



Pratique

Collecte passive d'informations – principes

Błażej Katak 

Degré de difficulté



Le fait de rendre publique trop d'informations peut mener à la violation des principes de la politique de sécurité, et de là faciliter l'attaque sur le système informatique d'une entreprise ou d'une institution. Nous verrons où et comment trouver facilement les données précieuses qui peuvent servir à compromettre le système de la protection d'une entreprise.

Les pentests. Dans les derniers temps, ce mot est devenu très populaire dans les journaux métier. Pour plusieurs personnes, qui basent leur connaissance sur les films du type *The Hackers*, où l'intrusion d'un système informatique consiste à voler dans l'espace virtuel parmi les tours brillantes semi-transparentes, les pentests sont de la *magie noire*. Mais il ne faut pas rendre le diable plus noir qu'il ne l'est... Il suffit de connaître quelques outils et méthodes de travail pour, avec un peu de bonne chance, compromettre un système de protection voulu.

Dans cet article, je ne veux pas expliquer la théorie du hacking, ni parler de l'éthique suivant laquelle agit un vrai hacker. Cet article ne sera pas non plus un simple guide à travers les outils, ni une liste de type TODO. Je voudrais montrer comment les connaissances acquises ou pouvant être acquises par la plupart des utilisateurs familiarisés avec les ordinateurs et le réseau, bien corrélées, peuvent servir à briser les systèmes de protection d'une grande partie des entreprises et institutions présentes sur Internet.

J'essaierai de montrer ce qu'il est possible de faire avec un navigateur, une chaise, de la musique et, bien sûr, notre intelligence sans laquelle

tout devient inutile. Je ne donne pas tous les détails techniques pour que le lecteur puisse expérimenter tout seul et ressentir de la satisfaction lorsque quelque chose *a finalement réussi*.

Ce texte est adressé, avant tout, aux utilisateurs débutant dans le domaine de la sécurité informatique, mais familiarisés avec les ordinateurs et Internet. Alors, au boulot.

Comme on fait son lit...

Au début, je parlerai des facteurs assurant, dans la plupart des cas, le succès final. La base de tous les pentests est l'environnement, commode et adopté aux besoins individuels.

Cet article explique...

- quelle est la première phase des tests de pénétration,
- comment se défendre contre une collecte passive des informations.

Ce qu'il faut savoir...

- utiliser un navigateur Web,
- connaître le modèle du réseau TCP/IP.

Tests de pénétration

Les tests de pénétration (audit de sécurité) est un processus de vérification du système de sécurité de l'infrastructure informatique par un groupe de personnes autorisées et qualifiées, par la simulation de différentes actions qui peuvent être entreprises par un intrus potentiel. Le but des tests est alors d'effectuer une attaque contrôlée contre les systèmes de production, de la détecter les failles et les éliminer et, par conséquent, d'élever le niveau de sécurité informatique d'un sujet (entreprise ou institution).

Du point de vue du savoir que peut posséder une équipe de pentests, les tests se divisent en ce qu'on appelle black box testing, alors aucune information sur l'objet examiné et white box testing – tous les détails techniques sont connus (les configurations, l'accès aux bases de données, au code source, etc.). Il existe aussi la division du point de vue de la localisation des auditeurs, c'est-à-dire les tests interne, effectués de derrière d'un objet examiné (par exemple un réseau) et internes – du point de vue, par exemple d'un employé.

Chaque test de pénétration est composé des phases suivantes :

- Collecte passive d'informations – le processus de recherche et de collecte des données concernant l'objet examiné en mode passif, sans pour autant fournir à l'objectif testé (par exemple une société) aucun prétexte qu'il est observé ;
- Scannage et mappage du réseau – l'analyse du trafic, l'examen des règles du pare-feu (en anglais firewalking) ;
- Fingerprinting – l'identification des types et des versions des systèmes d'exploitation utilisés dans le réseau ;
- Détection des failles et vulnérabilités dans la configuration – l'analyse des données collectées et définition des vecteurs d'attaque potentiels ;
- Intrusion – l'exploitation des failles et le cassage du système de sécurité ;
- Escalade des droits d'accès – l'obtention des droits d'accès dans les systèmes d'exploitation ;
- Reporting – la récapitulation de toutes les données sous forme d'un rapport, leur analyse avec le service technique de l'objectif examiné (par exemple d'une entreprise) et l'indication des méthodes permettant d'améliorer la sécurité de l'infrastructure IT.

Les pentests possèdent, en tant que méthodologie, leur propre standard : OSSTMM (*The Open Source Security Testing Methodology Manual*) de l'Institut ISECOM (*The Institute for Security and Open Methodologies*). Pour plus d'informations, référez-vous à l'adresse : <http://www.isecom.org/osstmm/>

Pour moi, la partie intégrale (voire indispensable) est mon navigateur préféré (*Firefox*), une bonne musique (pour moi, relaxante), un crayon, un bloc-notes (avec un grand nombre de pages) et une chaise commode ou un fauteuil dans lequel l'auditeur passera beaucoup de temps. Le temps est le dernier élément de ce puzzle et de sa quantité dépend le résultat final de nos actions. Pour nous faciliter la tâche, admettons que notre horizon temporel est l'infini (nous n'avons pas besoin de plus de temps). Cela nous permettra de nous concentrer sur les questions essentielles.

Après la configuration d'un environnement de travail convivial, nous

pouvons passer aux actions plus concrètes. Nos actions concerneront la phase préliminaire des tests de pénétration, c'est-à-dire la collecte passive (plus ou moins) des informations sur l'objectif potentiel (en anglais

Passive Information Gathering – Encadré Tests de pénétration). Imaginons que nous sommes un conseiller en sécurité qui doit recueillir la plus grande quantité d'informations sur une entreprise (ici, nous l'appellerons *Invulnérables S.A.*) en ne trahissant pas qu'une telle collecte a lieu.

Nous négligeons ici la question qui est le commettant éventuel (cela peut bien être cette même entreprise *Invulnérables S.A.* ou sa concurrence). C'est un élément non lié complètement à la commande en tant que telle nous nous concentrerons uniquement sur sa réalisation.

