http://www.phishing-initiative.com

Un projet de :

L'action de Phishing-Initiative en 2011

[1] http://creativecommons.org/licenses/by-nc-sa/3.0/

Sommaire

Introduction

Le 12 janvier 2012, le projet Phishing-Initiative, mené conjointement par le Cert-Lexsi (LEXSI Group) Microsoft et PayPal, soufflait sa première bougie.

Ce rapport annuel revient sur les actions entreprises par cette initiative collaborative et sans but lucratif de lutte contre le phishing. Le rapport dresse ainsi un état des lieux du phénomène du phishing en France pour l'année écoulée, d'après les soumissions qui ont été reçues et analysées au cours de la période.

Phishing-Initiative est une association loi 1901 depuis le 7 février 2011.

Remerciements

Nous remercions B. OURGHANLIAN et P. SAULIERE (Microsoft), B. LATHOUD (PayPal), T. ANDRIAMAHAZOSOA et W. POMIAN (Cert-Lexsi) pour leur aide dans la réalisation ou la relecture de ce document.

Nous remercions l'ensemble des contributeurs, particuliers, entreprises et administrations, qui nous ont remonté les adresses de sites suspects.

Synthèse

Plus de 12 500 des 20 000 adresses distinctes soumises en 2011 sur le site phishing-initiative.com relevaient d'une tentative de phishing, soit environ 35 par jour. Le ratio de sites validés par rapport aux signalements transmis s'établit donc à près de deux tiers d'adresses frauduleuses. De plus, un grand nombre de sites refusés l'ont été car les contenus suspects n'étaient plus disponibles au moment de notre vérification. 30 000 soumissions ont été enregistrées au total, avec une évolution mensuelle de plus 6% environ.

Plus de 6 000 individus ont a priori contribué, soit 5 soumissions en moyenne par personne. Les clients de FAIs français et détenteurs de comptes e-mail gratuits forment 90% des contributeurs. Mais un petit nombre d'organisations (banques et FAIs surtout) dont l'identité est usurpée dans ces tentatives d'escroquerie sont à l'origine de 20% du nombre de signalements.

Le temps médian pour caractériser et confirmer le caractère frauduleux d'un site soumis a été de 13 minutes environ. Plus de 70% des adresses suspectes ont été confirmées par Phishing-Initiative en moins de 30 minutes.

Le phishing visant les publics francophones est un phénomène massif, peu étudié à ce stade. Malgré l'efficacité de contre-mesures telles que celles lancées par Phishing-Initiative, sans réponse pénale associée, le phishing risque de perdurer voire de continuer d'augmenter.

Le portrait-robot d'une URL de phishing se présente d'après nos statistiques de la sorte :

une URL accessible en http sur le port 80,

avec un nom d'hôte composé d'un domaine enregistré au sein de l'extension « .com »,

hébergée sur un serveur possédant une résolution IP localisée aux Etats-Unis,

sous la responsabilité d'un prestataire d'hébergement légitime.

Méthodologie

La méthodologie utilisée pour ce bilan s'appuie sur les éléments suivants :

Période	12/1/2011-12/1/2012
Géolocalisation IP des sites	via Maxmind
Hébergeur (Nom et numéro d'AS détenteur des IPs)	via Cymru
Nombre d'adresses IPs annoncées par AS	via BGP
Sources	URLs soumises au sein du formulaire de phishing-initiative.com
Autres sources	Emails à destination de adresses contact@phishing-initiative.com et support@phishing-initiative.com Statistiques fournies par le Cert-Lexsi

Statistiques et résultats

Soumissions reçues via le formulaire

Nombre de contributions

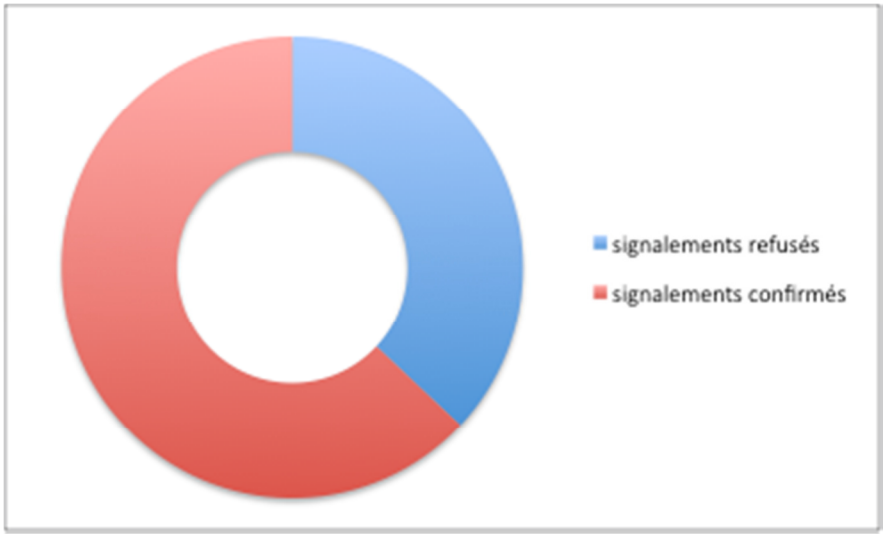
En 365 jours, près de 30 000 soumissions ont été enregistrées, représentant plus de 20 000 URLs distinctes de pages suspectées de faire partie d'une tentative de vol d'informations personnelles ou sensibles.

