L'hameçonnage social ou vol d'identité 2.0

Predrag.Viceic@epfl.ch, EPFL – DIT, chef de projet au KIS et concepteur de my.epfl

This article discusses some new threats to privacy arising from social networks and provides some thoughts on identity theft, social scamming, phishing and spamming.

L'article recueille un florilège de techniques de fraude informatique utilisant les réseaux sociaux comme médium de propagation. Il décortique les mécanismes utilisés par les cybercriminels et esquisse leurs motivations. Il ne propose pas de solutions, mais se permet de griffonner quelques maigres avertissements au réseauteur ...

Si l'on croit une étude récente du magazine PCWorld (2009), un tiers de réseauteurs ont au moins trois informations exploitables visibles sur leur profil, 59% ne savent pas qui peut voir leurs informations personnelles, 36% avouent qu'ils ne cachent aucune donnée personnelle et 28% d'entre eux acceptent les demandes d'amis des inconnus. Si de plus, nous admettons que, selon cette même étude, un tiers des réseauteurs utilise le même mot de passe pour tous leurs comptes en ligne, nous comprenons aisément pourquoi plus de 20'000 logiciels malicieux ont attaqué ou ont été déployés via les réseaux sociaux en 2008, c'est le chiffre fourni par **Kaspersky Lab**, l'entreprise russe spécialisée en sécurité informatique.

Nous sommes loin de la morale éducative, *ne parle pas à des inconnus, n'accepte pas leurs bonbons, ne dis pas ton nom ni adresse*, et autres commandements de la litanie qui a bercé notre enfance. Ceci peut s'expliquer par le prétendu anonymat de l'individu dans le flot d'informations présentes sur la toile. Après tout, tout le monde publie des informations personnelles sur le Net, n'est-ce pas précisément là le gage de l'innocuité de la démarche ? La confiance que nous avons dans cet anonymat est-elle réellement justifiée ?

Pour mieux esquisser les contours de la problématique, je commencerai par décrire les méthodes les plus courantes utilisées par les cyberarnaqueurs.

Nigerian 419

Tout le monde connaît la fameuse *arnaque nigériane*. Depuis des années, celle-ci s'est propagée par les courriels et consiste, dans sa version générique, à proposer une conséquente somme d'argent en échange d'un travail. Afin de récupérer l'argent, la victime est invitée à avancer une certaine somme au titre de frais quelconques. Il va sans dire que cette avance est perdue à jamais, et les fraudeurs volatilisés dans le cyberespace.

La version 2.0 de l'arnaque est beaucoup plus subtile et très difficile à détecter. Elle consiste à usurper l'identité de quelqu'un sur un réseau social, et d'envoyer une demande d'aide, financière, à sa liste d'amis. L'exemple donné, celui d'un employé de Microsoft résident de Seattle, a tout d'une légende urbaine, mais peut parfaitement servir d'illustration.

Ledit résident de Seattle, réseauteur à ses heures perdues, s'est donc réveillé un beau matin, pour constater qu'il ne peut plus se connecter sur son profil **Facebook**. De plus, il constate, que le dernier message de son profil indique qu'il a des soucis à l'étranger, et qu'il a besoin d'une certaine somme d'argent pour rentrer. Cette somme peut lui être versée sur un compte Western Union... Ne pouvant plus se connecter sur son profil, il ne peut pas modifier le message. Totalement réveillé à présent, notre héros essaie d'utiliser le profil de son épouse, qui est aussi son *amie* en ligne, pour afficher sur son mur que tout est en ordre et qu'il est victime d'une arnaque. Le seul problème, les fraudeurs ont pensé à enlever le statut d'ami à son épouse. Pour épicer le tout, il n'avait aucun autre moyen de contacter ses amis en dehors de Facebook...