Low hanging fruits

Par quoi commencer ? Certains visiteraient tout de suite le site www.invulnerables.com, ce qui est en opposition avec notre principe de base – rester cachés. Le réseau est plein d'endroits dans lesquels nous pouvons trouver beaucoup d'informations sur notre objectif. Les plus souvent, ce sont les services qui ont été créés il y a longtemps et étaient conçus pour faciliter la tâche des utilisateurs profitant d'Internet. Pourtant, comme on le sait de l'histoire de l'humanité, il s'est avéré que *tel est pris qu'il croyait prendre*. Ces informations, collectées et analysées de façon appropriée, peuvent donner une idée précise de ce qui se passe dans une entreprise X, quelle est sa structure, quels sont ses fournisseurs, qui la gère, quand elle travaille, etc. La liste est très longue. Bien sûr, il n'est pas possible de déterminer tous ces détails dans chaque cas, mais il faut au moins essayer car ce sont ce qu'on appelle

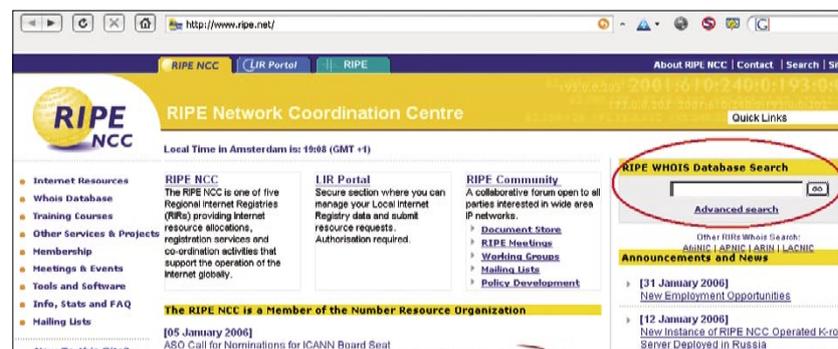


Figure 1. Le site RIPE de la base WHOIS



en anglais *low hanging fruits*, alors des choses que l'on peut avoir sans effort.

Bien. Commençons par déterminer où se trouve la société *Invulnérables S.A.*, quelles sont ses heures d'ouverture, quelle est l'adresse du site et prenons éventuellement les numéros de téléphone. Pour cela, nous avons donc besoin d'un annuaire des téléphones. Il n'est pas nécessaire d'en avoir un sous la main, bien qu'elle soit certainement disponible au bureau de poste le plus proche - dans le réseau, il y a des sites qui offrent ce type d'informations, par exemple : Annuaire des Entreprises (<http://www.pagespro.com>) ou bien *Pages Jaunes* (www.pagesjaunes.fr). Si vous voulez trouver sa localisation géographique, il suffit, par exemple maps.google.com ou www.multimap.com.

Il est conseillé de noter tout ce qu'on réussit à déterminer. Par exemple : les numéros de téléphones peuvent être utiles pour les attaques sociotechniques (si celles-ci seront exigées) ou pour *wardialing*, en utilisant le préfixe du numéro. Les adresses du courrier électronique montrent quel format de l'adresse est employé (par exemple : *Mr.Bean@invulnerables.com*).

Si notre objectif (la société *Invulnérables S.A.*) est coté en Bourse, nous pouvons vérifier quelles informations sont disponibles sur le site de la Bourse (<http://www.bourse.fr>) sur les portails financiers (par exemple : <http://www.boursorama.com>, <http://www.boursedirect.fr>, etc.).

Une fois les informations de base collectées, vérifions ce qu'on peut retrouver dans d'autres endroits. Commençons par le service Whois.

Qui est qui...

Whois est une base (Encadré *Service WHOIS*) contenant les informations sur les sujets Web enregistrés et a été conçue en vue de fournir les coordonnées à tous ceux qui ont besoin d'une telle information (par exemple quand on veut contacter un administrateur d'un réseau donné). Vu que nous appartenons aussi à ce

Listing 1. Les résultats de la requête effectuée dans la base WHOIS concernant le nom de l'entreprise *invulnerablessa.com*

```
% This is the RIPE Whois query server #2.
% The objects are in RPSL format.
%
% Note: the default output of the RIPE Whois server
% is changed. Your tools may need to be adjusted. See
% http://www.ripe.net/db/news/abuse-proposal-20050331.html
% for more details.
%
% Rights restricted by copyright.
% See http://www.ripe.net/db/copyright.html
% The object shown below is NOT in the RIPE database.
% It has been obtained by querying a remote server:
% (whois.snd.pl) at port 43.
% To see the object stored in the RIPE database
% use the -R flag in your query
%
Domain object:
domain:          invulnerablessa.com
registrant's handle: msk9999 (CORPORATE)
nservers:        ns2.invulnerablessa.com.[10.14.86.33]
                 nsl.invulnerablessa.com.[10.14.86.32]
created:         1999.12.02
last modified:   2005.12.13
registrar:       MLASK
74, rue Cléopatre
44444 Sainte-Ave
France/France
+33.22 5003333
help@snd.fr
option:          the domain name has not option
Subscribers Contact object:
company:         INVULNERABLES S.A.
Street:          150, QUAI D'ASTERIX ET OBELIX
city:            66666 Montjoly
location:        FR
handle:          msk9999
last modified:   2000.10.19
registrar:       MLASK
74, rue Cléopatre
44444 Sainte-Ave
France/France
+33.22 5003333
help@snd.fr
Whois database last updated: 2006.01.10
%REFERRAL END
```

groupe (bien que nous ne voulions pas contacter une personne de l'entreprise *Invulnérables S.A.*), voyons d'où vient le vent.

Pour cela, nous pouvons utiliser un outil très populaire appelé whois, disponible dans la plupart des systèmes Linux ou demander la base WHOIS directement à partir du navigateur Web (dans notre cas, le registre RIPE – <http://www.ripe.net/> – cf. la Figure 1). Nous nous servons de la deuxième option car elle est plus universelle et il n'est pas nécessaire d'utiliser un système d'exploitation concret.

Dans le champ sélectionné, il faut entrer le nom de domaine (par exemple : une requête sur *invulnerablessa.com*), le nom de l'hôte (par exemple <http://www.invulnerablessa.com/>) ou l'adresse IP de cet hôte.

Au début, nous ferons une requête sur le domaine *invulnerablessa.com*.

