Le phishing et son principal vecteur de diffusion le spam

Sommaire

2.1 Introduction au phishing . . . 14
2.1.1 Premier maillon : le site web contrefait . . . 14
2.1.2 Second maillon : la campagne de communication . . . 17
2.2 Anatomie d'un site de phishing . . . 19
2.2.1 Zoom sur l'URL . . . 19
2.2.1.1 Structure d'une URL . . . 19
2.2.1.2 Caractéristiques d'une URL de phishing . . . 20
2.2.2 Zoom sur la page web . . . 22
2.2.2.1 Synopsis d'une page web . . . 22
2.2.2.2 Caractéristiques d'une page de phishing . . . 23
2.3 Mise en œuvre du phishing . . . 25
2.4 Méthodes de détection/protection existantes côté client . . . 26
2.4.1 Le filtre idéal . . . 28
2.4.2 Au niveau des emails . . . 28
2.4.3 Au niveau du navigateur web . . . 29
2.4.3.1 Sécurisation de la connexion client-serveur avec le protocole HTTPS . . . 29
2.4.3.2 Alternatives pour la saisie des login/mot de passe . . . 30
2.4.3.3 Détection des signatures de pages webs contrefaites . . . 30
2.4.3.4 Barres d'outils anti-phishing . . . 30
2.5 Quelques alternatives côté réseau FAI/serveur web . . . 31
2.6 Synthèse du chapitre . . . 32

Difficile d'aborder la problématique du phishing sans évoquer son principal moyen de diffusion : le spam. En effet, les attaques de phishing sont majoritairement propagées/diffusées au travers de campagnes d'emails non sollicités. Les techniques de détection et de prévention associées s'en retrouvent donc souvent étroitement liées.

Dans ce chapitre, nous nous intéressons particulièrement au phishing. Nous démarrons en section 2.1 par une introduction à la problématique. En section 2.2, nous détaillons les propriétés d'un site de phishing. Puis, en section 2.3 nous discutons des vecteurs de propagations de ces attaques. Enfin, en section 2.4, nous traitons des différents travaux/méthodes de détection/protection qui s'y rapportent côté client. En complément, la section 2.5 discute des quelques mesures possibles côté réseau FAI/serveur web.

Ce chapitre fait partie de nos contributions : un Dossier Technique portant sur le spam a été publié dans la Collection *Sécurité des Systèmes d'Information* des *Editions Techniques de L'Ingénieur*, en Avril 2009 [Gas09].

2.1 Introduction au phishing

Avant de parler du phishing, il est nécessaire d'introduire le spam auquel il est étroitement lié. Un spam est un message électronique non sollicité, envoyé massivement à de nombreux destinataires, à des fins publicitaires ou malveillantes. L'ampleur du phénomène est conséquente puisque le spam représente aujourd'hui entre 88 et 91%[1] du volume total d'emails échangés [MAA11]. Les objectifs du spam sont divers et variés. Il sert notamment à propager[2] des publicités (p.ex. pour des produits contrefaits, pharmaceutiques), des canulars (c.-à-d. des fausses alertes à la population, des chaînes de solidarité censées prodiguer chance et bonheur, etc.), des logiciels malveillants : virus, chevaux de troie, etc. (p.ex. pour utiliser les postes d'Internautes comme membres d'un botnet[3]), des scams (qui visent à soutirer de l'argent grâce au leurre d'une forte rétribution financière à venir), ou encore des attaques de phishing.

Une attaque de phishing consiste à voler des informations confidentielles (p.ex. identifiant, mot de passe, numéro de carte bancaire, etc.) aux Internautes en usurpant l'identité de sites marchands et/ou bancaires. Pour ce faire, elle s'appuie sur deux maillons essentiels : la mise en ligne d'un site web contrefait qui usurpe l'identité d'un site légitime (p.ex. une banque, une plate-forme de jeux en ligne, etc.), et la mise en place d'une campagne de communication (typiquement basée sur du spam) afin d'attirer les Internautes.

2.1.1 Premier maillon : le site web contrefait

Techniquement, la mise en ligne du site contrefait s'appuie sur une page web d'accueil d'aspect (plus ou moins) ressemblant au site légitime. Meilleure sera la ressemblance et plus l'attaque sera efficace. Cette page web est accédée grâce à une URL également (plus ou moins) ressemblante à l'URL légitime. A nouveau, le degré de ressemblance de cette URL peut ajouter à l'efficacité de l'attaque. Il est toutefois indéniable que le rendu visuel global de la page sera plus décisif.

Pour illustrer ces propos, considérons deux exemples qui ciblent le site Facebook : les figures 2.1 et 2.2 montrent une première illustration de sites légitime/contrefait, tandis que les figures 2.3 et 2.4 en montrent une deuxième.

Le site contrefait développé dans le premier exemple est de prime abord, celui qui imite au mieux son pendant légitime. En effet, on ne détecte aucune erreur syntaxique, aucune modification de taille/type de polices de caractères, etc.). L'imitation apparaît parfaite. A contrario le deuxième exemple semble moins soigné, au sens visuel du terme. En effet, les zones 2 et 3 de la figure 2.4 démontrent des problèmes syntaxiques (c.-à-d. un manque de prise en charge des caractères spéciaux et accentués) ou des modifications de contenu (p.ex. on aperçoit que le site web contrefait est hébergé chez `T35 Hosting`, qui introduit des références publicitaires vers d'autres sites).

En ce qui concerne les liens accessibles depuis la page contrefaite, les zones 2 du premier exemple conduisent vers le site légitime. A contrario, les zones 3 et 4 de ce même exemple contiennent des liens contrefaits (cf. figure 2.2). Dans le deuxième exemple, tous les liens des zones 2 et 4 redirigent vers le site légitime (cf. figure 2.4). Seuls les liens amenés par l'hébergeur de site (cf. zone 3 de la figure 2.4) redirigent ailleurs (p.ex. `Free Domains` conduit à `http://www.domainsfree.org/`, `T35` redirige vers `http://www.t35.com`).

Enfin, si on s'intéresse à l'URL visitée (Zone 1 sur les figures 2.2 et 2.4), on voit que le deuxième exemple est cette fois-ci plus performant que le premier. En effet, l'URL contrefaite s'approche davantage de l'URL originale, en introduisant le nom de domaine légitime (c.-à-d. Facebook).

1. Ces chiffres - édités par le MAAWG (Message Anti-Abuse Working Group) - sont en provenance directe des FAIs. Ils portent sur plus de 500 millions de boîtes emails.
2. Liste non exhaustive.
3. ici, un botnet désigne un réseau de machines corrompues (appelées *bots* ou zombies) au travers d'Internet, utilisées comme rebond pour perpétrer des attaques.

FIGURE 2.1 – Zoom sur le site légitime italien de Facebook `http://it-it.facebook.com/`

FIGURE 2.2 – Zoom sur le site contrefait `http://genplus.altervista.org/` qui usurpe le site légitime Facebook illustré en figure 2.1

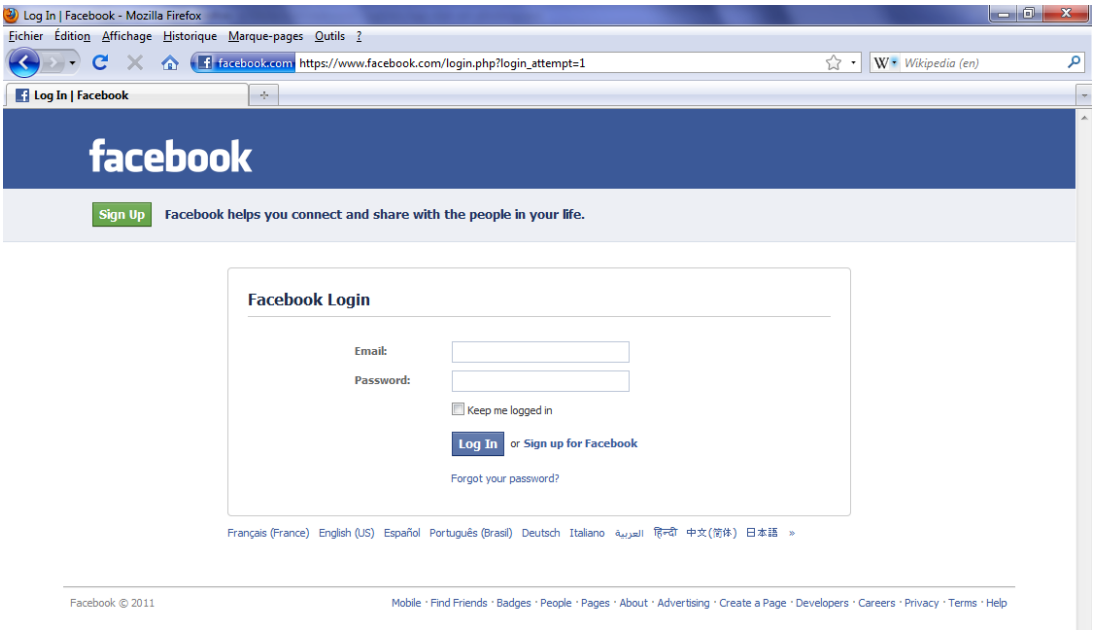


FIGURE 2.3 – Zoom sur le site légitime américain de Facebook `https://www.facebook.com/login.php?login_attempt=1`