Widget Warrior

Les arnaques nigérianes sont, en quelque sorte, du caviar parmi les arnaques. Elles demandent de la préparation, beaucoup de travail, de nombreux contacts avec la *clientèle*, impliquent plusieurs corps de métiers, du comptable au faussaire. De l'autre côté, nous avons les arnaques du type *Widget Warrior*, dont le seul objectif est de propager un code malicieux sur le plus d'ordinateurs possibles, en se servant des réseaux sociaux. Les bénéfices obtenus avec ce type d'arnaques sont moins évidents, mais leur impact est potentiellement beaucoup plus large.

La recette est très simple. En un, programmez un widget *super cool*, qui, par exemple, vous permet de trouver vos admirateurs virtuels. Ensuite, introduisez un code malicieux dans cet indispensable fleuron de la programmation. Pour finir, poussez à la distribution du *widget* à tous vos amis, ainsi qu'aux amis de vos amis et ainsi de suite. Voilà, le tour est joué. À la différence du mail, aucun filtre antipolluriel n'empêchera la distribution de ce cadeau à travers la toile. De plus, le *widget* est proposé par une connaissance, un ami, et, est donc au-delà de tout soupçon. Le code malicieux s'installera sur l'ordinateur de la personne visionnant la page et ensuite communiquera avec ses pairs sur d'autres ordinateurs des personnes également désireuses de connaître leurs admirateurs virtuels. Un réseau de machines zombies est né.

Koobface Virus

À la différence des arnaques du type Widget Warrior, dans lesquelles les réseauteurs servent de vecteurs de propagation, les infections du type **Koobface virus** se font de manière automatisée. Le mode de fonctionnement est simple. Le virus, dès son installation sur un ordinateur, utilise le navigateur Web pour se connecter sur le profil du propriétaire de l'ordinateur. Ensuite, il poste sur le *mur* un message incitant tous ses amis à aller consulter une vidéo *mégamarrante* (comme celle du chat qui tombe de la fenêtre...). Le site de la vidéo demande la mise à jour d'un composant logiciel quelconque, par exemple Flash, et voilà, le tour est joué. Le code malicieux est installé et peut recommencer son manège.

Communauté manipulée *(Contrived community)*

Contrairement à l'arnaque *nigériane* et aux deux méthodes de propagation de vers à travers les réseaux sociaux, l'arnaque de la *communauté manipulée* n'en est pas forcement une. Je vous laisse juger: un individu crée une communauté en ligne, p.ex **Informaticiens EPFL, volée 2002**. Ensuite, il invite toutes les personnes concernées à faire partie de la communauté. Enfin, il s'applique à récupérer le plus d'informations possible sur les membres de la communauté ainsi créée. Le seul hic est que notre individu n'a rien à voir avec les **Informaticiens EPFL, volée 2002** ! Son seul but est de créer un *yearbook* de la volée, avec les noms, photos, etc., et de le vendre aux personnes *inscrites*. Malin, non ?

Vol d'identité

Les fraudes décrites précédemment permettent de façon assez directe d'engranger les bénéfices, en recevant les versements d'argent de la part des victimes, ou en constituant un réseau d'ordinateurs infectés pouvant être utilisé pour les attaques de deni de service distribuées (DDOS) contre les infrastructures informatiques des organisations, entreprises et même des états. Le vol d'identité est toujours la première étape du processus. Ce vol peut consister à compromettre uniquement les données d'authentification, ou au contraire, à récupérer le plus d'informations possible sur la victime.
Le vol d'identité peut également être un objectif en soi. Ainsi, les données récoltées, comportant les noms, dates de naissance, amis, hobbies, animaux de compagnie, etc., peuvent ensuite être vendues à divers organismes criminels. Ceux-ci peuvent utiliser ces données de diverses manières. Les polluriels personnalisés peuvent être créés, semblant venir d'un ami ou d'une connaissance ou comportant une foule de données personnelles ajoutant à leur véracité. Ces informations peuvent également être utilisées pour deviner les réponses aux *questions personnelles* utilisées lors du changement de mot de passe d'un compte en ligne (animal de compagnie, nom de jeune fille, etc.).
En avril 2010, 1.5 million de comptes Facebook auraient été piratés par le dénommé **Kirllos**, selon le *Verisign iDefense Group*. Ces comptes auraient été mis en vente pour la modique somme de 25 USD à 45 USD, par tranche de 1'000 profils. 700'000 comptes auraient trouvé preneur, toutefois sans confirmation, selon Rick Howard, le directeur de *Verisign Cyber Intelligence*. L'entreprise derrière Facebook a nié l'importance des dégâts et dit avoir localisé **Kirllos**,... en Russie.
Indépendamment de la véracité ou non de cette histoire, elle est parfaitement crédible et montre les probables schémas d'attaque et de dissémination des données récoltées. Les histoires comme celle-ci deviendront la norme dans les années à venir, et sont une évolution logique des polluriels plus traditionnels, relativement bien arrêtés par les systèmes de filtrage de courriels actuels.