Le Listing 1 présente un enregistrement concernant l'entreprise en question. On voit que *Invulnérables S.A.* utilisent deux serveurs DNS portant les adresses 10.14.86.32 et 10.14.86.33, sont enregistré dans MLASK et ont leur siège à Montjoly

Service WHOIS

Le but principal du service WHOIS est de fournir des informations permettant de connaître le propriétaire et la disponibilité d'un nom de domaine quelle que soit son extension (d'une société, d'une institution, d'une organisation). La base est divisée en deux parties – la première partie est responsable des informations relatives à une plage d'adresses IP donnée (appelée *Network Service-based*), la deuxième partie – des noms de domaine (appelée *Name Service-based*). La base WHOIS contient, entre autres, les adresses IP affectées à un sujet donné, le numéro du système autonome (AS – utilisé pour le routage BGP), les données des personnes responsables du maintien de l'enregistrement et de plusieurs autres.

La base WHOIS a été divisé en quatre Registres Régionaux (en anglais *Regional Internet Registries*) :

- APNIC – Asie et Pacifique (*Asia-Pacific Network Information Center*)
– <http://www.apnic.net/>
- ARIN – Amérique du Nord (*American Registry for Internet Numbers*)
– <http://www.arin.net/>
- LACNIC – Amérique Latine et Caraïbes (*Latin American and Caribbean Internet Address Registry*) – <http://www.lacnic.net/>
- RIPE NCC – Europe (*Réseaux IP Européens Network Coordination Centre*)
– <http://www.ripe.net/>

ou Sainte-Ave, ce qui doit confirmer les données que nous avons obtenues dans le premier pas. Si non, probablement le siège de l'entreprise a changé, l'entreprise a fusionné avec une autre ou c'est une adresse d'une des filiales, responsables de la TI. Il est recommandé de noter ce fait.

Les adresses des serveurs DNS serviront comme une seconde requête (sur l'adresse : 10.14.86.32 ou 10.14.86.33) de la base WHOIS (Listing 2).

Et qu'est-ce qu'on obtient ? Premièrement, que le bloc d'adresses IP a été affecté à l'entreprise analysée (10.14.86.0/24), avec qui contacter (Mr.Bean – code JF6969-RIPE), l'adresse, le téléphone, l'email et le numéro AS (AS12345). Ce dernier indique si les Invulnérables S.A. ont un système autonome ou bien s'ils utilisent un autre (dans notre cas, ils sont connectés au réseau WARIA.FR). Chacune de ces informations peut être vérifiée dans la base WHOIS et peut nous fournir des données intéressantes. Maintenant, c'est ton tour, cher lecteur – va voir ce qu'on peut encore tirer de la base RIPE, et il y en a encore assez.

Vu que WHOIS n'est pas le seul endroit où l'on peut trouver des « fruits pendants aux arbres », l'étape suivante sera le service

très connu DNS, qui peut introduire beaucoup de confusion.

Et vous êtes...

DNS n'est rien d'autre qu'un ensemble de systèmes offrant la translation des adresses IP en nom et vice versa. Alors, il traduit les noms qui sont plus simples à mémoriser par l'utilisateur en adresses numériques, exigées dans la communication dans les réseaux basés sur le protocole IP. Qu'est-ce que cela signifie pour nous ? L'esprit humain est parfois prévisible et est souvent guidé par les schémas et les habitudes. Par exemple, le nom d'un serveur Web commencera par le préfixe *www*, le pare-feu sont souvent appelé *fw*, les *DNS* – *ns*, le courrier – *mail*, etc.

Les administrateurs se servent souvent d'un ensemble de noms issus d'une seule source, par exemple de la mythologie, de l'astronomie (par exemple les noms des planètes et de leurs lunes) ou d'un schéma admis (par exemple *dhcp13-14* peut signifier une station portant l'adresse terminée par les octets 13.14 et affectés à partir du serveur DHCP, *bud011122-01* – le premier ordinateur localisé dans le bâtiment n° 1 au 11ème étage dans le local 122.). C'est très utile et facilite l'administration du réseau, mais laisse beaucoup d'informations qui peuvent

être utiles pour un intrus potentiel. Il arrive parfois que le serveur DNS extérieur traduit les noms des ordinateurs se trouvant dans le réseau protégé !

Pour tester les serveurs DNS, nous pouvons utiliser des outils commodes. Sous Linux, il y en a plusieurs (*dig*, *nslookup*, *host*), sous Microsoft Windows, nous ne disposons que d'un outil standard (*nslookup*). Mais n'oublions pas que nous devons rester inaperçus. C'est pourquoi, il est recommandé d'utiliser un site Web qui demandera le serveur DNS et nous retournera les résultats sans envoyer un moindre paquet vers l'objet étudié. Pour cela, nous pouvons nous servir, par exemple du service <http://www.network-tools.com/>, offrant des options de requête avancées (Listing 2). Analysons ce que le serveur DNS retournera, si nous l'interrogeons sur le *invulnerables.com*.

La première chose qui frappe aux yeux est le fait que le serveur est administré par Jean Dubois (visible dans le champ email: de l'enregistrement SOA) et que le serveur DNS de base porte le beau nom Barbu. Après l'analyse de l'inscription entière il s'avère que ce serveur n'est rien d'autre que NS1, retourné par la base WHOIS et que son partenaire – NS2 – c'est *Chauve*, qui est en même temps le serveur de messagerie (enregistrement MX). Ces informations sont très importantes car si les serveurs qui donnent accès à ce type de services se trouvent physiquement sur la même machine, nous avons détecté une violation très grave des principes de sécurité. La compromission du serveur de messagerie et l'obtention des droits de superutilisateur à la fois, signifient la prise de contrôle du serveur DNS, ce qui peut avoir de conséquences très graves !

Bien sûr, il se peut qu'à la même adresse IP, il existe deux systèmes physiques différents. À cet effet, il est possible d'employer, par exemple du *loadbalancing* ou un autre système redirigeant les requêtes vers les différents services.

