Niebezpieczne Google – wyszukiwanie poufnych informacji

Michał Piotrowski

Artykuł opublikowany w numerze 3/2005 magazynu *hakin9*

Magazyn *hakin9*, Wydawnictwo Software, ul. Lewartowskiego 6, 00-190 Warszawa, *pl@hakin9.org*

Niebezpieczne Google – wyszukiwanie poufnych informacji

Michał Piotrowski

Informacje które powinny być chronione, bardzo często są dostępne publicznie. Ujawniają je nieświadomie – na skutek niedbalstwa lub niewiedzy – sami użytkownicy. Efekt jest taki, że poufne dane są na wyciągnięcie ręki, w Internecie. Wystarczy użyć Google.

Google odpowiada na około 80 procent wszystkich zapytań w Sieci, a tym samym jest najczęściej i najchętniej wykorzystywaną wyszukiwarką. Zawdzięcza to nie tylko wyjątkowo skutecznemu mechanizmowi generowania wyników, ale też bardzo rozbudowanym możliwościom zadawania pytań. Jednak należy pamiętać o tym, że Internet jest bardzo dynamicznym medium, przez co wyniki prezentowane przez Google nie zawsze są aktualne. Zdarza się, że niektóre odnalezione strony są mocno nieświeże, a jednocześnie wiele podobnych jeszcze nie zostało odwiedzonych przez Googlebota (skrypt-automat przeczesujący i indeksujący zasoby WWW).

Najważniejsze i najbardziej przydatne operatory precyzujące, wraz z opisem i efektem działania, zostały przedstawione w Tabeli 1, zaś miejsca w dokumentach, do których operatory się odnoszą w trakcie przeszukiwania zasobów Sieci (na przykładzie strony magazynu *hakin9*), prezentuje Rysunek 1. To tylko przykłady – umiejętne zadawanie pytań w Google umożliwia uzyskanie o wiele ciekawszych informacji.

Szukamy ofiary

Dzięki wyszukiwarce Google można dotrzeć nie tylko do powszechnie dostępnych zasobów

Z artykułu dowiesz się...

- jak przy użyciu Google wyszukiwać bazy danych osobowych i inne poufne informacje,
- jak odnaleźć informacje o podatnych na ataki systemach i usługach sieciowych,
- jak znajdować w Google publicznie dostępne urządzenia sieciowe.

Co powinieneś wiedzieć...

- powinieneś potrafić korzystać z przeglądarki internetowej,
- powinieneś mieć podstawową wiedzę o protokole HTTP.

O autorze

Michał Piotrowski, magister informatyki, ma wieloletnie doświadczenie w pracy na stanowisku administratora sieci i systemów. Przez ponad trzy lata pracował jako inspektor bezpieczeństwa w instytucji obsługującej nadrzędny urząd certyfikacji w polskiej infrastrukturze PKI. Obecnie specjalista ds. bezpieczeństwa teleinformatycznego w jednej z największych instytucji finansowych w Polsce. W wolnych chwilach programuje i zajmuje się kryptografią.