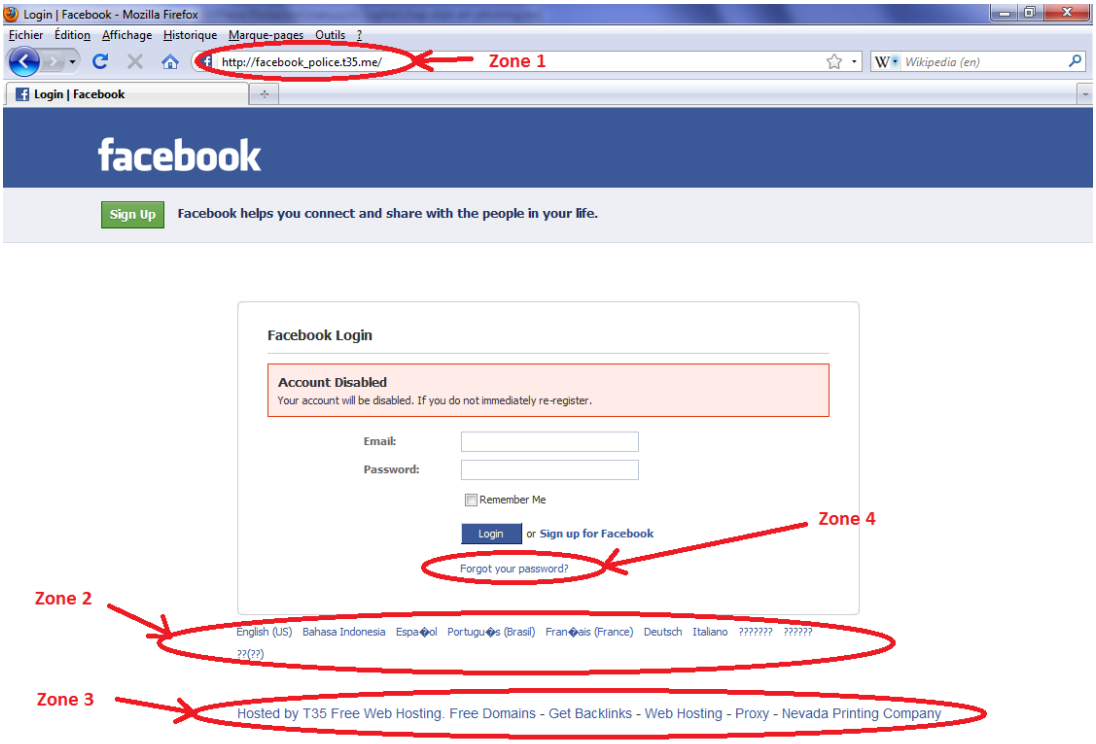


FIGURE 2.4 – Zoom sur le site contrefait `http://facebook_police.t35.me/` qui usurpe le site légitime Facebook illustré en figure 2.3

2.1.2 Second maillon : la campagne de communication

L'alliance d'un site web contrefait et d'une URL frauduleuse ne peut être suffisante pour attirer les victimes. En effet, les Internautes n'ont connaissance que de l'URL légitime. Il faut alors trouver un moyen de les amener à visiter l'URL frauduleuse. Pour ce faire, diverses techniques - exposées en section 2.3 - sont utilisables. Néanmoins l'une d'entre elles prédomine : la campagne de spam.

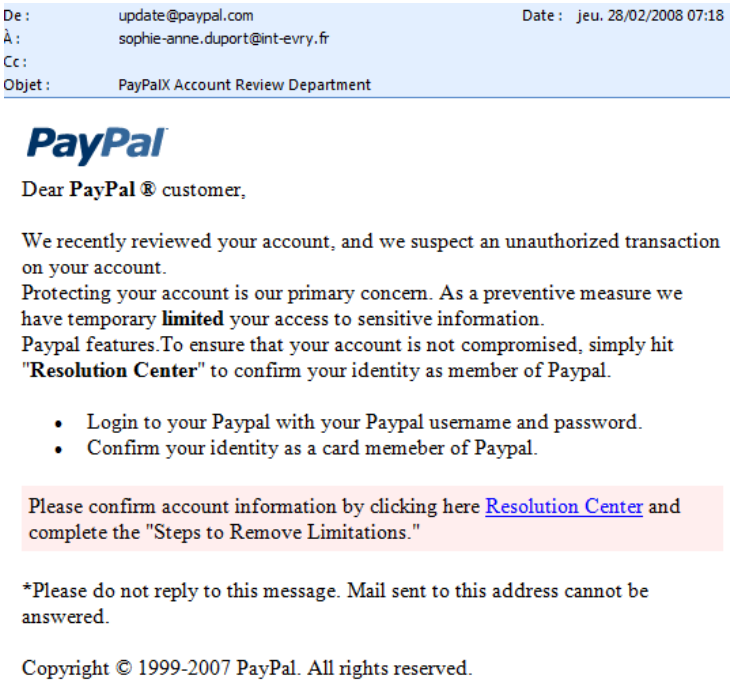
Ces spams qui servent à véhiculer les attaques de phishing sont eux aussi plus ou moins soignés. Prenons trois exemples :

- La figure 2.5 illustre un spam de haute qualité qui usurpe l'identité du site Paypal (`https://www.paypal.com`). Cet email ne comporte en effet que de minimes erreurs que bon nombre d'utilisateurs pourraient manquer. La majorité des indicateurs plaident en effet en faveur de la légitimité du site : le logo, l'émetteur affiché (c.-à-d. `update@paypal.com`), l'email destinataire qui apparaît valide, et l'URL frauduleuse utilisée qui - même si elle est masquée derrière l'élément `Resolution Center` - semblent faire référence au domaine légitime `http://paypal-secure-login.com/acc/login.php`. Seuls éléments d'alerte : un caractère en trop dans le titre du message (c.-à-d. le `X` après `Paypal`), ou l'URL de redirection qui utilise le protocole HTTP (et non HTTPS - pour *Hyper-Text Transfer Protocol Secure* -). De plus, l'email destinataire n'est pas celui du destinataire réel (bien qu'appartenant au même domaine) et enfin, en examinant le contenu de l'en-tête SMTP (pour *Simple Mail Transfer Protocol*) de l'email, on constate que l'émetteur réel est `akstcabilitamnsdgs@abilita.com`.

- La figure 2.6 illustre quant à elle un spam de qualité intermédiaire qui usurpe le Crédit Mutuel (`https://www.creditmutuel.fr/`). Les éléments qui plaident en faveur de la réussite de l'attaque sont : le logo légitime, des mentions de Copyright, un email émetteur affiché qui semble légitime même au sein de l'en-tête SMTP (c.-à-d. `service@creditmutuel.fr`), l'email destinataire ciblé et correct (c.-à-d. cohérent avec la boîte de réception où il est délivré), le lien de redirection masqué (derrière l'élément `Cliquez ici`) et enfin, un caractère d'urgence (grâce à l'email envoyé en priorité haute, et la menace explicite de résiliation de la carte bancaire sous 6 jours). A contrario, si on y regarde de plus près, le contenu du message est impersonnel et truffé de fautes. De plus, le lien vers lequel le client est emmené est `http://user33283.vs.easily.co.uk/credit/metuel/confirmation/compte/suspenstion/carte/bancaire/reconfirmation/informations/personelle/login.aspx/compte/`, sans rapport aucun avec le domaine légitime même s'il tente d'y faire vaguement référence dans le début de l'arborescence.

- Enfin, la figure 2.7 illustre un spam de phishing de basse qualité qui usurpe l'identité du site Paypal (`https://www.paypal.com`). En effet, l'email émetteur est imprécis et ne contient aucune référence au domaine légitime (c.-à-d. `carte de crXdit [systeme@security.net]`), l'email destinataire est impersonnel car masqué (`undisclosed-recipients`), le titre et le contenu du message sont truffés de fautes, ou encore l'URL de redirection est insuffisamment masquée : l'affichage indique en effet `definitefraudstart"www.artifizbox.com"definitefraudendhttps://www.paypal.com/fr/cgi-bin/webscr?cmd=_login-submit` pour une URL visitée `http://www.artifizbox.com/www.paypal.fr/cgi-bin/webscrcmd=_login-run/webscrcmd=_account-run/updates-paypal/confirm_paypal/`. Seul élément qui pourrait plaider en faveur de l'attaque : la mention d'un numéro de référence Paypal `PP-538-718-203`.

Pour terminer, notons également qu'un élément souvent utilisé par les attaques de phishing - afin d'accroître leur efficacité - est de limiter le temps de réflexion des Internautes. Pour ce faire, l'attaque introduit alors un caractère d'urgence avec lequel l'utilisateur doit réagir. Cette notion d'urgence peut aussi bien se trouver dans le spam (p.ex. dans la figure 2.6 où on voit la mention *Note :Si ce n'est pas achever le 10 Avril 2011, nous serons contraints de suspendre votre carte indéfiniment, car elle peut être utilisée pour des raisons frauduleuses...*) que dans le site web contrefait (p.ex. dans la figure 2.4 où on voit la mention *Account Disabled : Your account will be disabled. If you do not immediately re-register*).



De : update@paypal.com Date : jeu. 28/02/2008 07:18
À : sophie-anne.duport@int-evry.fr
Cc :
Objet : PayPalX Account Review Department

PayPal

Dear **PayPal ®** customer,

We recently reviewed your account, and we suspect an unauthorized transaction on your account.
Protecting your account is our primary concern. As a preventive measure we have temporary **limited** your access to sensitive information.
Paypal features.To ensure that your account is not compromised, simply hit "**Resolution Center**" to confirm your identity as member of Paypal.

- Login to your Paypal with your Paypal username and password.
- Confirm your identity as a card memeber of Paypal.

Please confirm account information by clicking here Resolution Center and complete the "Steps to Remove Limitations."

*Please do not reply to this message. Mail sent to this address cannot be answered.

Copyright © 1999-2007 PayPal. All rights reserved.