À partir d'un enregistrement TXT, qui contient les informations

**Listing 2. Les résultats de la requête de la base WHOIS sur les adresses IP de la société invulnerablessa.com**

```
% This is the RIPE Whois query server #2.
% The objects are in RPSL format.
%
% Note: the default output of the RIPE Whois server
% is changed. Your tools may need to be adjusted. See
% http://www.ripe.net/db/news/abuse-proposal-20050331.html
% for more details.
%
% Rights restricted by copyright.
% See http://www.ripe.net/db/copyright.html
% Note: This output has been filtered.
%       To receive output for a database update, use the "-B" flag
% Information related to '10.14.86.0 - 10.14.86.255'
inetnum:        10.14.86.0 - 10.14.86.255
netname:        INVULNERABLESINTPL-NET1
descr:          Invulnérables
country:        PL
admin-c:        JF6969-RIPE
tech-c:         JF6969-RIPE
status:         ASSIGNED PA "status:" definitions
mnt-by:         WARIA-MNT
source:         RIPE # Filtered
person:         Mr. Bean
address:         Invulnérables S.A.
Address:        150, QUAI D'ASTERIX ET OBELIX
address:        66666 Montjoly
address:        FRANCE
phone:          +33 55 5005555
fax-no:         +33 55 5005566
e-mail:         Mr.Bean@invulnerablessa.com
nic-hdl:        JF6969-RIPE
source:         RIPE # Filtered
% Information related to '10.14.86.0/22AS12345'
route:          10.14.86.0/22
descr:          WARIA.FR
origin:         AS12345
mnt-by:         WARIA-MNT
source:         RIPE # Filtered
```

texte, nous pouvons apprendre que le numéro de contact a changé (5005550). Cela pourra signifier que quelqu'un a fait une faute de frappe lors de la saisie des données dans l'enregistrement, ce qui est plutôt peu probable, si l'on prend en considération le fait que le chiffre 5 n'est pas à proximité de 0, ou bien c'est le numéro du support technique de notre entreprise. Et probablement Mr.Bean et Jean Dubois y travaillent. Ce sont de nouvelles données pour une attaque sociotechnique éventuelle.

Revenons à DNS. À partir du site http://www.ip-plus.net/tools/dns_check_set.en.html/, vous pouvez accéder à un outil qui, lors du questionnement du DNS, essaie de télécharger en exécutant des requêtes

le fichier de zone entier. Aujourd'hui, peu de serveurs sont vulnérables à ce type d'attaque. Elle consiste à lire tout le contenu de la base du serveur DNS de base pour le domaine donné (dans notre cas *invulnerablessa.com*) par le biais d'une seule requête ! Tous les enregistrements sont retournés – les noms des stations avec leurs adresses IP. Ces données sont très utiles pour déterminer la structure du réseau examiné.

Il est aussi possible d'aborder le problème d'un autre côté. Si nous avons à faire à un administrateur zélé, chaque inscription dans la base principale sera reflétée dans la base inversée. Pour accéder à ces informations, il suffit d'envoyer ce qu'on appelle résolution inverse (en

anglais *reverse lookup*), c'est-à-dire demander au serveur DNS sur le nom pour l'adresse IP donnée. Alors, nous envoyons une requête sur l'adresse IP connue (par exemple 10.14.86.32) et en réponse, nous obtenons le nom qui lui est affecté (par exemple : *ns1.invulnerables.com*, *barbu.invulnerables.com*). Si nous avons un peu de chances, nous pouvons obtenir la base DNS entière !

Pour que cela ne soit pas si facile que ça, admettons que notre admin est très harassé de travail, ou bien il parle tous les jours via IRC et en résultat, les bases DNS sont incomplètes. Il nous reste encore la façon la moins élégante, c'est-à-dire l'attaque par force brutale contre DNS. Il s'agit ici d'une simple devinette. Nous pouvons essayer de deviner si un nom donné est présent dans la base DNS (par exemple : *fw.invulnerables.com*, *ids.invulnerables.com*, *srv.invulnerables.com*, *srv1.invulnerables.com*, etc.).

Après l'analyse des résultats précédents, nous pouvons aussi essayer de restreindre l'étendue de la recherche et de questionner les noms qui sont liés aux noms déjà connus (par exemple les personnages de la mythologie). Dans le cas envisagé, cela peuvent être, par exemple : *poilu.invulnerablessa.com*, *chevelu.invulnerablessa.com*, *moustachu.invulnerables.com*, *boucle.invulnerablessa.com*, *hirsute.invulnerablessa.com*, etc. Tout dépend de notre invention et de notre imagination.

Google is your friend...

Justement. Jusqu'alors, nous avons profité des sources d'informations moins populaires. Passons alors à celles plus évidentes. Toute personne ayant cherché quelque chose dans le Réseau, a visité le site <http://www.google.com/>. C'est le moteur de recherche principal, en ce qui concerne la recherche de quelque chose sur n'importe quoi. Ce n'est pas sans raison que Google est le meilleur moteur de recherche sur Internet. Sa base comprend de centaines de millions de références, documents, photos et coordonnées. Dans le numéro

Chez votre marchand de journaux

PHP starter kit

PHP starter kit

PHP pour
les débutants

+CD PROGRAMMEZ EN PHP - ENVIRONNEMENT DE TRAVAIL COMPLET SUR LE CD

SUR LE CD

- La meilleure introduction à PHP5 : Bestseller **PHP5 Power Programming** en version électronique
- Les versions commerciales complètes IDE pour PHP : **phpED 3.3** et **TruStudio**
- Ensemble des outils complets pour réaliser des applications Web
 - éditeurs de programmation professionnels
 - environnements complets (Apache+PHP+MySQL) installés en un seul clic
- Atelier
Ensemble de meilleures applications les plus populaires en PHP
Quelles applications faut-il vraiment utiliser ?

À LA UNE !

À LA UNE !

Programmez en PHP

! Installer un environnement complet et créer une première application

Créer un portail pas à pas
Faire connaissance d'un système CMS : eZ publish de A à Z.

Boutique Internet en 5 minutes
osCommerce – la meilleure solution Open Source du type e-commerce

PHP et bases de données
Nous testons les bases Open Source pour PHP les plus populaires

Sécurité des applications PHP et du serveur Web
Vérifiez si vous tes menacés

Nouveau PHP5, complètement orienté objets
Vaut-il encore la peine de créer des projets en PHP4 ?

+ **REVENUS POUR LE WEBMASTER**
Comment peut-on gagner sur le Net ?

www.phpsolmag.org/fr

Également disponible sur shop.software.com.pl



3/2005 du magazine *hakin9*, Michał Piotrowski a publié l'article : *Google dangereux – recherche des informations confidentielles*, il n'est donc pas nécessaire de présenter les techniques connues comme *google hacks*. Je voudrais seulement mentionner quelques petits détails.