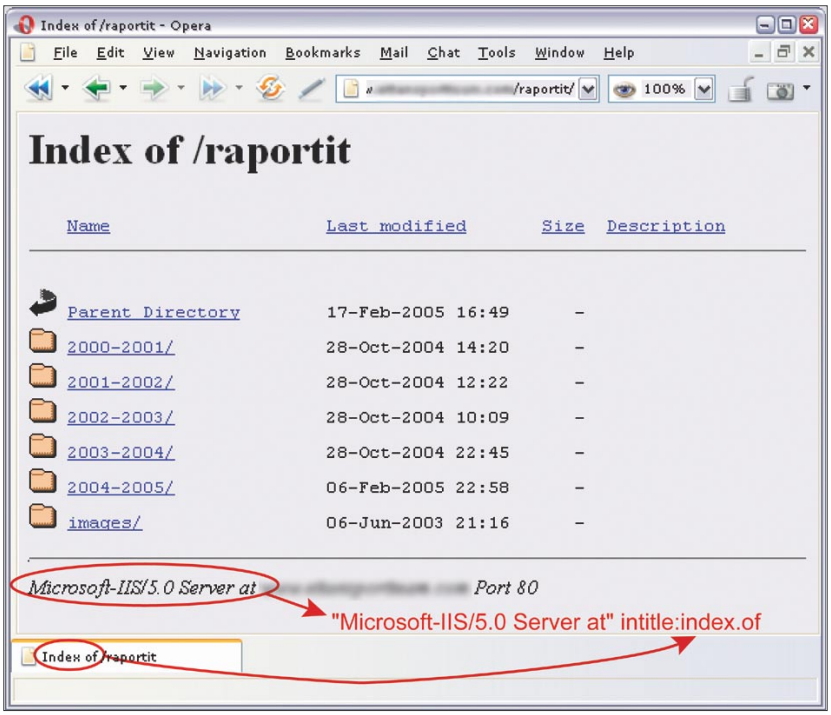
Tabela 1. *Operatory zapytań w Google*

Operator	Przeznaczenie	Przykład wykorzystania
`site`	ogranicza wyniki do stron znajdujących się w określonej domenie	`site:google.com fox` znajdzie wszystkie strony zawierające w tekście wyraz *fox*, które znajdują się w domenie **.google.com*
`intitle`	ogranicza wyniki do dokumentów zawierających podaną frazę w tytule	`intitle:fox fire` znajdzie strony zawierające wyraz *fox* w tytule i *fire* w tekście
`allintitle`	ogranicza wyniki do dokumentów zawierających wszystkie podane frazy w tytule	`allintitle:fox fire` znajdzie wszystkie strony zawierające w tytule wyrazy *fox* i *fire*; działa podobnie jak `intitle:fox intitle:fire`
`inurl`	ogranicza wyniki do stron zawierających podaną frazę w adresie URL	`inurl:fox fire` znajdzie strony zawierające w tekście wyraz *fire* i *fox* w adresie URL
`allinurl`	ogranicza wyniki do stron zawierających wszystkie podane frazy w adresie URL	`allinurl:fox` fire znajdzie strony zawierające w adresie URL wyrazy *fox* i *fire*; działa podobnie jak `inurl: fox inurl:fire`
`filetype, ext`	ogranicza wyniki do dokumentów o podanym typie	`filetype:pdf fire` zwróci dokumenty PDF zawierające wyraz *fire*, a `filetype:xls fox` zwróci dokumenty arkusza *Excel* zawierające *fox*
`numrange`	ogranicza wyniki do dokumentów zawierających w treści liczbę z podanego zakresu	`numrange:1-100 fire` zwróci strony zawierające liczbę z zakresu od 1 do 100 i wyraz *fire*. Identyczny efekt można uzyskać pytaniem: `1..100 fire`
`link`	ogranicza wyniki do stron zawierających odnośniki do podanej lokalizacji	`link:www.google.pl` zwróci dokumenty zawierające co najmniej jeden odnośnik do strony *www.google.pl*
`inanchor`	ogranicza wyniki do stron z odnośnikami zawierającymi w opisie podaną frazę	`inanchor:fire` zwróci dokumenty zawierające odnośniki, które posiadają wyraz *fire* w opisie (nie w adresie URL, na który wskazują, ale w podkreślonej części tekstu)
`allintext`	ogranicza wyniki do dokumentów zawierających podaną frazę w tekście i jednocześnie nie zawierające jej w tytule, odnośnikach i adresach URL	`allintext:"fire fox"` zwróci dokumenty, które posiadają frazę *fire fox* tylko w tekście
`+`	wymusza częste występowanie podanej frazy w wynikach	`+fire` segreguje wyniki zgodnie w dużą ilością występowania wyrazu *fire*
`-`	wymusza niewystępowanie podanej frazy w wynikach	`-fire` zwróci dokumenty nie zawierające wyrazu *fire*
`""`	pozwala wyszukiwać całe frazy, nie tylko wyrazy	`"fire fox"` zwróci dokumenty zawierające frazę *fire fox*
`.`	jest zastępowany pojedynczym znakiem	`fire.fox` zwróci dokumenty zawierające frazy *fire fox*, *fireAfox*, *fire1fox*, *fire-fox* itp.
`*`	jest zastępowany pojedynczym wyrazem	`fire * fox` zwróci dokumenty zawierające frazy *fire the fox*, *fire in fox*, *fire or fox* itp.
`|`	logiczne OR	`"fire fox" | firefox` zwróci dokumenty zawierające frazę *fire fox* lub wyraz *firefox*

Internetu, ale również do takich, które nigdy nie powinny zostać ujawnione. Jeśli zadamy odpowiednie pytanie, często otrzymamy naprawdę zadziwiające wyniki. Zacznijmy od czegoś prostego.