FIGURE 2.5 – Premier exemple de spam qui véhicule une attaque de phishing en usurpant Paypal

De : Crédit Mutuel [service@creditmutuel.fr] Date : lun. 04/04/2011 08:28
À : Sophie.Gastellier@it-sudparis.eu
Cc :
Objet : Votre Carte Bancaire a été suspendue

Crédit Mutuel
LA banque à qui parler

Bonjour client de Crédit Mutuel,

Votre Carte Bancaire a été suspendue .nous avons remarquer un problème sur votreCarte.

Nous avons déterminer que quelqu'un a peut-être utiliser Votre Carte sans votre autorisation. Pour votre protection, nous avons suspendue votre Carte de crédit.Pour lever cette suspension,**Cliquez ici**et suivez la procédure indiquerpourMettre a jour de votre Carte Crédit.

Note: Si ce n'est pas achever le 10 Avril 2011, nous serons contraints de suspendre votre carte indéfiniment, car elle peut être utilisée pour des raisons frauduleuses...

Nous vous remercions de votre coopération dans le cadre de ce dossier.

Merci,
Support Clients Service.

Copyright 1998-2011 Crédit Mutuel . Tous droits réserves.

FIGURE 2.6 – Deuxième exemple de spam qui véhicule une attaque de phishing en usurpant le Crédit Mutuel

Ce message a été envoyé avec une importance Haute.
De : carte de crXdit [systeme@security.net] Date : sam. 21/03/2009 13:38
À : undisclosed-recipients:
Cc :
Objet : Attention : Service-Veuillez retablire l'acess de votre compte PayPal .

Dans le cadre de nos mesures des securite, nous controlons regulierement les Activites
Encours dans le systeme PayPal. Au cours d'une recente verification,
Nous avons releve une probleme sur votre compte PayPal.

En etudiant votre compte, nous sommes rendu compte que nous avions besoin
D'informations supplementaires pour vous fournir un service securise.

definitefraudstart "www.artifizbox.com"definitefraudend https://www.paypal.com/fr/cgi-bin/webscr?cmd=_login-submit

Numero de reference : PP-538-718-203

PayPal

PayPal de sécurité et de la lutte anti-fraude Département.

FIGURE 2.7 – Troisième exemple de spam qui véhicule une attaque de phishing en usurpant Paypal

2.2 Anatomie d'un site de phishing

Un site web est un ensemble de pages webs, accédées au travers des URLs qui leur sont attachées. Analyser les caractéristiques d'un site de phishing - tel que développé au travers de plusieurs travaux précédents [PKKG10] [GPCR07] [PD06] [AHDT10] - revient donc à étudier les spécificités des URLs et/ou l'aspect/contenu de la page web contrefaite associée. Précisons en effet qu'un site web de phishing se résume en général à l'utilisation d'une unique page[1] contrefaite, qui conserve et fait appel à un maximum de redirections de liens du site légitime, typiquement pour l'accès à des informations complémentaires (p.ex. pour les pages d'aide).

2.2.1 Zoom sur l'URL

2.2.1.1 Structure d'une URL

Une URL est généralement composée des éléments suivants (cf. figure 2.8) :

1. Le protocole utilisé (p.ex. *http* pour une page web classique, *https* pour une page web de login, *ftp* pour un serveur de téléchargement, etc.), suivi des caractères *://* qui précèdent la désignation de l'emplacement de stockage de la ressource demandée.
2. Un nom d'hôte complet, appelé FQDN (pour *Fully Qualified Domain Name*) qui précise le nom de la machine qui héberge la ressource demandée. Ce FQDN est lui-même traditionnellement composé de 3 éléments : un nom d'hôte (p.ex. *www* pour un serveur web), un nom de domaine (p.ex. *yahoo*) et un TLD - pour *Top-Level Domain* - (p.ex. *fr*) qui est utilisé pour structurer/hiérarchiser la gestion des noms de domaines. Ce TLD est généralement représentatif d'une localisation géographique ou d'un type de site web (p.ex. EU pour Europe, .FR pour France, .COM pour des sites commerciaux, etc.). De plus amples détails sur la gestion des TLDs et des domaines sont disponibles en section 4.1.1.
3. Le chemin d'accès au sein de l'hôte spécifié, c.-à-d. l'arborescence de répertoires utilisés pour le stockage de la ressource, séparés par des /.
4. Le nom du fichier qui contient la ressource.

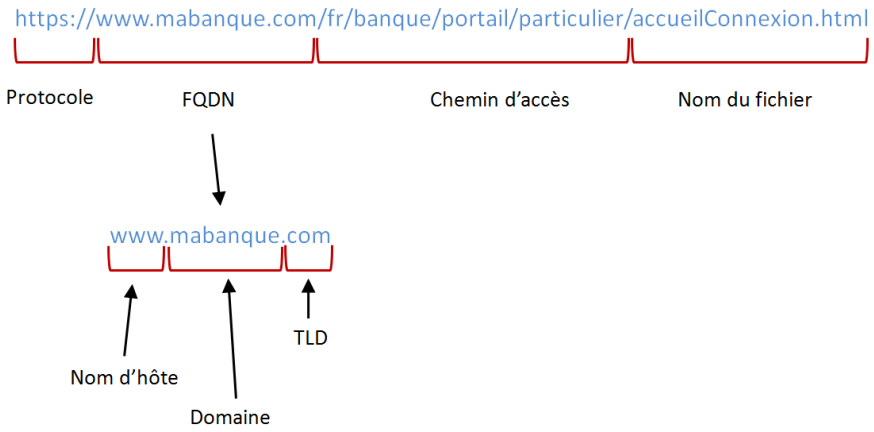


FIGURE 2.8 – Vue simplifiée de la décomposition d'une URL

Notons que le chemin d'accès et/ou le nom de fichier sont des éléments optionnels. Si l'hôte de destination a pré-configuré une page par défaut, il n'est pas nécessaire de spécifier ces éléments.

Précisons également que le schéma de décomposition d'une URL présenté ici est une version que l'on peut qualifier de "classique" et/ou simplifiée. Par exemple, la zone dite *nom d'hôte complet* est en réalité

1. ou un nombre très restreint de pages.

un peu plus complexe. En effet, sa syntaxe complète est : `<user>:<password>@<host>:<port>` [IET94], On y remarque la présence d'un caractère spécial : l'arobas(@)[1]. Cette syntaxe complète permet de spécifier des données d'authentification (c.-à-d. un nom d'utilisateur (`user`) et son mot de passe (`password`)) pour la connexion à un hôte (`host`) via un port destination choisi (`port`). Toutefois, seule la mention d'hôte est obligatoire. D'où la simplification courante de cette zone en *nom d'hôte complet*, puisqu'on y retrouve typiquement mention d'un FQDN. Néanmoins, d'autres variantes existent. On peut par exemple trouver une adresse IP en lieu et place du FQDN, ou une indication d'un numéro de port alternatif (c.-à-d. si le port destination est différent de 80 pour du HTTP, 443 pour du HTTPS, etc.).

URI, URL et URN : D'après le W3C (pour *World Wide Web Consortium*), URL, URI et URN sont trois notions différentes qu'il convient de distinguer [W3C01]. Un URI (pour *Uniform Resource Identifier*) est un terme global qui permet de désigner une ressource de façon unique. Un URI peut être une URL et/ou un URN.

Une URL (pour *Uniform Resource Locator*) désigne une ressource ainsi que sa localisation complète sur Internet et le moyen d'y accéder (c.-à-d. le protocole utilisé pour atteindre la ressource). Une URL est généralement utilisée pour la navigation web, l'accès aux mails, etc.

Un URN (pour *Uniform Resource Name*) est un identifiant permanent d'une ressource (p.ex. un numéro ISBN pour un livre), totalement décorrélé d'une quelconque notion de disponibilité.

2.2.1.2 Caractéristiques d'une URL de phishing

Plusieurs travaux précédents [PKKG10] [ZHC07] [GPCR07] [PD06] [CG06] [MG08] et notre étude approfondie des URLs de phishing font ressortir l'éventail des techniques utilisées par les attaquants afin de leurrer les utilisateurs. On peut notamment citer :

- **La substitution du FQDN par une adresse IP** : Certaines URLs de phishing utilisent une adresse IP en lieu et place d'un FQDN, ceci afin que le changement de FQDN - sans rapport aucun avec celui du site légitime - apparaisse moins visible. Par exemple, l'URL `http://74.220.215.65` est utilisée comme alternative au site contrefait `http://volleyballplayerz.com/` qui usurpe l'identité de la banque Natwest `http://www.natwest.com/`.

- **La déformation du domaine/FQDN légitime** : Il est très fréquent de rencontrer des URLs de phishing dont le FQDN est construit à partir d'une version déformée du nom de domaine ou FQDN (p.ex. via l'ajout, la modification ou le remplacement d'un - ou quelques - caractère(s)). Par exemple, l'URL de phishing `http://www.bhttle.net/` usurpe l'URL légitime `http://www.battle.net`, et l'URL contrefaite `http://faseboo.altervista.org/` usurpe l'URL légitime `http://www.facebook.com`.

- **L'utilisation de FQDNs/URLs longs** : Il est très fréquent de trouver des URLs de phishing constituées de FQDNs et/ou URLs à rallonge (p.ex. via l'utilisation de nombreux points (.)). Les URLs contrefaites `http://www.tsv1899benningen-ringen.de/chronik/update/alert/ibclogon.php` et `http://paypal.com.cg.ibin.webscr.cmd.login-submit.dispatch.5885d80a13c0db1f8e263663d3faee8d35d0e363192f28ea2.dgrrokpozefr.com/` en sont des exemples.