Nous savons qu'à l'aide d'une requête bien construite, nous pouvons obtenir des résultats très intéressants. Les opérateurs de type : `site:`, `inurl:`, `intext:`, `intitle:` etc... facilitent la recherche des données et limite l'étendue de la recherche, alors les résultats sont plus proches aux attentes. Mais que faire quand nous avons obtenu une référence intéressante ? Si nous cliquons sur celle-ci, notre adresse IP sera enregistrée dans les journaux du serveur Web examiné, et c'est une situation que nous voulons éviter. Nous pouvons profiter d'un serveur proxy gratuit, la liste des serveurs proxy gratuits est disponible par exemple à l'adresse www.proxy4free.com ou du paquet Tor tor.eff.org. Il existe aussi une solution plus élégante.

C'est le service Google qui nous vient à l'aide, en mettant à disposition son cache (Figure 2). La plupart des pages sont disponibles off-line, c'est-à-dire, que nous ne devons pas demander au serveur Web d'obtenir la page qu'il héberge. C'est ce que nous trouvons dans le cache qui est souvent la copie fidèle de la page originale. Mais il y a ici un petit truc. Lequel ? Consultons la Figure 3.

La figure présente l'exemple d'une requête sur le site du magazine *Hakin9* et sa page disponible à partir du cache. Pourtant, après l'analyse des journaux d'Ethereal (Figure 4), lancé lors du chargement de cette page, il s'est avéré que certaines requêtes ont été envoyées et réceptionnées au/à partir du serveur Web du magazine *hakin9* (adresse 62.111.243.84) ! Et c'est justement ce que nous voulions éviter. Que c'est-il passé ?

Après une brève analyse, la réponse est banale. Le fichier avec les styles CSS (paquets 31) et quelques images (paquets 78, 80, 130 et suivants) contenus sur la page ont été téléchargés à partir du serveur

Listing 3. Les résultats de requête rendus par serveur DNS de la société invulnerables.com

```
Nslookup Query the DNS for resource records
domain query type A - Address NS - Name server CNAME - Canonical name SOA
Start of authority MB - Mailbox domain MG -
Mail group member MR - Mail rename domain NULL - Raw data record WKS - Well-
known services PTR - Domain pointer HINFO - Host info MINFO - Mailing list
info MX - Mail exchange TXT - Text strings RP - Responsible person AFSDB -
AFS database X25 - X25 PSDN address ISDN - ISDN address RT - Route through
NSAP - NSAP address NSAP-PTR - NSAP-style pointer SIG - Security signature
KEY - Security key PX - X.400 mail mapping info AAAA - IPv6 address LOC -
Location NXT - Next domain SRV - Location of services NAPTR - Naming
authority pointer KX - Key exchange delegation UINFO -
User info UID - User ID GID - Group ID MAILB - Mailbox-related records ANY
- Any type server query class IN - Internet CH - CHAOS HS - Hesiod ANY
- Any class port timeout (ms)
no recursion advanced output
[10.14.86.32] returned an authoritative response in 156 ms:
Header
rcode: Success
id: 0 opcode: Standard query
is a response: True authoritative: True
recursion desired: True recursion avail: True
truncated: False
questions: 1 answers: 13
authority recs: 0 additional recs: 3
Questions
name class type
invulnerablesa.com ANY ANY
Answer records
name class type data time to live
invulnerablesa.com IN SOA server: barbu.invulnerablesa.com
email: jean.dubois@chauve.invulnerablesa.com
serial: 2005050508
refresh: 43200
retry: 3600
expire: 3600000
minimum ttl: 1209600
8100s (2h 15m)
invulnerablesa.com IN NS chauve.invulnerablesa.com 8100s (2h 15m)
invulnerablesa.com IN NS barbu.invulnerablesa.com 8100s (2h 15m)
invulnerablesa.com IN NS ns1.invulnerablesa.com 8100s (2h 15m)
invulnerablesa.com IN NS ns2.invulnerablesa.com 8100s (2h 15m)
invulnerablesa.com IN MX preference: 10
exchange: mail.invulnerablesa.com
8100s (2h 15m)
invulnerablesa.com IN A 10.14.86.33 8100s (2h 15m)
invulnerablesa.com IN TXT Invulnérables S.A. 8100s (2h 15m)
invulnerablesa.com IN TXT 150, Quai d'Asterix et Obelix 8100s (2h 15m)
invulnerablesa.com IN TXT FAX: +33 55 5005566 8100s (2h 15m)
invulnerablesa.com IN TXT TEL: +33 55 5005550 8100s (2h 15m)
invulnerablesa.com IN TXT 44444 Sainte-Ave, FRANCE 8100s (2h 15m)
invulnerablesa.com IN TXT RP: Admin <admin@invulnerablesa.com> 8100s (2h
15m)
Authority records
[none]
Additional records
name class type data time to live
chauve.invulnerablesa.com IN A 10.14.86.33 8100s (2h 15m)
barbu.invulnerablesa.com IN A 10.14.86.32 8100s (2h 15m)
ns1.invulnerablesa.com IN A 10.14.86.32 8100s (2h 15m)
ns2.invulnerablesa.com IN A 10.14.86.33 8100s (2h 15m)
mail.invulnerablesa.com IN A 10.14.86.33 8100s (2h 15m)
-- end --
URL for this output
```