Wyobraźmy sobie, że w pewnym powszechnie wykorzystywanym programie zostaje odnaleziona luka. Przypuśćmy, że dotyczy ona serwera Microsoft IIS w wersji 5.0 i że hipotetyczny napastnik chce znaleźć kilka maszyn z tym oprogramowaniem, aby je zaatakować. Oczywiście mógłby do tego celu użyć jakiegoś skanera, jednak woli skorzystać z Google – wpisuje więc następujące pytanie: `"Microsoft-IIS/5.0`

Rysunek 1. *Wykorzystanie operatorów w wyszukiwaniu na przykładzie witryny magazynu hakin9*



Rysunek 2. *Odnalezienie serwera IIS 5.0 przy użyciu operatora intitle*

`Server at" intitle:index.of` i w rezultacie otrzymuje odnośniki do poszukiwanych serwerów, a konkretnie do wylistowanych zawartości katalogów, znajdujących się na tych serwerach. Dzieje się tak dlatego, że w standardowej konfiguracji oprogramowanie IIS (i wiele innych) dodaje do niektórych dynamicznie generowanych stron banery zawierające swoją nazwę i wersję (widać to na Rysunku 2).

Jest to przykład informacji, która sama w sobie jest niegroźna; z tego powodu bardzo często jest ignorowana i pozostawiana w standardowej konfiguracji. Niestety jest to również informacja, która w pewnych okolicznościach może mieć dla napastnika istotne znaczenie. Więcej przykładowych pytań dla Google o inne typy serwerów zawiera Tabela 2.

Innym sposobem znalezienia konkretnych wersji serwerów WWW jest szukanie standardowych stron, które są z nimi dostarczane i dostępne po poprawnej instalacji. Może wydawać się to dziwne, ale w Sieci znajduje się mnóstwo serwerów, których domyślna zawartość nie została zmieniona po instalacji. Bardzo często są to słabo zabezpieczone, zapomniane maszyny stanowiące łatwy cel dla włamywaczy. Można je odnaleźć wykorzystując pytania zaprezentowane w Tabeli 3.

Ta metoda jest bardzo prosta i jednocześnie niezwykle użyteczna. Za jej pomocą można uzyskać dostęp do ogromnej ilości różnych serwisów sieciowych czy systemów operacyjnych wykorzystujących aplikacje, w których znaleziono błędy i których leniwi lub nieświadomi administratorzy nie usunęli. Za przykład niech posłużą dwa dosyć popularne programy: *WebJeff Filemanager* i *Advanced Guestbook*.

Pierwszy z nich jest webowym menadżerem plików, umożliwiającym przesyłanie plików do serwera oraz tworzenie, przeglądanie, usuwanie i modyfikowanie plików znajdujących się na serwerze. Niestety, *WebJeff Filemanager* w wersji 1.6 ma błąd, który umożliwia odczytanie zawartości dowolnego pliku znajdującego się na serwerze, do którego ma dostęp użytkownik uruchamiający demona WWW. Wystarczy więc, że intruz wpisze w podatnym systemie adres */index.php3?action=telecharger&fichier=/etc/passwd*, a uzyska zawartość pliku */etc/passwd* (patrz Rysunek 3). Oczywiście aby znaleźć podatne serwery napastnik wykorzysta Google zadając pytanie: `"WebJeff-Filemanager 1.6" Login`.

Druga aplikacja – *Advanced Guestbook* – jest napisanym w języku PHP programem korzystającym z bazy danych SQL, który umożliwia dodawanie ksiąg gości do serwisów WWW. W kwietniu 2004 roku została opublikowana informacja o luce dotyczącej wersji 2.2 tego programu, która umożliwia (dzięki wstrzyk-