Les analystes du Cert-Lexsi, en charge de la vérification du caractère frauduleux des soumissions, ont pu confirmer que plus de 12 500 de ces URLs relevaient d'une tentative de phishing. 12 667 URLs, soit près de 35 URLs par jour, ont ainsi été validées et transmises aux partenaires techniques (Microsoft, Google) de Phishing-Initiative pour ajout aux listes noires opérées par ces organisations et blocage dans les navigateurs participants.

Le Cert-Lexsi a également diffusé à ces partenaires plus de 5000 autres URLs de phishing non identifiés et remontées par les contributeurs de Phishing-Initiative.

Ratio de contributions confirmées

Environ 2/3 des contributions se sont donc révélées être de nature frauduleuse. Ce ratio déjà élevé est très conservateur, puisque plusieurs centaines, voire quelques milliers d'URLs parmi les 7 500 soumissions refusées n'ont en fait simplement pas pu être confirmées, car inaccessibles ou suspendues au moment de la vérification.



Répartition des signalements selon leur statut après vérification

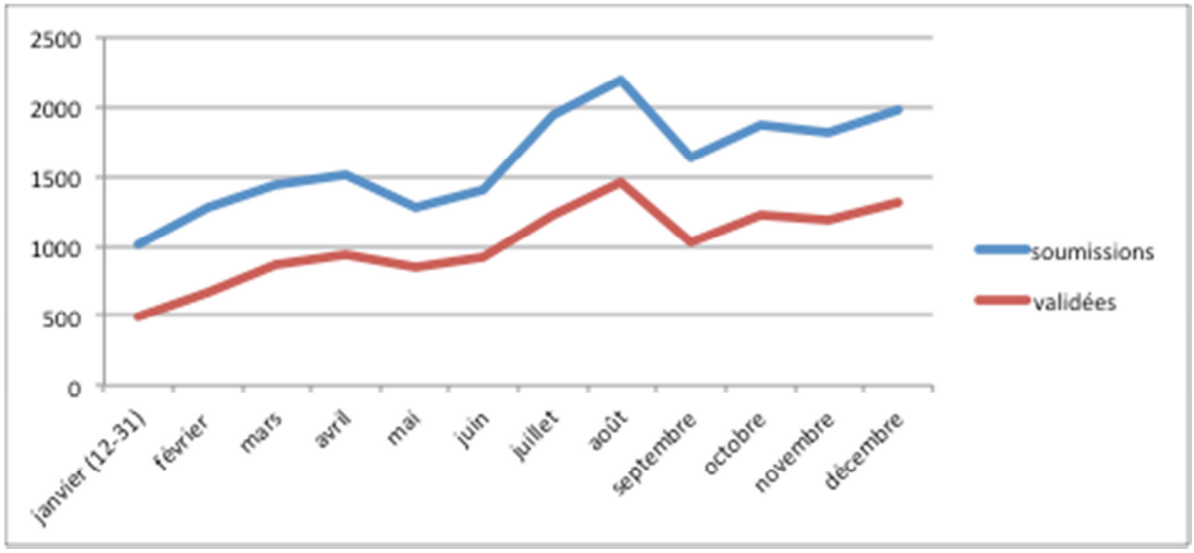
Nous estimons que plus de 15,000 URLs de phishing ont *a minima* été reçues au cours des 12 derniers mois sur notre plateforme.

Evolution mensuelle des soumissions

Nous constatons une augmentation régulière et concomitante du nombre de soumissions et de phishing validés. Une tendance lourde peut donc être identifiée. Deux raisons principales peuvent expliquer cet état de fait :

- le phénomène gagne de l'ampleur et le nombre d'attaques croît,
- de plus en plus d'internautes sont informés de l'existence de cette initiative et nous rapporte les attaques.

L'impact de chacune de ces hypothèses sur cette évolution est impossible à estimer. On note cependant un pic de soumissions en juillet-août, potentiellement du fait d'actions de communications entreprises par les membres du projet (communiqué de presse du 12 juillet 2011) et par un FAI français.



Evolution mensuelle des adresses (nombre d'URLs) soumises et validées

Analyse des URLs de phishing confirmées

Ports TCP/IP : HTTP vs. HTTPS

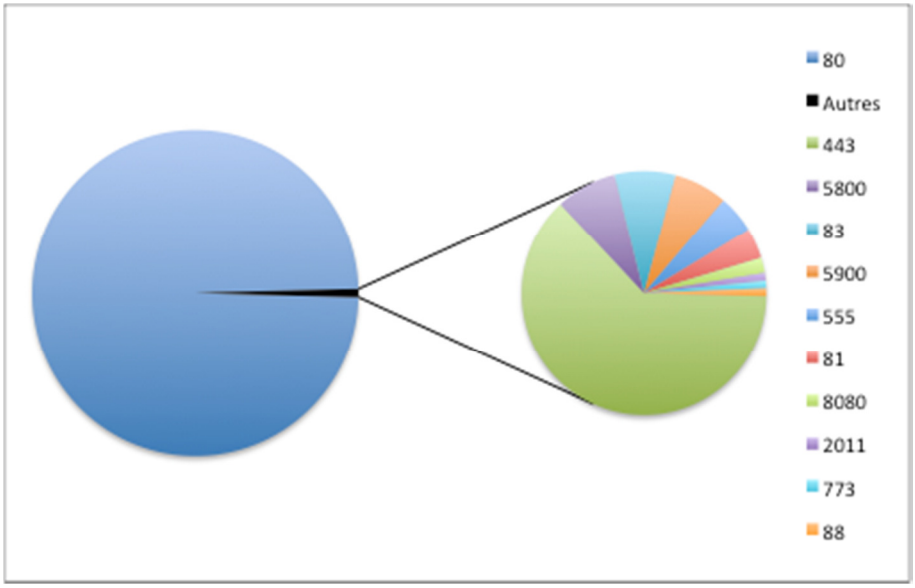
La quasi-totalité (99,2%) des URLs de phishing répondait sur le port 80, le port par défaut pour le protocole HTTP.

