ;

- elle doit être maintenue à jour.

Comme cette information caractérise des relations diverses dont certaines sont d'ordre commercial, certains acteurs considèrent qu'il n'est pas souhaitable de les publier sur un registre opposable en vertu du secret des affaires.

Une fois le SPAM envoyé, l'annonce est abandonnée et le préfixe sombre à nouveau dans l'oubli. Il est désormais terni par une mauvaise réputation dans les listes de blocage et ne devrait plus être utilisé avant l'expiration de ces enregistrements.

L'opération peut se poursuivre sur cette infrastructure en utilisant un autre bloc. Dans le cas « Telelatina » (AS15078), plus de 1000 blocs ont ainsi été utilisés entre juillet et septembre 2014 **[10]**. Le motif était alors très reconnaissable puisqu'environ 8 préfixes étaient annoncés simultanément, utilisés, puis retirés alors que d'autres étaient annoncés, et ce plusieurs dizaines de fois à des intervalles de 8 à 24h.

6 Victime collatérale, complicité ou simple chatouillis ?

Une constante est visible dans ces « attaques » : la confiance naturelle entre les opérateurs - induite par le fonctionnement même de BGP - est abusée par un des AS intermédiaires. Quelles actions sont envisageables ?

Les guillemets sont importants ici : il ne s'agit pas d'une attaque, car les usurpations ne perturbent pas – dans le cas de leur usage par les spammeurs, et tant que le bloc ciblé n'était pas utilisé – le fonctionnement d'Internet ou d'un des réseaux qui relayent les annonces.

Il n'y a pas non plus nécessairement de violation de clauses contractuelles, puisque les assignations des blocs d'adresses, les interconnexions entre opérateurs, et même le raccordement à certains points d'échanges ne sont pas forcement soumis à un contrat.

Enfin, il ne semble pas y avoir de crime ou délit défini en droit français ou américain qui puisse s'appliquer à l'utilisation illégitime d'un bloc d'adresses dormant. Seuls certains moyens d'y parvenir sont actuellement répréhensibles (fraude, faux et usage de faux, usurpation d'identité).