Un système automatisé de vol d'identité sur les réseaux sociaux

Les réseaux sociaux regorgent d'informations personnelles sur les millions de réseauteurs. Les concepteurs de ces sites l'ont compris et s'évertuent à mettre en place les mécanismes rendant la collecte automatique de ces données difficile. Ainsi, les sites fourniront les informations personnelles uniquement aux *amis* de l'utilisateur. Ainsi, on restreint le nombre de personnes ayant l'accès aux informations. Les cybercriminels utilisent différentes méthodes, relevant souvent de l'ingénierie sociale, pour rentrer dans le cercle des *ayants droit*. Le défaut de toutes ces méthodes est qu'elles nécessitent beaucoup de temps et ne sont pas très rentables, sauf dans les économies où l'heure de travail est très bon marché. Un système qui arriverait à automatiser la pénétration du cercle de confiance des réseauteurs serait une arme redoutable, mais est-il faisable ? Malheureusement, oui.
Une étude récente faite à l'EURECOM propose de construire, pas à pas, un système automatisé de vol d'identité sur les réseaux sociaux. Le but de l'attaquant est simple: récolter le plus d'informations personnelles sur un large échantillon de réseauteurs. Comment ? En devenant ami avec un grand nombre d'utilisateurs des réseaux sociaux afin d'obtenir l'accès à leurs informations confidentielles.
L'étude propose deux types d'attaque sur les réseaux Facebook, XING, StudiVZ et MainVZ. La première consiste à cloner un compte existant sur le réseau. L'homonymie dans les noms et prénoms étant très courante, les sites des réseaux sociaux ne peuvent l'interdire. Le programme malicieux récupère ensuite la photo du profil authentique et la rajoute sur le profil falsifié. Il invite ensuite les amis et les contacts du propriétaire du profil cloné à devenir ses *amis* dans le nouveau profil, en prétextant une erreur de manipulation qui lui aurait fait perdre l'usage de l'ancien compte. Les informations personnelles des nouveaux contacts deviennent ainsi disponibles. C'est le tour de la victime suivante !
Il est important de constater qu'à partir d'une information publique partielle, à savoir nom, prénom, la photo et la liste des contacts, le programme a pu devenir l'*ami* des contacts du réseauteur et pénétrer ainsi leur cercle de confiance.
La seconde attaque est très similaire à la première. Au lieu de cloner un profil existant sur un réseau social, le code malicieux clone un profil existant sur un autre réseau. Ainsi, la supercherie est encore plus indécelable, car la demande de contact du programme pirate est très crédible dans les yeux de ses victimes. Après tout, quoi de plus naturel que d'avoir les mêmes listes d'amis sur des réseaux différents ?