Environ 0,5% des URLs étaient accessibles en HTTPS, mais seule une poignée disposait potentiellement d'un certificat SSL spécifiquement créé pour tromper les utilisateurs. Les autres

pages relevaient d'espaces d'hébergement, de services en ligne légitimes ou de sites compromis dont une page en HTTPS servait à héberger le contenu frauduleux.

Seuls 2 cas mettant en scène des domaines frauduleux déposés pour l'occasion sont suspectés d'avoir utilisé un certificat SSL pour crédibiliser une tentative de phishing. Malheureusement les certificats n'ont pu être analysés plus avant par nos soins.

Outre les 0,5% d'URL sur le port 443 (HTTPS), 0,3% des sites (moins d'une quarantaine) étaient accessibles via un port différent. Les ports suivants ont ainsi été ainsi identifiés :



Répartition des URLs de phishing validées d'après leur port d'écoute

Il est à noter que les ports 5800 et 5900 sont réservés traditionnellement au service « VNC » de prise de main à distance sur un poste.

Nom d'hôte : adresse IP vs. domaine

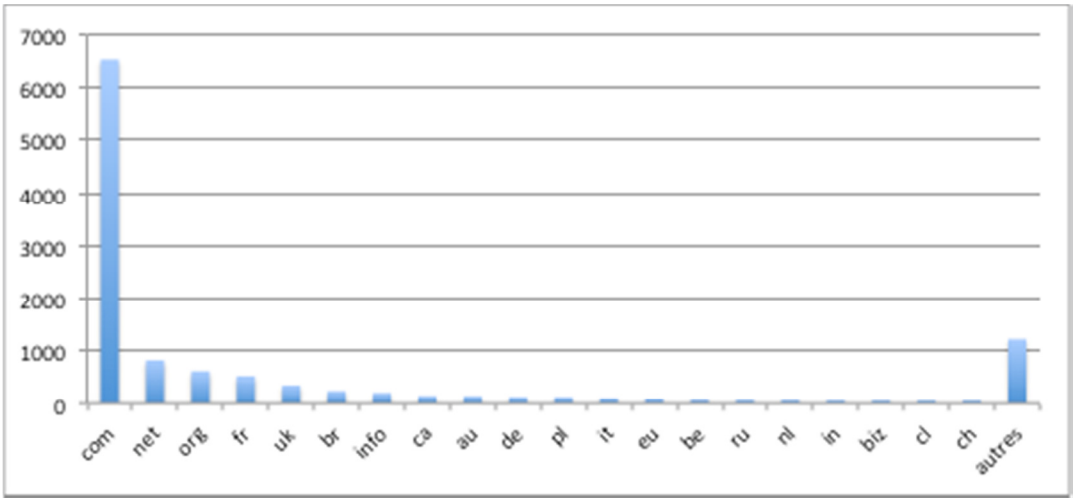
Moins de 5% des URLs validées avaient dans leur nom d'hôte une adresse IP au lieu d'un nom de domaine. Les domaines ainsi utilisés pour héberger ces contenus peuvent être classés en deux grandes catégories:

- les domaines légitimes compromis par le pirate,
- les domaines enregistrés frauduleusement par le pirate
  Cette catégorie comprend elle-même plusieurs sous-catégories (domaine de deuxième ou troisième niveau, déposé gratuitement ou non, etc.)

Domaine : répartition par extension de domaine

Plus de 6 000 domaines différents ont été recensés sur l'ensemble des près de 12 000 pages de phishing comprenant un nom de domaine (qu'il soit compromis ou spécifiquement déposé pour l'occasion), soit un peu moins de 2 pages par domaine.

La répartition de ces domaines d'après leur extension géographique se présente de la sorte :



20 principales extensions de domaines (TLD) par nombre d'URLs de phishing

Plus de la moitié des domaines possédaient une extension en « .com ». Avec les deux autres extensions génériques « .net » et « .org », ils forment ensemble plus des deux tiers des domaines rencontrés.