Figure 2. L'accès au cache du service Google

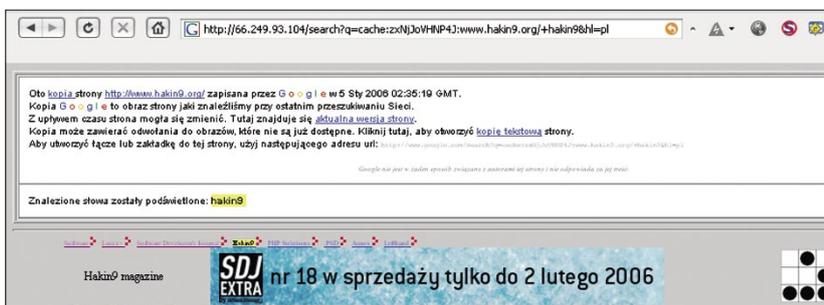


Figure 3. La page cache du service Google

No.	Time	Source	Destination	Protocol	Info
6	2006-01-14 22:49:05.78409	192.168.3.242	66.249.93.104	HTTP	GET /search?q=cache:zxnj30vHNp4J:www.hakin9.org/+hakin9&hl=pl HTTP/1.1 200 OK (text/html)
9	2006-01-14 22:49:05.96859	66.249.93.104	192.168.3.242	HTTP	Continuation of non-HTTP traffic
10	2006-01-14 22:49:06.03520	66.249.93.104	192.168.3.242	HTTP	Continuation of non-HTTP traffic
13	2006-01-14 22:49:06.03762	66.249.93.104	192.168.3.242	HTTP	Continuation of non-HTTP traffic
16	2006-01-14 22:49:06.04090	66.249.93.104	192.168.3.242	HTTP	Continuation of non-HTTP traffic
17	2006-01-14 22:49:06.10764	66.249.93.104	192.168.3.242	HTTP	Continuation of non-HTTP traffic
20	2006-01-14 22:49:06.11110	66.249.93.104	192.168.3.242	HTTP	Continuation of non-HTTP traffic
23	2006-01-14 22:49:06.30449	192.168.3.242	62.111.243.84	HTTP	GET /style.css HTTP/1.1
78	2006-01-14 22:49:07.79703	192.168.3.242	62.111.243.84	HTTP	GET /images/flag_en.gif HTTP/1.1
80	2006-01-14 22:49:07.79713	192.168.3.242	62.111.243.84	HTTP	GET /images/flag_pl.gif HTTP/1.1
82	2006-01-14 22:49:07.79721	192.168.3.242	62.111.243.84	HTTP	GET /p1/img/gif/der.png HTTP/1.1
84	2006-01-14 22:49:07.79731	192.168.3.242	62.111.243.84	HTTP	GET /adv1/ew.php?what=zone:156&ma=7f70931 HTTP/1.1 404 Not Found [Unassembled Packet]
86	2006-01-14 22:49:07.79739	192.168.3.242	62.111.243.84	HTTP	GET /images/logo.gif HTTP/1.1
88	2006-01-14 22:49:07.79747	192.168.3.242	62.111.243.84	HTTP	GET /style/klocki.gif HTTP/1.1
119	2006-01-14 22:49:07.84656	62.111.243.84	192.168.3.242	HTTP	HTTP/1.1 304 Not Modified
123	2006-01-14 22:49:07.85274	62.111.243.84	192.168.3.242	HTTP	HTTP/1.1 200 OK (GIF89a) [Unassembled Packet]
130	2006-01-14 22:49:07.86670	62.111.243.84	192.168.3.242	HTTP	Continuation of non-HTTP traffic
131	2006-01-14 22:49:07.87243	62.111.243.84	192.168.3.242	HTTP	Continuation of non-HTTP traffic
134	2006-01-14 22:49:07.89626	62.111.243.84	192.168.3.242	HTTP	Continuation of non-HTTP traffic
137	2006-01-14 22:49:07.89992	62.111.243.84	192.168.3.242	HTTP	Continuation of non-HTTP traffic
138	2006-01-14 22:49:07.90498	62.111.243.84	192.168.3.242	HTTP	Continuation of non-HTTP traffic
141	2006-01-14 22:49:07.92838	62.111.243.84	192.168.3.242	HTTP	Continuation of non-HTTP traffic
146	2006-01-14 22:49:07.93190	62.111.243.84	192.168.3.242	HTTP	Continuation of non-HTTP traffic
147	2006-01-14 22:49:07.93591	62.111.243.84	192.168.3.242	HTTP	Continuation of non-HTTP traffic
150	2006-01-14 22:49:07.93928	62.111.243.84	192.168.3.242	HTTP	Continuation of non-HTTP traffic
151	2006-01-14 22:49:07.94299	62.111.243.84	192.168.3.242	HTTP	Continuation of non-HTTP traffic
154	2006-01-14 22:49:07.94909	62.111.243.84	192.168.3.242	HTTP	Continuation of non-HTTP traffic
155	2006-01-14 22:49:07.95287	62.111.243.84	192.168.3.242	HTTP	Continuation of non-HTTP traffic
216	2006-01-14 22:49:09.38469	192.168.3.242	62.111.243.84	HTTP	GET /images/flag_it.gif HTTP/1.1

Figure 4. Le journal d'Ethereal démontrant la cause du questionnement du serveur source

source. Que faire alors ? Comment accéder au cache sans rumeur ?

Et encore une fois Google nous surprend. Quand nous lisons attentivement l'en-tête du cache, nous pouvons voir un message en petits caractères informant que la page ne peut être affichée qu'en mode texte (Figure 5). La référence au cache a un peu changé. À la fin, elle contient un court paramètre : `&strip=1`. C'est lui qui est responsable de notre anonymat. Alors, pendant la recherche dans le cache, il suffit de copier le résultat de la recherche de Google dans le presse-papier, le coller dans

le champ URL, ajouter `&strip=1` magique et appuyer sur Entrée. Simple et élégant.

Si l'on pense à des moteurs de recherche, en premier lieu, on se réfère automatiquement à Google. N'oublions pas qu'il existe encore d'autre services de ce type (search.msn.com, www.yahoo.com, www.clusty.com, etc.). Je recommande particulièrement le service Clusty, qui, outre la vitesse du questionnement, présente les résultats finaux d'une façon très intéressante en les triant en sous-catégories (Figure 6).

Il ne faut pas non plus oublier les fichiers `robots.txt` placés sur les serveurs Web pour éviter l'indexation des certaines pages. Chaque moteur de recherche se sert d'un mécanisme de recherche et d'enregistrement automatique des sites Web (appelé *webcrawler*). *Webcrawler* analyse le code HTML trouvé, toutes les références qu'il contient et essaie de les suivre. Pour empêcher l'indexation des pages souhaitées, on crée dans le service un fichier spécial nommé `robots.txt`, qui continent les instructions pour webcrawler quels revois il doit négliger. Mais le fichier `robots.txt` en tant que tel est enregistré – et de ce fait peut nous fournir d'autres attractions dans nos pentests.

