

Le hameçonnage n'a pas de rapport avec les IDN

Stéphane Bortzmeyer

<stephane+blog@bortzmeyer.org>

Première rédaction de cet article le 3 Novembre 2009

<http://www.bortzmeyer.org/idn-et-phishing.html>

L'annonce par l'ICANN, le 30 octobre <<http://www.icann.org/en/announcements/announcement-30oct09-htm>>, de la possibilité d'avoir bientôt des IDN dans la racine, c'est-à-dire des domaines de tête en Unicode, a suscité la marée attendue de réactions provinciales <<http://www.trigeminal.com/samples/provincial.html>> et étroites de tous ceux qui regrettent qu'on ne puisse pas forcer ces insupportables chinois ou russes à utiliser la même écriture que nous. Un record dans ce domaine a été établi par l'ultra-nostalgique « *ICANN Approves Domain Names We Can't Type* » <http://www.pcworld.com/businesscenter/article/181094/icann_approves_domain_names_we_cant_type.html> ». Un des arguments avancés, très classique mais complètement faux, serait que les IDN favorisent le hameçonnage. Qu'en est-il ?

L'argument est que IDN permet des noms de domaine **homographes**, c'est-à-dire qui s'écrivent de manière visuellement identique (ou très proche). On risque alors de ne pas pouvoir distinguer `PAYPAL.com` et `P[Caractère Unicode non montré1]YPAL.com` (dans le second, il y a un alpha grec).

Les homographes existent bien, et ils n'ont pas attendu les IDN. Par exemple, `google.com` et `google.com` sont quasi-homographes (regardez bien). Unicode multiplie leur nombre car les écritures humaines n'ont pas été conçues par des technocrates rationnels mais sont issues d'une longue évolution distribuée sur toute la planète. Unicode est donc complexe, car le monde est complexe.

Mais le problème n'est pas dans l'existence d'homographes. Il est dans le fait que ce problème n'a rien à voir avec le hameçonnage. Je reçois beaucoup de rapports de hameçonnage au bureau et aucun ne dépend jamais d'homographes. La plupart du temps, le hameçonneur ne fait aucun effort pour que l'URL soit vraisemblable : il utilise un nom comme `durand.monfai.net`, voire une adresse IP. Et pour cause, très peu d'utilisateurs vérifient la barre d'adresse de leur navigateur, ne serait-ce que parce qu'ils ne comprennent pas ce qu'elle contient et qu'ils n'ont eu aucune formation sur les noms de domaines. Le hameçonneur, escroc rationnel, ne se fatigue donc pas.

1. Car trop difficile à faire afficher par L^AT_EX

Parfois, il faut quand même un petit effort. On voit des noms comme `secure-societegenerale.com` pour tromper sur `societegenerale.com`. À part les spécialistes du DNS, qui comprennent sa nature arborescente et le rôle de chaque label du nom, qui verra que `secure-societegenerale.com` n'a aucun rapport avec `societegenerale.com`?

Si vous ne faites pas confiance à mon vécu, regardez les nombreux articles de psychologie qui ont été consacrés au hameçonnage. Comme cette activité coûte cher aux banques, de nombreuses études scientifiques ont été faites pour mieux comprendre ce phénomène et pourquoi les utilisateurs se font avoir. Toutes concluent que le nom de domaine ne joue aucun rôle (et que donc l'argument d'homographie contre les IDN est du FUD). Voici une petite bibliographie (avec mes remerciements à Mike Beltzner de Mozilla) :

- « *"Decision Strategies and Susceptibility to Phishing"* <http://cups.cs.cmu.edu/soups/2006/proceedings/p79_downs.pdf> » de Downs, Holbrook & Cranor,
- « *"Why Phishing Works"* <http://people.ischool.berkeley.edu/~hearst/papers/why_phishing_works.pdf> » de Dhamija, Tygar & Hearst (lien alternatif <http://www.deas.harvard.edu/~rachna/papers/why_phishing_works.pdf>),
- « *"Do Security Toolbars Actually Prevent Phishing Attacks"* <<http://www.simson.net/ref/2006/CHI-security-toolbar-final.pdf>> » de, Wu, Miller & Garfinkel,
- « *"Phishing Tips and Techniques"* <<http://www.cs.auckland.ac.nz/~pgut001/pubs/phishing.pdf>> » de Gutmann.

On peut aussi citer, sur une problématique proche, « *"So Long, And No Thanks for the Externalities : The Rational Rejection of Security Advice by Users"* <<http://www.bortzmeyer.org/rational-security.html>> » de Cormac Herley, dont la section 4 parle de l'analyse d'URL. Et le rapport 2010 de l'Anti-Phishing Working Group <http://www.antiphishing.org/reports/APWG_GlobalPhishingSurvey_1H2010.pdf> dont la page 16 parle des IDN.

Mais, si toutes les études scientifiques montrent qu'il n'y a pas de connexion entre IDN et hameçonnage, pourquoi trouve-t-on tant d'articles faisant ce lien ? La principale raison est l'ignorance : les journalistes ne connaissent pas leur sujet, ne cherchent pas (pas le temps) et se recopient donc les uns les autres, sans questionner l'« information » originale. Ce phénomène n'est d'ailleurs pas réservé aux journalistes professionnels et s'étend aux blogueurs, aux twitteurs, etc.

Mais il y a une autre raison : beaucoup, parmi les « élites mondialisées », qui parlent anglais, regrettent, plus ou moins explicitement, qu'il ne soit pas possible de faire rentrer toutes les langues humaines dans le lit de Procuste d'une seule langue. Les vagues accusations d'hameçonnage contre les IDN, jamais étayées, ne sont donc que l'expression en douceur de leur refus de l'internationalisation.