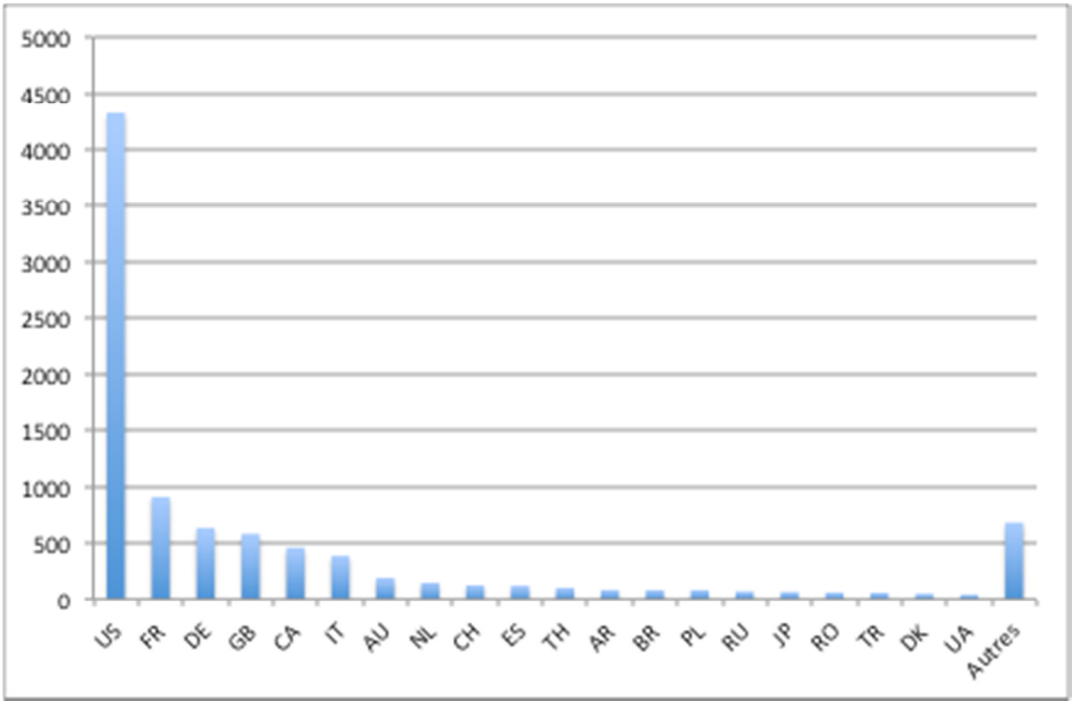
Les URLs comprenant un domaine « .fr » n'arrivent ainsi « qu'en » 4ème place dans environ 5% des cas. Au total, près de 3 500 URLs frauduleuses étaient hébergées sur une extension nationale (ccTLD[2]) contre environ 8 250 sur des extensions génériques. 140 extensions nationales différentes ont néanmoins été identifiées en 1 an, mais seules une vingtaine d'entre elles hébergeaient plus d'1 cas par mois.

Adresse IP : localisation géographique des sites frauduleux

L'hébergement des sites de phishing est concentré au sein d'un petit nombre de pays. Près de la moitié des adresses IP répondant aux sites de phishing étaient localisées aux Etats-Unis, contre 1 sur 10 pour la France. Les 3 premiers pays hébergeaient ensemble près des deux tiers des sites et les 20 premiers plus de 90% des cas.

Entre mars et décembre (soit 9 mois pour lesquels nous disposons de l'information), les adresses IP correspondant aux enregistrements DNS de type A des URLs de phishing validées étaient réparties géographiquement de la façon suivante :

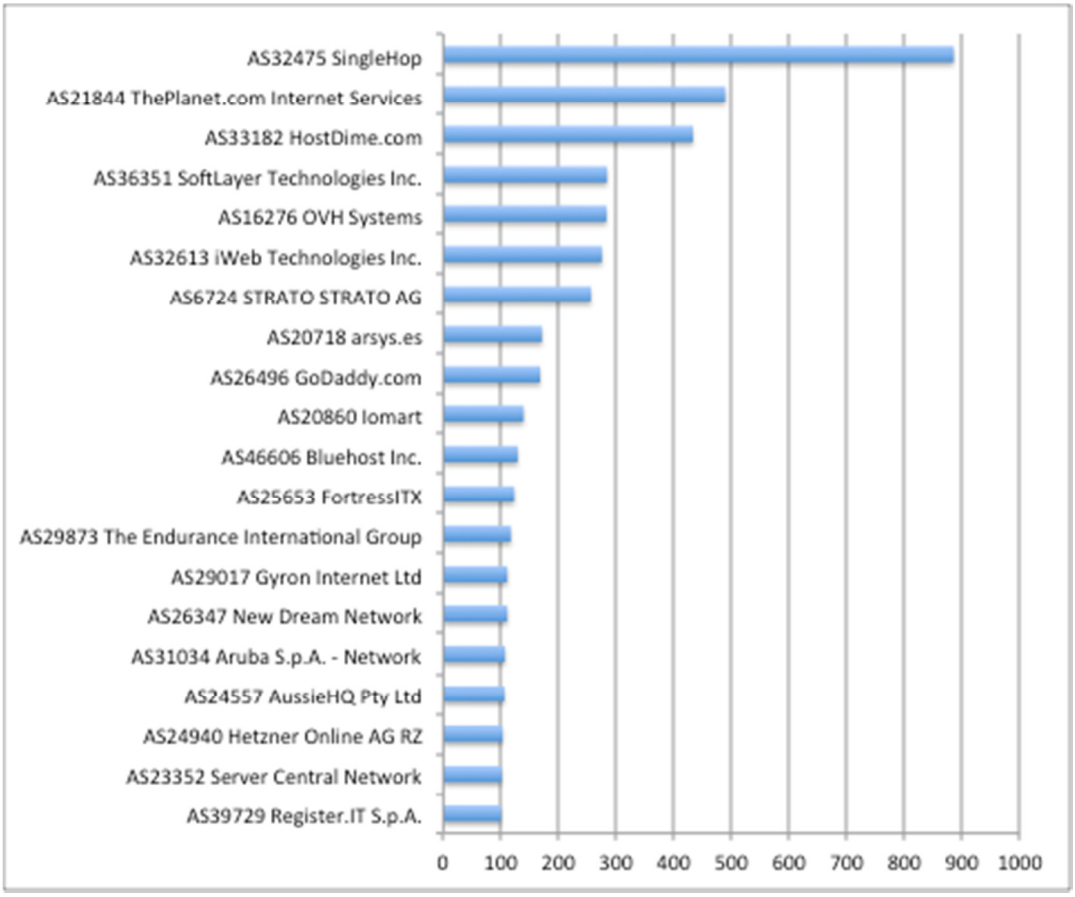
[2] ccTLD : country-code Top Level Domain



20 principaux pays associés aux adresses IP d'après le nombre d'URLs de phishing

Adresse IP : répartition par détenteur des adresses IP (volume brut)