Il vaut la peine de consulter les groupes de discussion. Parfois, il arrive que les administrateurs du réseau analysé participent aux différents groupes ou aux forums de discussion. Si l'on visite `groups.google.com`, on peut trouver les discussions très intéressantes du personnel technique. Souvent, les questions concernant les détails techniques d'une configuration ou des descriptions des problèmes rencontrés dans le travail peuvent nous aider à déterminer quels systèmes sont employés dans l'entreprise. Il arrive que les éléments de la configuration des périphériques ou des services sont révélés par les employés insouciantes de la TI. Dans notre cas, il faudrait chercher Mr. Bean et Jean Du-bois et vérifier s'ils participaient à une discussion quelconque (en utilisant leurs prénoms et noms de familles accompagné d'`invulnerables.com`, ou des adresses email).

Très souvent, les entreprises informatiques présentent dans leurs portefeuilles les informations sur les projets terminés et réussis. Il est possible de trouver ici les données sur les systèmes installés, avec les détails, tels que : le type et la version du système d'exploitation, la structure du réseau interne, la version de la base de données, etc. Toutes ces informations, avant d'être publiées sur Internet, doivent être autorisées par le client. Alors, ses secrets sont révélés avec son consentement !



D'où vient le vent ?

Que nous reste-t-il outre les moteurs de recherche ? Toute une panoplie de possibilités.

Le service Netcraft (<http://www.netcraft.com/>) fournit les statistiques concernant les sites Web. Mais il informe aussi sur d'autres détails très importants. Par exemple, si nous demandons sur le site du magazine *hakin9* (Figure 7), nous obtenons les données sur la localisation, le serveur DNS, l'adresse IP, le nom de retour, le système d'exploitation sur lequel fonctionne le Web, et même l'information sur la version de ce serveur. Si les requêtes adressées à Netcraft sont bien construites, nous pouvons retrouver certains noms DNS introuvables sur le serveur de noms, encore une tâche pour le lecteur. La lutte contre le spam et les bases RBL (en anglais *Realtime Blackhole List*) est aussi une épée à deux tranchants. Le site *openrbl.org* nous fournit les informations sur les spammers potentiels, mais il offre aussi une fonction très intéressante : la recherche des adresses email pour la déclaration du spam provenant d'une adresse IP donnée.

Il existe aussi quelques sites (par exemple *www.sampspade.org*, *www.dnsstuff.com*) qui fournissent des outils complets pour la recherche des informations. Par exemple : DNSStuff permet d'effectuer les attaques de type email brute forcing. L'un des outils demande au serveur de messagerie si tel ou tel mail sera accepté. Si le destinataire n'existe pas, une erreur est retournée (Figure 8). Dans ce cas, deux explications sont possibles : le format de l'adresse donnée est incorrect, ce qui est peu probable, vu que dans les tests précédents, on pouvait connaître la forme appropriée ou cette personne ne travaille plus dans l'entreprise. Alors, si l'on connaît le format de l'adresse email, en se servant des dictionnaires des noms et prénoms et d'un simple script, il est possible de déterminer la liste des personnes qui travaillent dans une entreprise donnée.

Mais il faut se rendre compte que le fait de *frapper à la porte* avec tant d'obstination peut éveiller des soupçons. Si un administrateur consulte

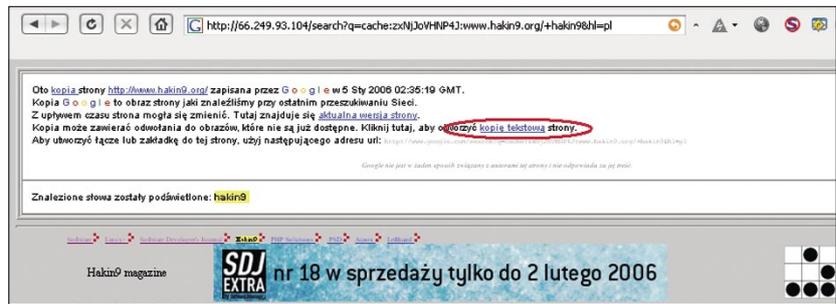


Figure 5. Le cache de Google avec les renvois à la version texte du document



Figure 6. Le résultat retourné par le service *www.clusty.com*

de temps en temps les fichiers journaliers et trouve quelques tentatives de connexion échouées provenant d'une adresse IP, il commencera à s'intéresser à ce qui s'est passé, ce qui n'est pas bon pour nous. Je voudrais recommander ici le logiciel de la société *VisualWare*, qui dessine sur la carte du monde la trace d'un point vers l'autre. Le site *visualroute.visualware.com* donne accès à une version de démonstration, mais il faut s'enregistrer. Vous pouvez essayer de vous enregistrer ou profiter de la base *www.bugmenot.com* qui comprend les données pour l'autorisation.

Vous avez un message...

Dans le numéro 5/2004 du magazine *hakin9*, Tomasz Nidecki dans l'article *Tracer l'expéditeur d'un email* décrit les méthodes d'extraction des informations à partir des en-têtes du courrier électronique. Ces en-têtes contiennent les données sur le trajet d'un message, les systèmes de messagerie utilisés, la protection antispam, quel client de messagerie a été employé par l'expéditeur, quel

adressage IP est utilisé à l'intérieur du réseau de l'entreprise, etc. Vous obtenez tout cela à partir d'un seul email reçu à partir de l'objet analysé. Vous pouvez profiter des comptes de messagerie gratuits et envoyer une demande d'une nouvelle offre commerciale et attendre la réponse. Vous pouvez aussi soumettre la recherche dans les forums Internet.