- **L'utilisation d'URLs très courtes** : A contrario du cas précédent, il est possible de rencontrer parfois des URLs très courtes. Pour ce faire, les attaquants ont recours à des services webs tels que celui de TinyURL [tin] qui crée des alias minimalistes redirigeant vers des URLs beaucoup plus longues. Par exemple, l'URL `http://www.amazon.com/Kindle-Wireless-Reading-Display-Globally/dp/B003FSUDM4/ref=amb_link_353259562_2?pf_rd_m=ATVPDKIKX0DER&pf_rd_s=center-10&pf_rd_r=11EYKTN682A79T370AM3&pf_rdt=201&pf_rd_p=1270985982&pf_rd_i=B002Y27P3M` qui se compose de 224 caractères peut être ramenée à l'URL `http://tinyurl.com/2enearw` de 26 caractères. Cette technique peut être utilisée afin de leurrer les moteurs de détection qui analysent les URLs à la recherche de comportements anormaux (c.-à-d. typiquement des tests heuristiques).

1. A ne pas confondre avec la présence d'un arobas (@) dans l'arborescence de l'URL qui permet d'apporter des indications complémentaires - appelées *attributs* - pour l'affichage de la page web (p.ex. pour spécifier un emplacement au sein de la page visitée, etc.)

- **L'utilisation de techniques de redirection** : Certains attaquants font appel à des techniques de redirection - plus ou moins masquées[1] - grâce à l'utilisation de caractères spéciaux (c.-à-d. @, //). Le caractère (@) doit alors être placé dans le FQDN, tandis que les caractères (//) doivent quant à eux être placés dans l'arborescence de l'URL, typiquement derrière la mention `http:`.
  L'URL contrefaite `http://usa.visa.com/track/dyredir.jsp?rDirl=http://200.251.251.10/.verified/` [CG06] est un exemple de redirection utilisant les caractères (//) : l'utilisateur croit visiter le site `http://usa.visa.com`, mais il est en réalité redirigé vers le serveur web possédant l'adresse IP 200.251.251.10.

- **L'utilisation de techniques d'encodage au sein de l'URL** : Tel qu'amorcé dans le point précédent, certaines URLs de phishing ont recours à des techniques d'encodage pour mieux masquer l'utilisation de caractères réservés, employés à des fins de redirections. Ces techniques sont utilisées tant pour éviter d'éveiller les soupçons de l'utilisateur, que pour leurrer d'éventuelles techniques de détection [Oll04]. Pour ce faire, les caractères réservés (p.ex. (://) et (@)) sont remplacés par leurs valeurs encodées. L'URL contrefaite `http://www.blizzard-accountlogin-security.com/login.asp?ref=https%3A%2F%2Fus.battle.net%2Faccount%2Fmanagement%2Fbeta-profil e.xml&app=bam&rhtml=y&rhtml=true` en est une illustration. On voit en effet apparaître la mention `https%3A%2F%2` dans l'arborescence de l'URL. En encodage ASCII, la valeur hexadécimale `%3A` correspond au (:), et la valeur `%2F` représente le (/).

- **La création de multiples dérivés d'une URL de phishing** : Les attaquants génèrent fréquemment de multiples URLs de phishing depuis une URL source, grâce à la modification/l'ajout de caractères tant au sein du FQDN que de l'arborescence (p.ex via le rajout de caractères aléatoires, de caractères d'indexation, etc.). Par exemple, les URLs contrefaites `http://fasemook.altervista.org/`, `http://faseboox.altervista.org/`, `http://fasebo.altervista.org/` et `http://faseboo.alterv ista.org/` usurpent le site Facebook (`http://www.facebook.com`) en altérant quelques caractères dans le FQDN. Un autre exemple : les URLs contrefaites `http://terabaap.hdfree.in/d.html`, `http://terabaap.hdfree.in/c.html` et `http://terabaap.hdfree.in/b.html` usurpent le site Orkut (`http ://www.orkut.com/`). Entre ces trois URLs, on voit que seul 1 caractère a été modifié, dans le nom de fichier.
  Précisons que des outils existent sur le web pour aider à la création de FQDN dérivés. On peut citer par exemple *Domain Typo Generator* [She].

- **La référence au FQDN/domaine légitime dans l'URL contrefaite** : Pour leurrer au mieux les utilisateurs, certaines URLs contrefaites mentionnent le FQDN/domaine légitime - dans le FQDN ou l'arborescence de la contrefaçon - sans pour autant effectuer une quelconque redirection. On peut citer par exemple l'URL de phishing `http://221.165.190.119/www.paypal.com/ws/www/x-us/webscr.html?cmd=x_login-run` qui utilise le FQDN de Paypal (`www.paypal.com`) - le site usurpé - en guise de nom de répertoire. Le site contrefait `http://verifymyfacebook.700megs.com/Index.html` mentionne quant à lui le nom de domaine légitime Facebook dans son FQDN.
  Précisons également que ce rappel du FQDN/domaine légitime est parfois réalisé via l'usage de l'arobas (@) dans l'arborescence de l'URL.

- **L'utilisation de multiples TLDs au sein du FQDN** : Il apparaît fréquent de rencontrer l'utilisation de multiples TLDs au sein des URLs de phishing, aussi bien dans le FQDN que dans l'arborescence de l'URL (typiquement lors de l'utilisation de techniques de redirection). On peut notamment citer les exemples d'URLs contrefaites : `http://www.ialp.org.br` qui utilise les TLD .ORG (pour *Organisation*) et .BR (pour *Brésil*), ou `http://www.click-here.us.ly/preview_login.htm.htm` qui utilise les TLD .US (pour *United States*) et .LY (pour la *Lybie*).

- **L'utilisation du protocole HTTP en lieu et place du protocole HTTPS** : Comme Ludl et al. [LMKK07] l'ont relevé dans leur étude, nous avons remarqué que la quasi-totalité des URLs de phishing rencontrées utilisent le protocole HTTP. Il est en effet plus rapide et moins dangereux pour un attaquant de ne pas utiliser de connexion sécurisée. En effet, il est ainsi moins exposé (vs. s'il

1. selon s'il y a encodage ou non.

devait obtenir un certificat valide) et évite toute alerte de sécurité émise par le navigateur et/ou le système d'exploitation en cas d'utilisation d'un certificat invalide. Les attaquants tirent ainsi avantage de la difficulté qu'ont les utilisateurs à distinguer un environnement sécurisé d'un environnement non-sécurisé [DT05] [FHH+02].

- **L'utilisation d'un numéro de port alternatif** : Tel que le rapporte régulièrement l'APWG[1] [APW10], certaines URLs de phishing font appel à une redirection de port dans la zone de FQDN (cf. section 2.2.1.1). Une URL contrefaite qui illustre cette redirection est : `http://186.97.10.96:8081/https/bancolombia.olb.todo1.com/olb/Init.php`, où l'on voit l'utilisation du port TCP 8081 pour accéder à la ressource demandée. Ce port 8081 - ou le port 8080 - sont des ports alternatifs au port 80, traditionnellement utilisé par le protocole HTTP.
  Cette redirection de port se justifie typiquement par le besoin de faire tourner un serveur web sur une machine corrompue (cf. section 3.4.1.1.4 pour plus de détails).

- **L'utilisation de mots-clés** : Il apparaît également que les URLs de phishing - qui usurpent des sites de login - utilisent de manière récurrente des mots-clés (p.ex. *login*, *signin*) en rapport avec la catégorie de sites usurpés au sein de leurs URLs. Par exemple, l'étude de Garera et al. [GPCR07] a retenu des mots-clés d'un minimum de 5 lettres (p.ex. *webscr*, *secure*, *banking*, *account*) pour détecter les URLs de phishing.

Notons que l'ensemble des techniques exposées ici peuvent être combinées entre elles pour construire une URL de phishing.

Précisons enfin que ces techniques ne sont pas l'apanage des URLs de phishing. Certaines d'entre elles sont également régulièrement rencontrées dans des URLs légitimes (p.ex. l'utilisation de mots-clés, de techniques de redirection ou d'encodage).

2.2.2 Zoom sur la page web

2.2.2.1 Synopsis d'une page web

Une page web est généralement constituée d'un document HTML (nommé *code source HTML*) auquel sont attachés un certains nombre de fichiers (p.ex. les images affichées, des scripts complémentaires pour l'ouverture d'encarts publicitaires en provenance d'un autre domaine, des feuilles de style, etc.). L'ensemble ainsi constitué est nommé *page web complète*.

Le *code source HTML* constitue le socle de la page web visitée. A ce titre, il est révélateur d'une grande partie de son contenu. Il permet par exemple de définir l'URL du site, donner des informations sur son propriétaire, définir les index de référencement pour les moteurs de recherche, ou encore il détaille le contenu visualisé par l'utilisateur (texte, images, etc.), que celui-ci soit intégré dans le corps du code source ou récupéré depuis d'autres ressources.

Un document HTML est constitué d'un assemblage de texte et de balises (ou *tags* en anglais). Les balises sont indiquées par des mots-clés (p.ex. TITLE, META SCRIPT, IMG, etc.) placés entre les caractères < et >[2]. On en dénombre plus de 90 sortes différentes.

Selon les standards définis par le W3C, un document HTML se décompose en 3 parties [W3Cb] (cf. figure 2.9) :

1. En préambule du code source, on trouve une déclaration de DTD (pour *Document Type Definition*). Celle-ci permet de préciser la version HTML utilisée pour le codage de la page. D'usage non-obligatoire, elle permet d'obtenir un meilleur rendu d'affichage de la page dans le navigateur web du client.
2. Un en-tête, nommé *HEAD* et identifié par une balise de même nom, qui contient des informations générales. On y retrouve typiquement :
   - le titre du document (balise TITLE),
   - des spécifications (p.ex. des mots clés, le nom d'auteur, une description générale de la page, des effets graphique associés au chargement de la page, etc.) indiquées dans les balises META.