Sur la même période de 9 mois (3-12/2011), les détenteurs des adresses IP hébergeant les contenus frauduleux ont été analysés d'après la correspondance de l'adresse IP identifiée avec son numéro d'AS. Le Top 20 des prestataires ainsi recensés se présente comme suit :



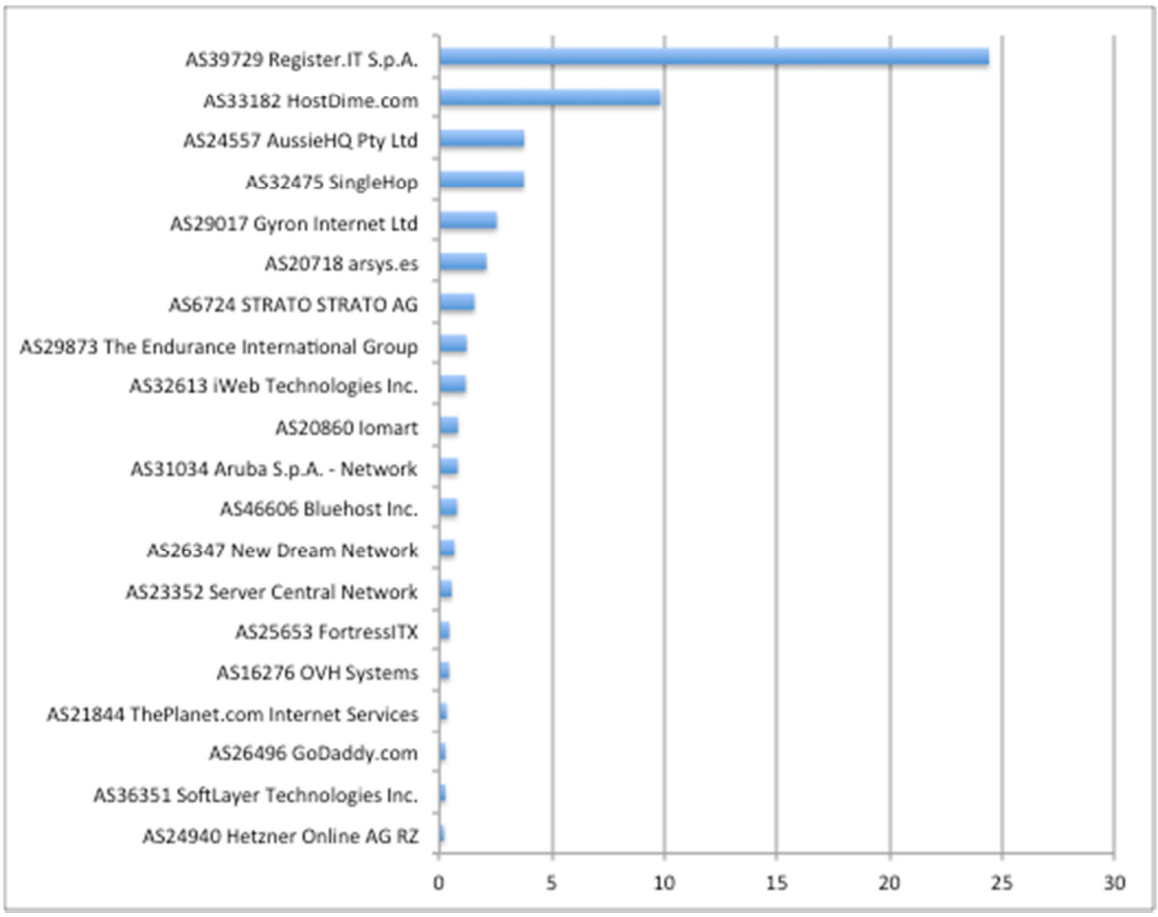
Top20 des hébergeurs d'URLs de phishing (nombre de cas)

L'hébergeur américain SingleHop détenait près de 10% des adresses IP associées aux URLs de phishing validés recensées. Les 10 premiers AS ont hébergé plus du tiers des cas de phishing, les prestataires du top 20 (ci-dessus) comptant pour près de la moitié des cas. Les 100 premiers AS possédaient quant à eux 80% des adresses IP identifiées.

Adresse IP : répartition par détenteur des adresses IP (d'après leur taille)

Nous avons dans un premier temps calculé le nombre d'adresses IP associées aux cas de phishing pour les principaux AS identifiés ci-dessus (Top20) d'après leur taille respective. Pour ce faire, le nombre d'adresses IP annoncées par chaque AS a été pris en compte[3].

[3] NDLR : le nombre d'adresses IP annoncées a été déterminé à un instant « t » (9 mars 2012, 17h GMT), et non à la date d'identification du phishing et de son IP associée. Ce classement « à postériori » est donc à titre informatif et pas forcément parfaitement représentatif, puisque les annonces de classes d'adresses IP (nombre et même détenteur effectif) évoluent constamment.