La seule difficulté restant à surmonter se trouve en amont, comment pirater le CAPTCHA? L'étude mentionnée utilise les algorithmes très simples de détection de bord et de recherche de correspondances entre les formes ainsi obtenues et les jeux de caractères dans les différentes polices et tailles. Comme les systèmes d'authentification déployés sur les sites en ligne proposent souvent plusieurs essais à l'utilisateur, même des méthodes rudimentaires se sont avérées redoutablement efficaces.
Les résultats de l'étude sont stupéfiants. En moyenne, 40'000 profils ont pu être collectés par jour, ramenant le chiffre total à 5 millions de profils avec les informations de contact, et 1.2 million de profils contenant les informations complètes sur leurs propriétaires. L'acceptation des demandes de contact par les profils supposément connus par la victime était de plus de 60%! En comparaison, pour les invitations faites par les profils choisis au hasard, sans lien avec la victime, c'était moins de 30%.

Quel avenir?

Il serait illusoire de croire que les cyber mafias ne déploient pas déjà ce genre de systèmes automatisés à travers les réseaux sociaux. Je vous épargnerai les conseils que nous connaissons tous et qui pourraient se résumer à la litanie du début de cet article. Néanmoins, il est important de savoir que nous ne pouvons pas éviter que les données personnelles que nous dévoilons sur la toile deviennent visibles et exploitables. Il est regrettable que les grands sites de réseautage donnent l'impression d'assurer la sécurité des informations personnelles mises en ligne, alors que c'est faux et techniquement impossible à garantir. Le plus judicieux reste de considérer les données ainsi fournies comme publiques et de les dévoiler avec parcimonie et en connaissance de cause. **Les grands frères vous regardent.**

Berlin Alexander Platz, juin 2010

Références

- *Beware: Identity Thieves Harvest Social Networks*, PCWorld, www.pcworld.com/article/167511/beware_identity_thieves_harvest_social_networks.html
- *All your Contacts Are Belong to Us: Automated Identity Theft Attacks on Social Neworks*, www2009.org/proceedings/pdf/p551.pdf

GLOSSAIRE

CAPTCHA: un *captcha* est une forme de test de Turing permettant de différencier de manière automatisée un utilisateur humain d'un ordinateur. La vérification utilise la capacité d'analyse d'image ou de son de l'être humain. Un *captcha* usuel demande que l'utilisateur tape les lettres et les chiffres visibles sur une image distordue qui apparaît à l'écran.

DDOS: attaque par déni de service par exemple en bloquant un serveur de mail par l'envoi d'un très grand nombre de messages.

demande d'ami: action de rajouter quelqu'un dans son cercle de confiance sur un réseau social.

filtre antipolluriel: logiciel détectant et supprimant les polluriels.

hameçonnage (*phishing*, parfois filoutage): c'est une technique utilisée par des fraudeurs pour obtenir des renseignements personnels dans le but de perpétrer une usurpation d'identité.

ingénierie sociale: *(social engineering)* c'est une forme d'escroquerie utilisée en informatique pour obtenir un bien ou une information. Cette pratique exploite l'aspect humain et social de la structure à laquelle est lié le système informatique visé. Utilisant ses connaissances, son charisme, l'imposture ou le culot, le pirate abuse de la confiance, l'ignorance ou la crédulité de personnes possédant ce qu'il tente d'obtenir.

machine zombie: en sécurité informatique, une machine zombie est un ordinateur contrôlé à l'insu de son utilisateur par un pirate informatique. Ce dernier l'utilise alors le plus souvent à des fins malveillantes, par exemple afin d'attaquer d'autres machines en dissimulant sa véritable identité.

mur: dans Facebook, c'est la partie du profil permettant au réseauteur d'afficher une information d'actualité le concernant.

polluriels: le spam, pourriel ou polluriel est une communication électronique non sollicitée, en premier lieu via le courrier électronique. Il s'agit en général d'envois en grande quantité effectués à des fins publicitaires.

profil: compte sur un site de réseautage.

widget: petite application pouvant être rajoutée dans l'interface d'un site Web.