1. cf. section 2.3 pour plus de détails sur cet organisme.
2. un début de balise est indiqué par <*mot-clé*>, tandis qu'une fin de balise est indiquée par</*mot-clé*>.

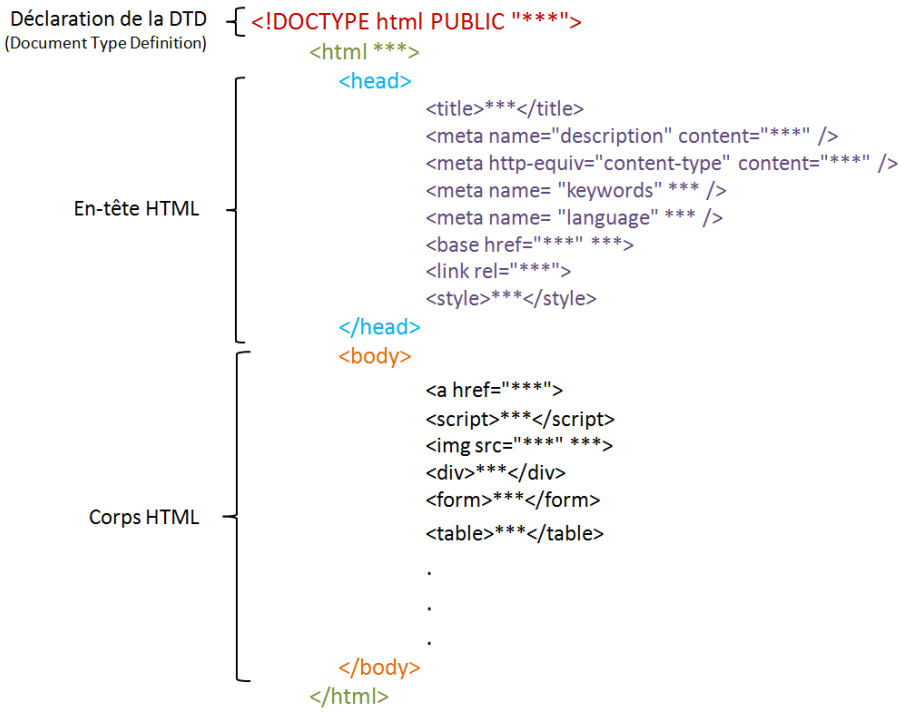


FIGURE 2.9 – Structure type du code source d'une page HTML

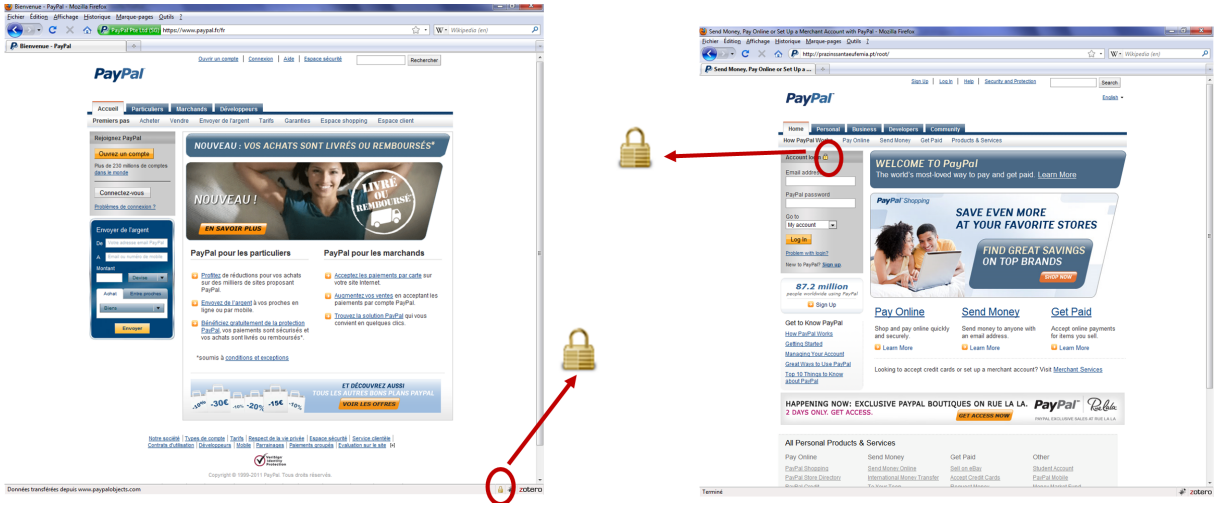
- une feuille de style (balise STYLE). Pour faciliter d'éventuelles modifications, on lui préfère plutôt une feuille de style spécifiée dans un fichier séparé, appelée par un lien hypertexte.
- le chemin d'accès absolu de tous les liens indiqués ultérieurement dans le document (balise BASE).
- des importations de documents externes (p.ex. pour des composants de feuille de style) via la balise LINK.

3. Le corps du document, nommé *BODY* et identifié par une balise de même nom, qui contient du texte, des tableaux, des images, des zones de saisie utilisateur, etc. Parmi les balises les plus utilisées dans cette partie du document, on peut citer notamment :
   - la balise A qui permet d'indiquer un lien hypertexte interne ou externe au document.
   - les balises FORM, INPUT utilisées pour définir des formulaires (typiquement des zones de saisie utilisateur).
   - la balise IMG pour insérer une image.
   - la balise DIV pour mettre en forme des données.
   - la balise SCRIPT pour insérer un bloc de code type Javascript.

L'arborescence constituée par l'assemblage et l'imbrication des différentes balises du code source HTML peuvent être désignées sous la terminologie structure/arbre DOM [W3Ca] (pour *Document Object Model*).

2.2.2.2 Caractéristiques d'une page de phishing

L'une des techniques préférées des attaquants - afin d'établir leurs contrefaçons - est d'utiliser des outils d'aspiration de sites webs dans l'objectif de s'approprier un maximum de contenu de la page légitime [Jam06]. On peut par exemple citer des outils comme Wget [Fou] ou WebWhacker [Squ] qui aident à la création de sites miroirs. Cette technique a le double avantage de simplifier l'élaboration du site contrefait, tout en rendant l'attaque moins détectable via l'utilisation d'un maximum de redirections légitimes et/ou l'utilisation de la structure originale. Notons néanmoins que quelques précautions peuvent être prises côté serveur web, afin de complexifier cette tâche (cf. section 2.5).



(a) Site légitime en HTTPS (b) Site contrefait en HTTP

FIGURE 2.10 – Utilisation de l'icône de cadenas dans Mozilla Firefox par deux sites Paypal

Des études [LMKK07] [PD06] [CG06] se sont attachées à distinguer les spécificités des pages webs de phishing. En complément, notre étude approfondie nous a laissés entrevoir que les pages contrefaites se perfectionnent, atteignant parfois une qualité d'imitation visuelle exceptionnelle, qui pourrait leurrer jusqu'aux utilisateurs les plus avertis.

Parmi les composants utilisés par les attaquants dans leurs contrefaçons de pages webs, nous pouvons commencer par citer ceux qui concernent l'aspect visuel :

- **L'intégration des logos et images du site légitime**.

- **L'utilisation du cadenas de sécurité au sein de la page web** : Toute page légitime affichée au travers d'une connexion sécurisée de type HTTPS, introduit automatiquement la présence d'un icône de cadenas dans la barre d'état du navigateur web client (typiquement en bas à droite du navigateur - cf. figure 2.10). Cet icône symbolise le certificat utilisé par le serveur web, élément indispensable au chiffrement des données client-serveur. A contrario, une page de phishing - qui utilise majoritairement une connexion non sécurisée - est dénuée de ce cadenas. Néanmoins, les attaquants ont trouvé une parade : pour leurrer les Internautes, ils insèrent l'icône au sein de la page web visitée, telle une image (cf. figure 2.10).

- **L'intégration de logos de sécurité** : Il est très fréquent de trouver au sein des pages webs de login légitimes, bon nombres de logos de sécurité utilisés comme gage d'une sécurité renforcée. La figure 2.11 en illustre quelques exemples. Les attaquants ont donc également recours à cette même technique.

- **L'utilisation de la structure globale de la page légitime** en conservant tailles et positionnements des images, textes, tableaux, formulaires, etc. du site original.

- **La conservation d'un maximum de redirections du site original**. Nous avons en effet remarqué que les sites de phishing conservent une - plus ou moins grande - proportion de liens originaux (cf. exemples détaillés en section 2.1). Ceci permet de minimiser le travail de l'attaquant et limiter les soupçons de l'utilisateur. En effet, nous remarquons sinon que bon nombre de liens additionnels renvoient vers une page d'erreur.

En regardant de plus près le contenu de la page web contrefaite (c.-à-d. son code source HTML), on peut également noter des caractéristiques typiques d'une contrefaçon :

FIGURE 2.11 – Exemples de logos de sécurité présents dans les pages webs

– **Les balises <TITLE> et <FORM> ne correspondent pas au domaine visité** : Par exemple le site contrefait (`http://www.top-pharmacies.com/ePHARMACIES_languages/English/admin/help/chaseupdate/chaseupdate/chaseupdate/Signon.htm?section=signinpage&=&cookiecheck=yes&=nba/signin`) qui usurpe le site légitime de la banque Chase contient *Chase Personal Banking Investments Credit Cards Home Auto Commercial Small Business Insurance* en guise de balise <TITLE>, ce qui est décorrélé du FQDN visité `www.top-pharmacies.com`.