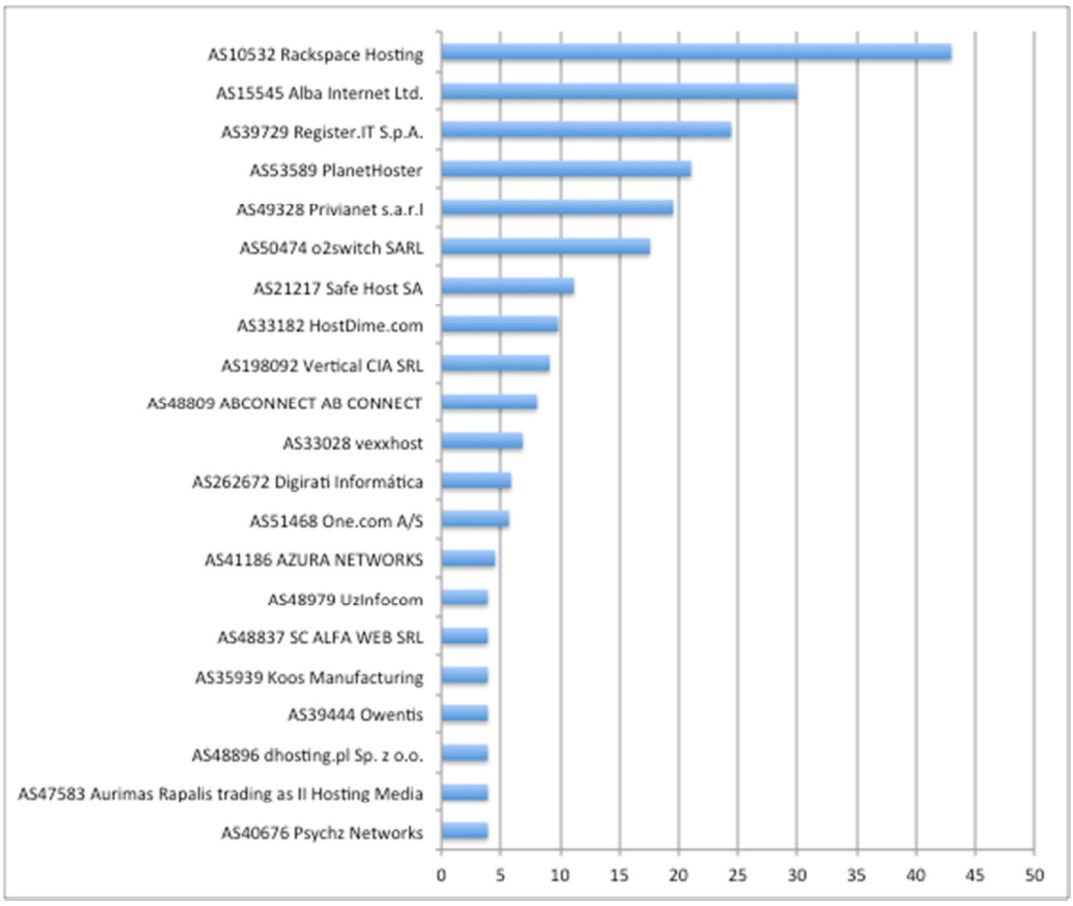
Top20 des hébergeurs d'URLs de phishing selon leur taille relative

(nombre de cas de phishing pour 1000 adresses IP annoncées)

Le classement est de fait très différent, Register IT passant de la 20ème à la première place en nombre de phishing hébergés par millier d'adresses IP annoncées, devant HostDime. SingleHop rétrograde quant à lui à la quatrième place de ce Top20.

Ensuite, l'ensemble des 730 détenteurs différents a été pris en considération pour le calcul par taille relative, soit y compris pour ceux qui n'ont hébergé qu'un seul cas de phishing sur la période. Au total plus de 730 numéros d'AS distincts ont ainsi été identifiés, mais plus du tiers n'ont été concernés que par 1 seule URL.

Le classement s'établit dès lors de la sorte :



Classement de tous les hébergeurs d'URLs de phishing selon leur taille relative

(nombre de cas de phishing pour 1000 adresses IP annoncées)

La moitié des ces organisations sont des hébergeurs modestes ou petits, annonçant moins de 1000 adresses IPs. De fait, elles sont plus susceptibles d'entrer dans ce classement dès lors que quelques cas de phishing sont identifiés sur l'une de leurs adresses IP.

Seuls Register IT et HostDime se retrouvent à la fois dans le Top20 en volume brut de nombre de cas hébergés (respectivement 100 et 433 URLs) et par taille relative (respectivement 24 et 10 cas pour 1000 IPs annoncées).

On peut ainsi dire que le ratio de phishing hébergé a tendance à baisser avec la taille des hébergeurs. Ce fait est probablement en partie lié aux moyens plus importants pouvant être mis en œuvre pour empêcher des sites de phishing de voir le jour par les principaux hébergeurs (vérifications de l'identité lors de la souscription d'un espace d'hébergement, surveillance des contenus et des accès, mises à jour applicatives des serveurs, etc.).

La communauté Phishing-Initiative

Un nombre croissant de contributeurs

Rappel

Le nombre de contributeurs est une estimation puisqu'il provient des adresses e-mail renseignées lors des soumissions. Or il s'agit d'une donnée qui n'est pas vérifiable car n'importe quelle adresse peut être fournie. Nous pensons qu'une très large majorité des contributeurs renseigne cependant une adresse e-mail personnelle existante.