Défense

Comment se défendre contre une collecte passive d'informations ? Voici quelques méthodes recommandées :

- ne pas révéler le format de l'adresse utilisé à l'intérieur de l'organisation, par exemple pour la base WHOIS, il faut créer un nouveau compte (*whois@invulnerableness.com*) ;
- partout où c'est possible, utilisez un seul numéro de téléphone pour l'organisation entière – il sera plus difficile de deviner la plage de numéro affectée par l'opérateur de téléphonie, ce qui rendra plus difficiles les attaques de type wardialing ;

Site report for www.hakin9.org				
Site	http://www.hakin9.org	Last reboot	unknown	
Domain	hakin9.org	Netblock owner	LEWARTOWSKIEGO JOZEFA 6	
IP address	62.111.243.84	Site rank	53809	
Country	PL	Nameserver	ns.software.com.pl	
Date first seen	August 2003	DNS admin	hostmaster@ns.software.com.pl	
Domain Registry	publicinterestregistry.net	Reverse DNS	host-ip64-243.crowley.pl	
Organisation	ul. Konstruktorska 6, Warszawa, 02-673, Poland	Nameserver Organisation		
Check another site: <input type="text"/>				
Hosting History				
Netblock Owner	IP address	OS	Web Server	Last changed
LEWARTOWSKIEGO JOZEFA 6 WARSZAWA Connected by Crowley Data Poland Sp. z o.o.	62.111.243.84	Linux	Apache/2.0.52 Aurox Linux	20-Mar-2005
LEWARTOWSKIEGO JOZEFA 6 WARSZAWA Connected by Crowley Data Poland Sp. z o.o.	62.111.243.84	FreeBSD	Apache/2.0.52 Aurox Linux	12-Mar-2005
LEWARTOWSKIEGO JOZEFA 6 WARSZAWA Connected by Crowley Data Poland Sp. z o.o.	62.111.243.84	FreeBSD	Apache/2.0.47 Aurox Linux	13-Jan-2005
LEWARTOWSKIEGO JOZEFA 6 WARSZAWA Connected by Crowley Data Poland Sp. z o.o.	62.111.243.84	FreeBSD	unknown	12-Jan-2005
LEWARTOWSKIEGO JOZEFA 6 WARSZAWA Connected by Crowley Data Poland Sp. z o.o.	62.111.243.84	FreeBSD	Apache/2.0.47 Aurox Linux	8-Jul-2004
Crowley Data Poland PROVIDER Local Registry	62.111.243.84	FreeBSD	Apache/2.0.47 Aurox Linux	23-Feb-2004
Crowley Data Poland PROVIDER Local Registry	62.111.243.84	FreeBSD	Apache/1.3.26 Unix Debian GNU/Linux PHP/4.1.2 mod_fastcgi/2.2.10	3-Sep-2003

Figure 7. Le résultat de la requête concernant le domaine hakin9.org dans le service Netcraft.com

E-mail Tester results for <u>jas.fasola@google.com</u>		
Generated by www.DNSSstuff.com		
Getting MX record for google.com (from local DNS server, may be cached)... Got it!		
Host	Preference	IP (s) [Country]
smtp4.google.com.	10	66.102.9.25 [US]
smtp1.google.com.	10	216.239.57.25 [US]
smtp2.google.com.	10	64.239.167.25 [US]
smtp3.google.com.	10	64.239.189.25 [US]
Step 1: Try connecting to all of these (in a random order, per RFC1129 5.2.4):		
smtp4.google.com.	-	66.102.9.25
smtp1.google.com.	-	216.239.57.25
smtp2.google.com.	-	64.239.167.25
smtp3.google.com.	-	64.239.189.25
Step 2: If still unsuccessful, queue the E-mail for later delivery.		
Trying to connect to all mailservers:		
smtp4.google.com.	- 66.102.9.25	[Could not connect: Got an unknown RCPT TO response: 550 5.5.3 ... Invalid id]
smtp1.google.com.	- 216.239.57.25	[Could not connect: Got an unknown RCPT TO response: 550 5.5.3 ... Invalid id]
smtp2.google.com.	- 64.239.167.25	[Could not connect: Got an unknown RCPT TO response: 550 5.5.3 ... Invalid id]
smtp3.google.com.	- 64.239.189.25	[Could not connect: Got an unknown RCPT TO response: 550 5.5.3 ... Invalid id]

Figure 8. L'erreur retournée par le service DNSSstuff.com après la saisie d'une adresse email incorrecte

À propos de l'auteur

Błażej Kantak travaille en tant que *network troubleshooter* pour une grande institution financière. Outre les réseaux qui sont sa spécialité, il s'occupe aussi des questions relatives à la sécurité informatique, en particulier au Wi-Fi, VPN, FW, VoIP et à la *compromission des périphériques* Cisco. Son dernier succès du domaine Physical Security était une attaque DoS réussie sur l'ascenseur.

- verrouiller la possibilité de transférer le fichier de zone à partir du serveur DNS ;
- la résolution inverse DNS ne doit être appliquée dans des cas particuliers ;
- utiliser les noms DNS qui ne suggèrent pas à quoi sert un serveur donné. Il est recommandé

- d'appliquer une nomenclature contenant les données identifiant l'hôte, mais qui ne seront pas compris pour les personnes en dehors de l'organisation ;
- limiter au minimum la diffusion des bannières de services (par exemple SMTP, Web, etc...) ou modifier leur contenu de façon

à ce qu'elles suggèrent un autre système. Parfois, cette opération nécessitera des modifications dans les fichiers de configuration, et même dans le code source, si vous en disposez (open source) ;

- désactiver les messages d'erreur retournés sur les pages Web. Ces erreurs peuvent être appelées par l'intrus à l'aide des données d'entrée incorrectes et peuvent révéler des détails concernant l'application Web ;
- si une page contenant les données critiques ou confidentielles a été indexée dans le navigateur, il faut contacter le service technique pour supprimer cette page de la base et du cache ;
- ne pas autoriser les informations détaillées sur les projets actuels et terminés appliqués dans l'infrastructure informatique ;
- ne pas utiliser robots.txt – au lieu, de déterminer l'autorisation auprès des pages Web critiques avec le chiffage SSL.

Conclusion

Le processus de la protection des systèmes informatiques ne finit pas par la configuration d'un pare-feu, l'application d'un correctif au serveur de messagerie, la mise à jour de la base antivirus et l'enregistrement des fichiers journaux. En fait, ce processus n'en finit jamais. Dans chaque situation, il faut bien réfléchir si le fait de rendre publique trop d'informations (par exemple sur la mise en oeuvre du système PKI dans l'entreprise) ne violera pas les principes de la politique de sécurité, et de cela, la sécurité des employés et des ressources de la société.

Dans cet article, j'ai tenté de démontrer comment les données, à première vue, banales et peu importantes, peuvent mener à une compromission du système. Grâce à ces données, l'intrus est capable de toucher notre point le plus faible. N'oubliez pas la protection d'une infrastructure informatique signifie que tous ses éléments soient résistants à l'attaque. Pour l'intrus, un seul élément faible suffit. Et le plus c'est souvent c'est un homme. ●