– **Une large majorité des liens ne correspondent pas au domaine visité** : Typiquement des liens d'image (balise <IMG>), de redirection (p.ex. <A HREF>), etc. font appel au domaine légitime. En utilisant l'exemple du point précédent, on constate notamment que bon nombre de liens débutent par `http://www.chase.com/`.

Précisons toutefois que ces deux derniers points doivent faire l'objet de précautions particulières puisque nous avons également rencontré des sites légitimes qui répondaient à ces mêmes caractéristiques, faute de contenus adéquats ou explicites dans leurs balises. Par exemple, la balise <TITLE> de la page légitime de login Hotmail (`http://www.hotmail.com`) contient *Sign In*.

Notons enfin que l'ensemble des techniques exposées ici peuvent être combinées entre elles, et/ou avec les caractéristiques énoncées en section 2.2.1.2, pour élaborer le site de phishing.

2.3 Mise en œuvre du phishing

Diffusion des attaques : Le principal moyen de diffusion des attaques de phishing est le spam. Néanmoins, d'autres techniques existent pour attirer les utilisateurs sur des sites contrefaits ou détourner leurs données confidentielles [Oll04] [Emi05] [Jam06]. On peut citer par exemple l'utilisation :

– de bannières publicitaires insérées dans des sites webs,

– de publicités véhiculées sur les réseaux de messagerie instantanée,

– des moteurs de référencement pour indexer au mieux le site contrefait,

– des réseaux de Voix sur IP (VoIP) pour véhiculer des faux appels téléphoniques automatisés. Typiquement, ceux-ci informent l'utilisateur d'un blocage/d'une utilisation frauduleuse de son compte ou de sa carte bancaire, et l'invitent à rappeler un numéro qui leur demandera de saisir les informations confidentielles recherchées. Cette attaque est plus souvent désignée sous le terme de *vishing*, pour la contraction de VoIP et phishing.

– des réseaux mobiles via l'envoi de SMS/MMS (pour *Short Message Service*/*Multimedia Messaging Service*) de phishing, une technique connue sous la dénomination *smishing*, pour la contraction de SMS et phishing. Là encore, les messages véhiculés sont de nature similaire au point précédent : le danger d'une utilisation frauduleuse d'un moyen de paiement / d'un compte bancaire en l'absence de réaction de l'utilisateur. Celui-ci est alors invité à se rendre sur un site contrefait ou appeler une plate-forme frauduleuse pour résoudre le problème.

– d'outils malveillants tels que des *keylogger* ou *screenlogger* - récupérés typiquement par les Internautes lors de leur navigation web - qui espionnent les frappes clavier ou effectuent des captures écrans. On peut également mentionner des scripts malveillants qui visent à superposer de fausses zones de saisie de login sur les zones de saisie originales des pages légitimes, lors de la navigation web de l'Internaute [Oll04].

Localisation des sites de phishing : L'étude de McGrath et al. [MG08] met en exergue le fait que les sites de phishing ne sont généralement pas hébergés dans le pays annoncé par leur TLD. Leurs travaux démontrent en effet que près de 20% des domaines de phishing testés dans leur étude sont hébergés sur de multiples machines dispersées à travers le monde. Ils indiquent par ailleurs qu'un pourcentage significatif des adresses IP associées à ces sites de phishing sont utilisées par des particuliers (p.ex. 14% des sites contrefaits évalués sont hébergés par des particuliers situés aux États-Unis). Ceci s'explique notamment par la corruption des machines de ces utilisateurs, détournées afin d'être utilisées comme membres d'un réseau de botnet.

Sensibilisation au phishing et recensement d'URLs contrefaites : Face à la prolifération des attaques de phishing, les acteurs d'Internet (p.ex. FAI, banques) se mobilisent et diffusent très régulièrement des campagnes d'informations/de sensibilisation à destination des utilisateurs. La figure 1.1 - présentée dans le Chapitre 1 - et la figure 2.12 présentée ci-après en sont des exemples. En effet, la première figure illustre deux messages d'alerte diffusés par des banques sur leurs pages de login, tandis que la seconde figure illustre un email de sensibilisation émis par un FAI (c.-à-d. Orange) à destination de ses clients.

Au-delà de ces messages d'alerte, plusieurs organismes recensent et publient des listes d'URLs de sites contrefaits. Ces listes, plus connues sous la terminologie *liste noire* - ou *blacklist* en anglais - sont typiquement utilisées pour la détection des attaques de phishing (cf. section 2.4). Deux organismes particulièrement connus sont l'APWG [apw] et Phishtank [phi].

L'APWG est un consortium d'industriels ouvert - moyennant adhésion - aux banques, FAIs, éditeurs/constructeurs du domaine de la sécurité, ou organismes de recherche et/ou gouvernementaux. Pour la réalisation des études présentées dans ce mémoire, l'Institut Télécom/Télécom SudParis a donc signé un accord avec l'APWG, ce dernier s'autorisant à contrôler toute divulgation des données recueillies.

A contrario, Phishtank est une plate-forme d'accès libre où les utilisateurs soumettent et vérifient eux-même des URLs dites de phishing. Cette plate-forme est gérée par OpenDNS [ope].

Une durée de vie éphémère : Dans leurs travaux, McGrath et al. [MG08] ont calculé que la durée de vie moyenne d'une campagne de phishing était de 3 jours, 31 minutes et 8 secondes. Ils mentionnent par ailleurs que certaines campagnes ont une durée de vie extrêmement courte (33% d'entre elles n'excèdent pas 55 minutes), tandis 25% d'entre elles peuvent subsister durant 12 jours. De leur côté, Sheng et al [SWW+09] ont constaté que 63% des sites de phishing qu'ils ont testés ont vécu moins de 2 heures.

Globalement, on peut donc considérer que les sites de phishing ont une durée de vie très éphémère - et donc les listes noires associées également - d'où la multiplication d'URLs/domaines contrefaits.

2.4 Méthodes de détection/protection existantes côté client

Les techniques de détection/protection anti-phishing côté client découlent des différentes caractéristiques exposées en section 2.2. Précisons que ces techniques sont relativement similaires à celles rencontrées pour le spam, du fait que les deux attaques sont étroitement liées.

De manière générale, il est vivement recommandé aux utilisateurs d'équiper leurs postes clients avec les traditionnels anti-virus, anti-spyware, anti-spam, etc. en association avec la mise en œuvre de quelques règles de sécurité minimales (p.ex. restriction des paramètres d'acceptation des cookies et des contrôles ActiveX pour la navigation web, mise à jour régulière du système d'exploitation et

de :	"Orange votre service clients internet" <noreply@mailforge.orange.fr>
à :	[redacted]@orange.fr
date :	12/08/11 00:16
objet :	Emails frauduleux ? Ayez les bons réflexes !

voir l'en-tête complet

orange

Orange vous informe

Tentatives de phishing

Chère cliente, Cher client,

Nous attirons votre attention sur le fait que **des tentatives d'escroqueries par e-mail, aussi appelées "phishing", sont régulièrement menées.**

Le "phishing" est une technique consistant à usurper l'identité d'un tiers de confiance (banques, administrations, fournisseurs d'accès internet, etc.) pour soutirer des renseignements personnels à des fins malveillantes.

Ces emails reprennent généralement l'identité visuelle de différentes sociétés et vous invitent à modifier vos informations (coordonnées postales ou bancaires, identifiants et mots de passe) à partir d'un lien hypertexte.

Orange ne vous demandera en aucun cas par e-mail vos coordonnées bancaires, identifiants ou mots de passe.

Pour en savoir plus sur le phishing, l'assistance orange.fr met à votre disposition des compléments d'information accessibles depuis orange.fr > assistance > internet > sécurité > phishing.

Nous vous recommandons d'être vigilant lors de la consultation de vos e-mails et vous invitons à transférer tout message suspect à notre cellule spécialisée, à l'adresse suivante : abuse@orange.fr.

Adoptez les bons réflexes :

- **Traiter tout mail suspect comme indésirable,** sans y répondre, ni cliquer sur les liens.
- Par précaution, nous recommandons de **protéger votre ordinateur des risques liés à internet** : orange.fr > assistance > internet > sécurité > risques et prévention > protéger mon ordinateur

Nous vous remercions de votre confiance.

Votre service clients internet

FIGURE 2.12 – Le FAI Orange diffuse un email de sensibilisation aux attaques de phishing, à ses clients

des applications utilisés, etc.) [Oll04]. En complément d'autres techniques de détection/protection plus ciblées sur le phishing peuvent être utilisées. Elles sont exposées ci-après.

2.4.1 Le filtre idéal

Avant de détailler ces techniques, il est nécessaire d'introduire la notion de filtre anti-phishing idéal. Le doux rêve de tout(e) technique/outil qui cherche à s'attaquer à la problématique du phishing est de pouvoir détecter 100% des sites contrefaits, tout en autorisant 100% des sites légitimes.

Techniquement, cela se traduit par : 100% de sites contrefaits détectés, 0% de faux-positifs et 0% de faux-négatifs. Les faux-positifs (FPR pour *False Positive Rate*) correspondent au nombre/taux de sites légitimes classifiés à tort, comme sites contrefaits. Par similitude, les faux-négatifs (FNR pour *False Negative Rate*) correspondent au nombre de/taux de sites contrefaits classifiés à tort, comme sites légitimes.

Aucune technique de détection/protection existante à l'heure actuelle n'est capable d'atteindre cet objectif idéal. D'où la multitude d'approches proposées - qu'elles s'appliquent à l'URL, à la page web ou au moyen de diffusion de ces attaques - pour tenter de s'en approcher.