Volume de contributeurs

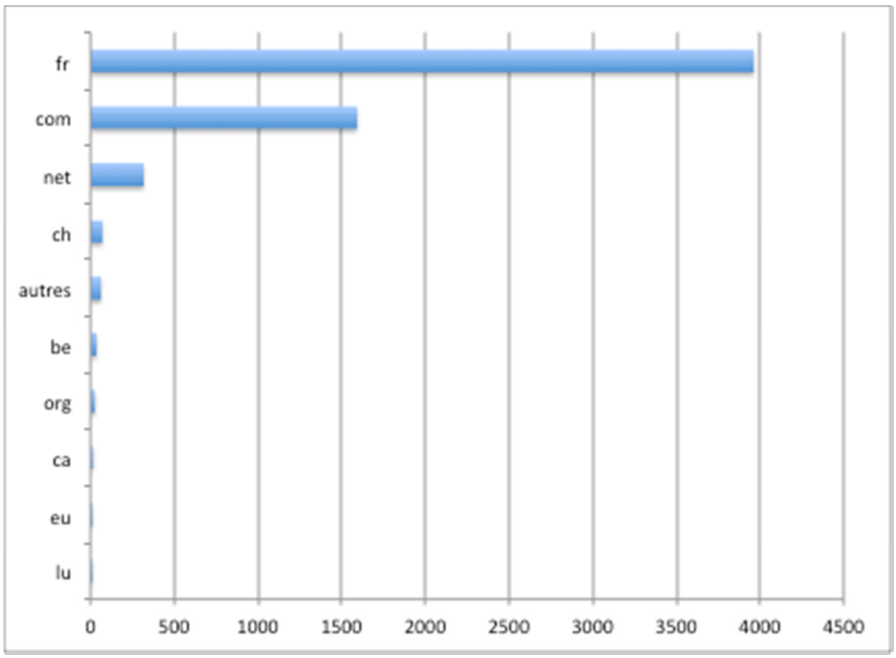
Plus de 6 000 contributeurs différents ont participé au signalement des cas suspects. La moyenne se situe donc à environ 5 soumissions par an par personne, mais cache de fortes disparités. En effet, plus des deux tiers des individus n'ont remonté qu'un seul cas à ce stade.

Le nombre de contributeurs ne représente qu'environ 1 Français sur 10 000 à l'heure actuelle. Il est cependant utile de rappeler que dès lors qu'un site est remonté et confirmé, les actions mises en œuvre au niveau des navigateurs partenaires permettent de protéger l'ensemble des autres Internautes visitant ce site frauduleux.

Donc il est nécessaire pour que Phishing-Initiative soit efficace que l'information nous parvienne le plus tôt possible. Obtenir par la suite cette même information par un grand nombre d'autres internautes permet en plus de dresser des tendances et analyser les volumétries des campagnes d'attaques elles-mêmes.

Extensions de domaine des adresses e-mail des contributeurs (informatif)

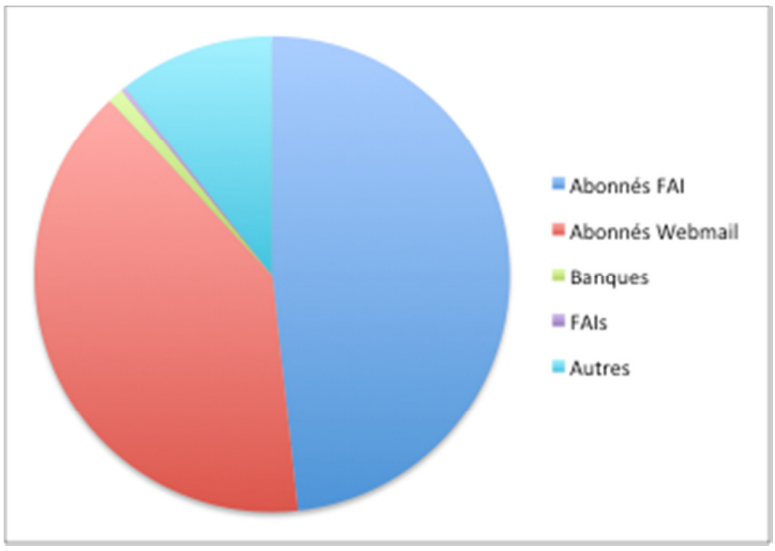
Les contributeurs disposent en majorité d'adresses e-mail comprenant un domaine dans l'extension nationale française, puisque 2/3 de celles-ci possédaient une extension de domaine en « .fr » :



Extension de domaine (TLD) des adresses e-mail des contributeurs

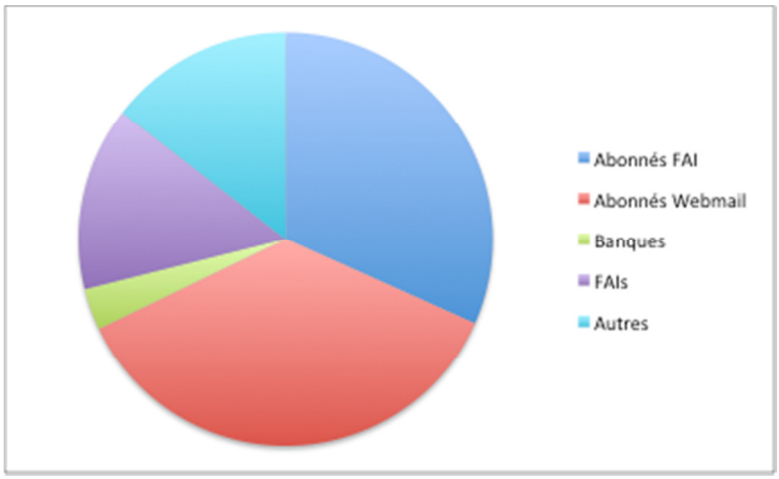
Typologie de contributeurs (informatif)

Les abonnés des FAIs français et les détenteurs de compte e-mail gratuit en ligne (Webmail) forment ensemble plus de 90% des participants.



Typologie des contributeurs (estimation)

Mais une vingtaine d'adresses e-mail appartenant à des employées d'organisations fortement ciblées (banques et surtout FAIs) ont à eux seuls contribué pour près de 20% des signalements.



Nombre de soumissions par typologie de contributeurs

Des partenaires mobilisés

Il faut en moyenne un peu moins de 13 minutes pour qu'un site soit étudié et classifié par les analystes en charge de la vérification. 54% des soumissions ont été traitées dans les 15 minutes et plus de 70% dans les 30 minutes suivant leur réception.