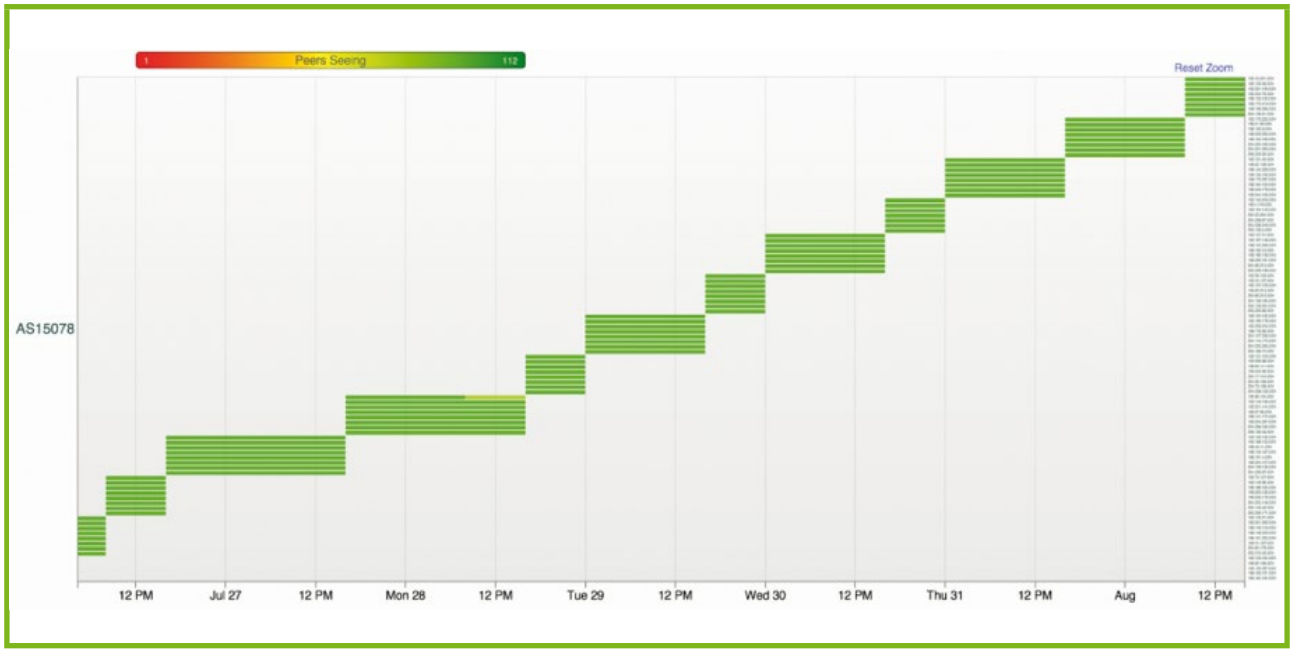

Figure 2 : Annonce de blocs usurpés par Telelatina.

Internet repose tout entier sur des conventions, qu'elles soient établies de gré à gré pour les interconnexions ou sous une forme plus institutionnelle au travers de l'IETF ou du RIPE par exemple. Le non-respect des conventions ou *Best Current Operationnal Practices*, n'est pas fondamentalement répréhensible. Simplement mal vu par la communauté qui fait fonctionner le réseau.

À première vue, il n'y a donc pas d'autre recours qu'un traitement de l'anomalie par la communauté elle-même. La première réponse à un abus est la mise en place de filtres qui ne l'avaient pas été jusque-là, ou bien directement le *de-peering* (rupture unilatérale d'interconnexion).

La perte de confiance de la communauté à l'égard d'un complice – volontaire ou non – d'une campagne d'IP-squatting ne prendrait donc a priori pas d'autre forme qu'une méfiance finalement pas forcement malvenue.

Il n'en va pas de même lorsqu'une annonce intentionnellement frauduleuse perturbe le fonctionnement d'un réseau, par exemple en annonçant un bloc déjà utilisé ailleurs dans le réseau, mais une tolérance s'applique lorsque l'intention n'est pas manifeste. La plupart des pannes provoquées par des usurpations en BGP sont historiquement dues à des erreurs de configuration, pas à des attaques.

7 D'autres applications possibles

Un grand nombre de facteurs combinés permettent ce type de détournement. Pour résumer, citons les principaux :

- faible qualité de certains registres, aggravée par la marchandisation des blocs d'adresses en période de pénurie ;
- nature des interconnexions entre opérateurs, historiquement basée sur la confiance entre un nombre relativement limité d'acteurs ;
- complexité opérationnelle d'un réseau d'opérateur dont la mise à jour d'un routeur entraîne généralement des interruptions de service ;
- manque d'automatisation de ces réseaux, qui ne fonctionneraient de toute façon qu'avec des registres de bonne qualité ;
- inertie des équipementiers qui vendent très cher des implémentations tardives, voire incomplètes, des moyens de contrôle.

L'IP-squatting est donc globalement aisé pour quiconque ayant accès à une ou plusieurs interconnexions de confiance (lire « laxiste »).

Lorsqu'il est temporaire, souvent faute de logs des annonces quotidiennes de la table routage globale, il est difficile à repérer a posteriori. Il apparaît clairement aux acteurs majeurs de la lutte contre le spam que cette pratique a le vent en poupe.

Une attaque bien construite (comme Pilosov et Kapela **[11]**) permet d'aller encore plus loin que le simple SPAM, en détournant le trafic à destination d'un bloc pour implémenter un MiTM à grande échelle. C'est de cette manière que pourraient s'opérer des campagnes d'espionnage industriel **[12]**, car le chiffrement correct des mails est difficile si ce n'est impossible avec les systèmes de messagerie employés au sein des entreprises et institutions visées. ■

Retrouvez toutes les références accompagnant cet article sur http://www.miscmag.com