2.4.2 Au niveau des emails

Une première approche vise à bloquer/détecter les attaques de phishing au moment de leur diffusion en s'attaquant à leur principal vecteur de propagation : le spam.

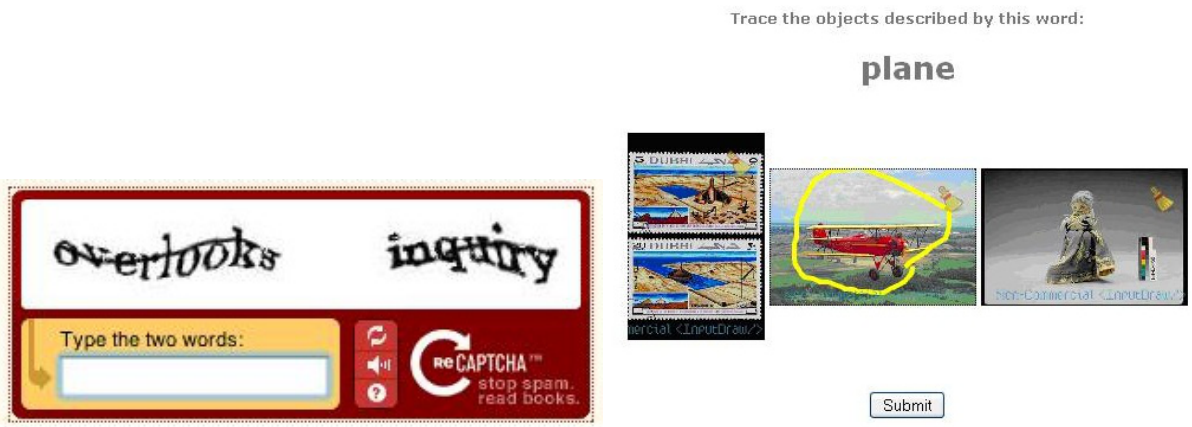
Pour ce faire, de nombreuses pistes sont proposées côté client. On peut notamment citer [Gas09] :

- un filtrage des emails - qui s'applique au corps du message et/ou à son en-tête SMTP - qui effectue :
  - la recherche de mots-clés interdits considérés comme caractéristiques d'une attaque, contenus dans un dictionnaire pré-établi.
  - un filtrage bayésien [Gra03] qui, après un apprentissage préliminaires sur *bons* et *mauvais* emails, classifie un email entrant selon les types de mots trouvés (c.-à-d. *bons* ou *mauvais*) et leur quantité.
  - une recherche de comportements dits anormaux - cette technique est également plus connue sous le terme de *tests heuristiques* - tels que : un email écrit en MAJUSCULES, un nom de domaine expéditeur anormalement long, un champ expéditeur vide, le couplage de plusieurs langues au sein du texte, etc. L'ensemble de ces comportements anormaux sont additionnés (selon les règles de pondération préalablement définies) et, si l'ensemble dépasse le seuil choisi, l'email est estampillé spam.
  - un filtrage selon le type de caractères/langues utilisés.
  - une analyse des images contenues dans le message [DGE07]. En effet, pour mettre en échec les techniques usuelles de filtrage par analyse du texte, certains attaquants encapsulent celui-ci dans des images.
  - une recherche de signatures. A l'image de la détection anti-virale, des signatures sont générées et pré-enregistrées à partir d'emails de spams avérés.
  - un filtrage à partir de listes noires (pour bloquer) et/ou blanches (pour autoriser) portant sur l'expéditeur (p.ex. nom de domaine, adresses IP, adresse email) ou les URLs contenues dans les emails. Ces listes peuvent être personnelles (c.-à-d. définie par l'utilisateur et/ou l'administrateur réseau auquel il appartient), ou émises par la communauté Internet.

- des techniques de vérification du domaine expéditeur de l'email qui s'assurent que la machine qui a expédié le message, est bien autorisée à le faire par le domaine pour lequel elle prétend agir. Ces techniques s'appuient typiquement sur des requêtes DNS[1].

- un filtrage dit par détection humaine, via l'utilisation des travaux de Turing [Tur50], pour essayer de détecter si le message a été expédié par un robot (p.ex. un attaquant) ou un être humain. En

1. cf. section 4.1.1 pour plus de détails.



(a) (b)

FIGURE 2.13 – Exemples de CAPTCHA utilisables pour une protection anti-spam [cap]

effet, pour leurs envois massifs de messages, les attaquants s'appuient sur des outils de génération automatique d'emails. Le filtrage mentionné ici est une technique amont qui vise à bloquer temporairement un expéditeur, jusqu'à ce qu'il réussisse un test que les robots ne peuvent théoriquement pas accomplir (p.ex. via un CAPTCHA - pour *Completely Automated Public Turing test to tell Computers and Humans Apart* -, qui est un test relativement simple et principalement visuel. Des exemples sont illustrés en figure 2.13). Ce test doit être réalisé par chaque expéditeur, uniquement lors d'un premier envoi d'email à destination d'un utilisateur protégé par cette technique. Si le test est réussi, l'expéditeur est ajouté à la liste blanche du destinataire.

– la signature des emails via l'utilisation de protocoles type S/MIME (pour *Secure/Multipurpose Internet Mail Extensions*) [Mai03] qui permettent d'authentifier l'expéditeur d'un message.

Ces pistes qui s'intéressent à bloquer les spams - et donc les attaques de phishing véhiculées par ce biais - sont certes intéressantes, mais insuffisantes. En effet la mise en œuvre de protocoles visant à signer les emails au niveau du poste client est peu facile d'accès pour des utilisateurs non-avertis. Ces protocoles s'en retrouvent donc insuffisamment déployés et, par conséquent, peu efficaces face à la problématique du phishing. Par ailleurs, les autres pistes évoquées ici ne sont ni infaillibles ni dépourvues de FPR/FNR. Enfin, les attaques de phishing pouvant être véhiculées par d'autres biais, cette approche ne peut être suffisante.

2.4.3 Au niveau du navigateur web

D'autres approches se focalisent quant à elles plus autour du navigateur web du client.

2.4.3.1 Sécurisation de la connexion client-serveur avec le protocole HTTPS

L'une d'entre elles repose sur l'établissement d'une connexion sécurisée avec le serveur web visité pour l'échange des données. Il s'agit ici des pages affichées via la mention HTTPS, accédées au travers du port TCP 443. Le protocole HTTPS n'est ni plus ni moins qu'une association des protocoles HTTP, et SSL ou TLS (pour *Secure Socket Layer* et *Transport Layer Security*). Il permet à la fois le chiffrement des données échangées entre client-serveur et l'authentification de ce dernier[1], grâce au(x) certificat(s) généré(s) par un tiers de confiance. Ce protocole HTTPS est typiquement utilisé pour le e-commerce.

Malheureusement, plusieurs études ont démontré le manque d'efficacité des indicateurs de sécurité affichés dans le navigateur web en cas de connexion HTTPS (cf. section 2.2.1.2). Le succès des attaques de phishing - qui usurpent en grande majorité des sites HTTPS grâce à des URLs contrefaites en HTTP - en est d'ailleurs la plus belle preuve.

1. au minimum. En effet, dans la majorité des cas, seul le serveur est authentifié. Pour certains usages il arrive que le client le soit également.

2.4.3.2 Alternatives pour la saisie des login/mot de passe

Pour pallier les problèmes des *keylogger* ou *screenlogger* (cf. section 2.3), certains sites d'e-commerce - typiquement les sites bancaires pour l'accès à la gestion des comptes sur Internet - proposent des zones de saisie des mots de passe à base de claviers virtuels (cf. exemple du site de login BNP Paribas, en figure 1.1 dans le Chapitre 1). D'autres sites ont quant à eux recours à l'utilisation de mots de passe à usage unique (ou OTP pour *One-Time Password* en anglais).

D'autres pistes [KK05] [RKKF07] [CHL08] proposent de stocker une base de données d'authentification côté client (c.-à-d. comprenant un couple login/mot de passe associé à une URL de site web légitime), en vue d'émettre des alertes dès lors que ces données sont saisies sur un serveur différent de celui qui est pré-enregistré. Dans la même veine, Dhamija et al. [DT05] ont proposé l'utilisation d'une fenêtre séparée pour la saisie des données d'authentification. Cette fenêtre dite de confiance repose sur l'utilisation d'une image de fond, préalablement choisie par le client et affichée par le serveur lors de la visite de ce dernier. Ross et al. [RJM+05] proposent quant à eux une extension pour navigateur web qui améliore la sécurité des mots de passe choisis par l'utilisateur. En effet, le plug-in s'interpose de façon transparente entre le client et les serveurs web visités, afin que chaque mot de passe utilisé soit différent et renforcé via une fonction de hash. D'autres alternatives similaires existent [WML06] [YS06]. Les défauts majeurs de ces solutions résident dans le besoin de stockage d'une base de login/mots de passe ou d'images choisies par l'utilisateur, qui introduisent des zones de vulnérabilités supplémentaires côté client ou côté serveur.

Enfin, une alternative proposée par Yue et al. [YW08] vise à fondre le mot de passe saisi par l'utilisateur sur un site de phishing, dans une multitude de mots de passe fantômes générés par l'outil et envoyés au site suspect. Cette méthode s'appuie sur une détection amont de la présence d'un site de phishing (p.ex. via une barre d'outils anti-phishing). Elle s'en retrouve donc exposée aux mêmes vulnérabilités.