Microsoft (et Google) récupèrent toutes les 5 minutes maximum les nouvelles adresses de sites frauduleux à ajouter à leurs listes noires, ce qui permet de rapidement protéger l'ensemble des utilisateurs finaux. Le blocage des sites de phishing est notamment activé par défaut dans les versions 8 et suivantes des navigateurs Internet Explorer.

Plusieurs centaines de questions auxquelles nous avons essayé de répondre ont par ailleurs été reçues sur nos adresses de contact. Les interrogations récurrentes de ces contributeurs portent surtout sur les moyens à même de faire cesser l'arrivée de ces escroqueries dans leurs boîtes aux lettres.

La forme d'incompréhension la plus courante porte sur les informations à remonter. En effet, un grand nombre d'internautes nous écrivent pour signaler l'adresse e-mail d'expédition (c'est à dire le champ « De : » ou « From : », en lieu et place de l'adresse de la page Web frauduleuse destinée à voler les informations). Or, celle-ci est intégralement configurable par l'attaquant, qui use de ce stratagème pour usurper le nom et l'identité des organisations utilisées pour crédibiliser son attaque. Cette information ne représente donc pas un élément utile à notre action.

Le phishing francophone : un phénomène massif

Evolution annuelle

Nous n'avons à ce stade pas de visibilité sur l'évolution du phishing en France ces dernières années. La tendance est fortement à la hausse selon nos partenaires du Cert-Lexsi, qui a enregistré un quasi doublement entre 2010 et 2011 des cas visant un public francophone.

Les statistiques de Phishing-Initiative fin 2012 permettront de déterminer une tendance et confirmer le cas échéant l'évolution constatée par le Cert-Lexsi.

Hypothèses d'estimation du phénomène

Si nous ajoutons aux plus de 12 500 cas confirmés via Phishing-Initiative :

- Quelques milliers d'URLs impossibles à confirmer, car inaccessibles au moment de notre vérification,
- Près de 6 000 autres URLs visant un public francophone identifiées par le Cert-Lexsi, ses clients ou partenaires, (source : Cert-Lexsi),

il est probable que plus de 20 000 URLs de phishing visant un public francophone ont a priori existé en 2011, soit plus de 50 par jour !

Coefficients réducteur ou multiplicateur de l'estimation

Néanmoins, le nombre d'URLs frauduleuses est différent du nombre d'attaques distinctes menées. Il est donc nécessaire d'affecter au nombre d'URLs confirmées certains facteurs augmentant ou diminuant l'estimation du nombre de cas uniques.

D'un côté, l'estimation ne prend par exemple en compte que les cas de phishing :

- identifiés par Phishing-Initiative et Cert-Lexsi,
- impliquant l'existence d'un site frauduleux, et non :
    - les tentatives de vol d'informations par réponse directe à un e-mail (visant notamment les identifiants de comptes de messagerie Hotmail, Gmail, etc.)
    - les autres formes de phishing avec récupération des informations :
        - par compromission de l'ordinateur (pharming, man-in-the-middle, etc.)
        - par téléphone (vishing),
        - par SMS (smishing),
        - via la messagerie instantanée ou les réseaux sociaux,
    - etc.

De plus, une nouvelle campagne de spams faisant la promotion plusieurs jours plus tard d'une même URL précédemment identifiée ne sera pas comptabilisée une nouvelle fois du côté de Phishing-Initiative.

D'un autre côté, plusieurs URLs sont parfois détectées et remontées à Phishing-Initiative pour une même campagne de phishing :

- adresses de redirection automatique,
- multiples pages comprenant le formulaire où renseigner les données,
- URL différente pour la récupération des données volées,
- etc.

Règle de calcul améliorée

Postulat de base :

Nombre d'URLs confirmées par différentes sources fiables (Phishing-Initiative, Cert-Lexsi, etc.)

+ Ajout liés aux :

- attaques non remontées donc non détectées
- attaques suspectées, mais non confirmées
- autres formes d'attaques hors périmètre (vishing, smishing, etc.)
- réapparitions d'une attaque sur une même URL (site remis en ligne)
- autres facteurs

- Diminution du fait des :

- attaque comprenant une ou plusieurs URLs de redirection vers un site final
- attaque comprenant différentes URLs pour le site et la récupération des données volées
- attaque visant un public non uniquement francophone, mais inclus dans le décompte
- campagne de spam unique, mais diffusant des URLs différentes
- autres facteurs

= Nombre d'attaques de phishing contre les internautes francophones.

Plus de recherches sont nécessaires pour déterminer une potentielle pour ces différents facteurs, et ainsi obtenir une vue plus exacte du phishing visant le public francophone.

On peut néanmoins estimer que le nombre d'attaques perpétrées contre les Internautes francophones en 2011 est considérable.

Conclusion

Plus l'initiative sera connue du public

Plus nombreuses seront les personnes contribuant sur le site

Plus tôt les adresses de phishing suspectées seront reçues et analysées

Plus vite elles seront transmises aux différents partenaires pour action et blocage

Plus complète sera la protection des internautes francophones contre les attaques de phishing

Microsoft, PayPal et Lexsi demeurent ainsi mobilisés pour poursuivre la mission engagée par Phishing-Initiative.

Nous souhaitons désormais accueillir en 2012 de nouveaux membres issus du secteur public et privé, et souhaitant participer à la lutte contre le phishing. Cette ouverture contribuera à pérenniser le projet en France voire à le développer à terme à l'étranger.