2.4.3.3 Détection des signatures de pages webs contrefaites

Un autre style d'approche vise à comparer les pages de phishing à des bases de données de pages légitimes, en s'intéressant à leur aspect visuel (p.ex. structure générale, zones de texte, images, etc.). Par exemple, l'approche de Medvet et al. [MKK08] génère une base de signatures à partir des images de différents sites légitimes, auxquelles sont rattachées un certain nombre de mots-clés caractéristiques de leurs contenus/propriétaires. Ensuite, dès lors qu'une URL de phishing est suspectée, le moteur de recherche intégré à l'outil entre en action afin de retrouver la page légitime associée (c.-à-d. grâce aux mot-clés pré-enregistrés). Une signature de la page suspecte est alors générée pour comparaison avec la signature légitime sélectionnée. Dès lors que ces deux signatures sont trop similaires, une alerte est notifiée à l'utilisateur. Dans la même lignée, on peut également citer les approches de Hara et al. [HYM09], Wenyin et al [WHX+05] ou Chen et al. [CDM10].

Les inconvénients majeurs de ces approches résident à la fois dans le maintien de bases de données d'images et/ou de signatures d'images (dont découlent des zones de corruptions potentielles), et dans la quantité de faux-positifs générés dès lors que les pages webs ont recours à des contenus dynamiques.

2.4.3.4 Barres d'outils anti-phishing

Une dernière approche de détection se présente sous la forme d'une barre d'outils anti-phishing qui s'intègre dans le navigateur web. Dès lors que l'Internaute visite un site suspect, une alerte lui est notifiée. Le type d'alarme émise est très variable selon les éditeurs, allant d'un simple changement de couleur/texte affiché dans la barre d'outils (on parle alors de notification passive) jusqu'à un blocage de la navigation de l'Internaute dans l'attente d'une action de sa part (on parle alors de notification active). Ces barres d'outils anti-phishing - dont quelques exemples sont visibles en figure 2.14 - reposent sur l'utilisation de listes noires (pour vérifier l'URL visitée) et/ou de tests heuristiques (qui s'appliquent à l'URL et/ou au contenu de la page web). Diverses études ont été menées pour évaluer l'efficacité de ces barres anti-phishing [ZECH07] [WMG06] [ECH08] et/ou définir/évaluer les méthodes de détection qu'elles utilisent [LMKK07] [SWW+09] [GPCR07] [MSSV09]. Au global, il en ressort que cette approche - comme toutes les autres - n'est pas infaillible ni dépourvue de FPR/FNR. Néanmoins, il apparaît que les notifications actives sont incontournables et que la combinaison des deux types de techniques de

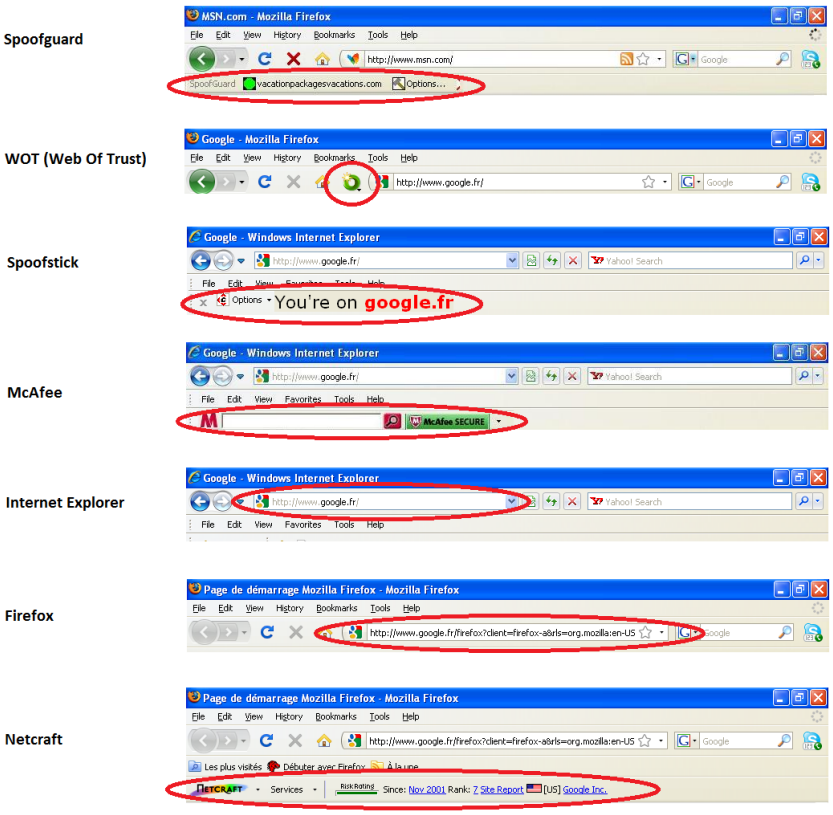


FIGURE 2.14 – Aperçu de plusieurs barres anti-phishing

détection (c-à-d. listes noires et tests heuristiques) sont un gage d'une meilleure détection (c.-à-d. qui limite les FPR/FNR). Une analyse plus détaillée de ces techniques de détection est menée dans le Chapitre 3.

Notons enfin que d'autres outils/mesures s'intéressent à la sécurité du navigateur web en général. Certains d'entre eux, qui peuvent répondre à la problématique des attaques de phishing diffusées via des outils/scripts malveillants (cf. section 2.3), sont évoqués en section 4.1.3.3.

2.5 Quelques alternatives côté réseau FAI/serveur web

Côté réseau FAI/serveur web, quelques mesures alternatives sont disponibles pour contrer les attaques de phishing. Néanmoins, celles-ci s'articulent majoritairement autour du spam.

L'une de ces alternatives concerne la sensibilisation des utilisateurs aux dangers de ces attaques, telle que nous l'avons déjà évoquée en sections 1.1 et 2.3.

Une deuxième alternative réside dans la mise en œuvre de mesures réseaux qui ciblent le spam. Sur ce point, de nombreuses approches sont évoquées et/ou utilisées par les FAIs ou entreprises [Gas09]. On peut par exemple citer la mise en place de *greylisting* pour rejeter temporairement les emails entrants, dans l'attente d'une réémission par le serveur de messagerie expéditeur (ce que les serveurs emails attaquants savent rarement faire). On peut également parler de mesures d'authentification des emails entre serveurs de messagerie (pour éviter cette tâche au niveau du poste client). Enfin, on peut également évoquer des règles de gestion de trafic telles que : A/ une limitation par les FAIs de l'utilisation du port 25 (qui correspond au protocole SMTP). Cette mesure cible principalement les particuliers, afin de réduire l'émission de spam via des machines corrompues, B/ une limitation du nombre/volume des

emails par expéditeur, C/ une limitation du nombre de destinataires d'un email, etc.

Néanmoins, cette alternative a ses limites puisque comme nous l'avons déjà évoqué, le spam n'est pas l'unique vecteur de propagation du phishing.

Enfin, une troisième alternative s'articule autour du serveur web visité. Au delà des mesures d'authentification vues en section 2.4.3, certaines approches consistent à compliquer la tâche des attaquants qui aspirent les sites webs légitimes. On peut par exemple citer l'utilisation de techniques d'encodage dynamiques - p.ex. en remplaçant des caractères ASCII par leur valeur hexadécimale - pour compliquer l'interprétation des liens contenus dans le code source de la page web [Fac05]. On peut également parler de l'utilisation de techniques qui visent à limiter la collecte d'adresses emails sur les pages webs, dans l'objectif de réduire le spam (p.ex. masquer/modifier les adresses emails en leur ajoutant des caractères vides ou en remplaçant le caractère @ par les caractères *(at)*, piéger les robots des spammeurs qui scrutent les pages webs dans des boucles dynamiques infinies, etc.) [Gas09].

2.6 Synthèse du chapitre

Ce chapitre a présenté les attaques de phishing ainsi que les mesures qui aident à s'en prémunir et/ou les détecter. Il apparaît évident qu'aucune des mesures exposées ici n'est suffisamment efficace ou infaillible pour éradiquer le problème. Néanmoins, parmi les techniques qui ciblent le phishing côté client, les barres d'outils anti-phishing se révèlent être un(e) outil/mesure particulièrement facile d'accès aux Internautes.

Comme pour toute mesure de détection, la véracité de la décision qu'elles délivrent est capitale. Les faux-négatifs sont évidemment bien plus dangereux que les faux-positifs. Néanmoins ces derniers auront tendance à lasser les utilisateurs, qui risquent alors de se détourner de leur solution anti-phishing dans les plus brefs délais.

La décision de légitimité établie par les barres d'outils anti-phishing s'appuie sur des listes noires et/ou des tests heuristiques. A l'image de la durée de vie des sites contrefaits, les listes noires semblent relativement éphémères. De plus, leur forte dépendance à une base de données - personnelle ou en provenance d'une tierce partie - fiable (c.-à-d. non corrompue) et suffisamment réactive, nous amène à penser qu'il serait dangereux de se limiter à leur simple utilisation (car trop de faux-négatifs potentiels). A contrario, les tests heuristiques semblent moins exposés aux vulnérabilités car plus autonomes. Mais la grande diversité des caractéristiques des sites webs (tant au niveau de l'URL que du contenu de la page web) peut les amener à délivrer des décisions plus enclines aux erreurs (c.-à-d. quelques faux positifs et faux-négatifs). Néanmoins, on peut imaginer qu'il y a une multitude de tests heuristiques à étudier : au moins autant que de signes caractéristiques d'un site de phishing. Quelle est leur efficacité ? Que vaut-il mieux analyser : l'URL ou le contenu de la page web ? Les tests heuristiques sont-ils tous égaux dans la détection des sites légitimes et contrefaits ? Quelle peut être leur pérennité ? La suite de notre étude s'essaie à y répondre.