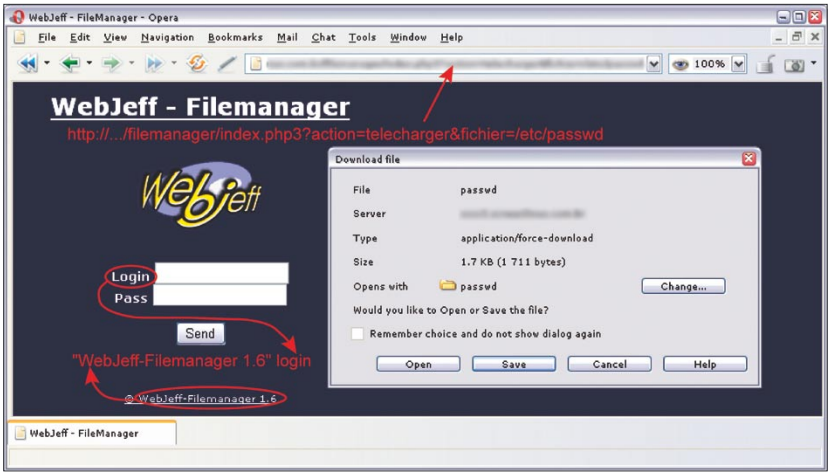
Tabela 2. *Google – pytania o różne rodzaje serwerów WWW*

Pytanie	Serwer
`"Apache/1.3.28 Server at" intitle:index.of`	Apache 1.3.28
`"Apache/2.0 Server at" intitle:index.of`	Apache 2.0
`"Apache/* Server at" intitle:index.of`	dowolna wersja Apache
`"Microsoft-IIS/4.0 Server at" intitle:index.of`	Microsoft Internet Information Services 4.0
`"Microsoft-IIS/5.0 Server at" intitle:index.of`	Microsoft Internet Information Services 5.0
`"Microsoft-IIS/6.0 Server at" intitle:index.of`	Microsoft Internet Information Services 6.0
`"Microsoft-IIS/* Server at" intitle:index.of`	dowolna wersja Microsoft Internet Information Services
`"Oracle HTTP Server/* Server at" intitle:index.of`	dowolna wersja serwera Oracle
`"IBM _ HTTP _ Server/* * Server at" intitle:index.of`	dowolna wersja serwera IBM
`"Netscape/* Server at" intitle:index.of`	dowolna wersja serwera Netscape
`"Red Hat Secure/*" intitle:index.of`	dowolna wersja serwera Red Hat Secure
`"HP Apache-based Web Server/*" intitle:index.of`	dowolna wersja serwera HP

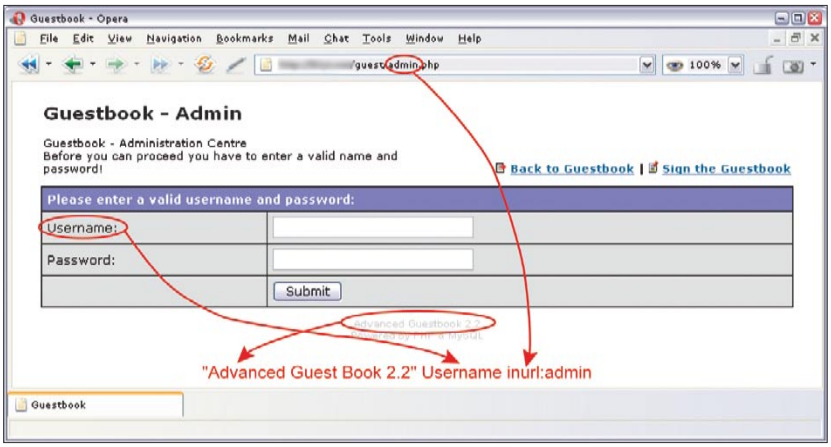
Tabela 3. *Pytania o standardowe poinstalacyjne strony serwerów WWW*

Pytanie	Serwer
`intitle:"Test Page for Apache Installation" "You are free"`	Apache 1.2.6
`intitle:"Test Page for Apache Installation" "It worked!" "this Web site!"`	Apache 1.3.0–1.3.9
`intitle:"Test Page for Apache Installation" "Seeing this instead"`	Apache 1.3.11–1.3.33, 2.0
`intitle:"Test Page for the SSL/TLS-aware Apache Installation" "Hey, it worked!"`	Apache SSL/TLS
`intitle:"Test Page for the Apache Web Server on Red Hat Linux"`	Apache w systemie Red Hat
`intitle:"Test Page for the Apache Http Server on Fedora Core"`	Apache w systemie Fedora
`intitle:"Welcome to Your New Home Page!" Debian`	Apache w systemie Debian
`intitle:"Welcome to IIS 4.0!"`	IIS 4.0
`intitle:"Welcome to Windows 2000 Internet Services"`	IIS 5.0
`intitle:"Welcome to Windows XP Server Internet Services"`	IIS 6.0

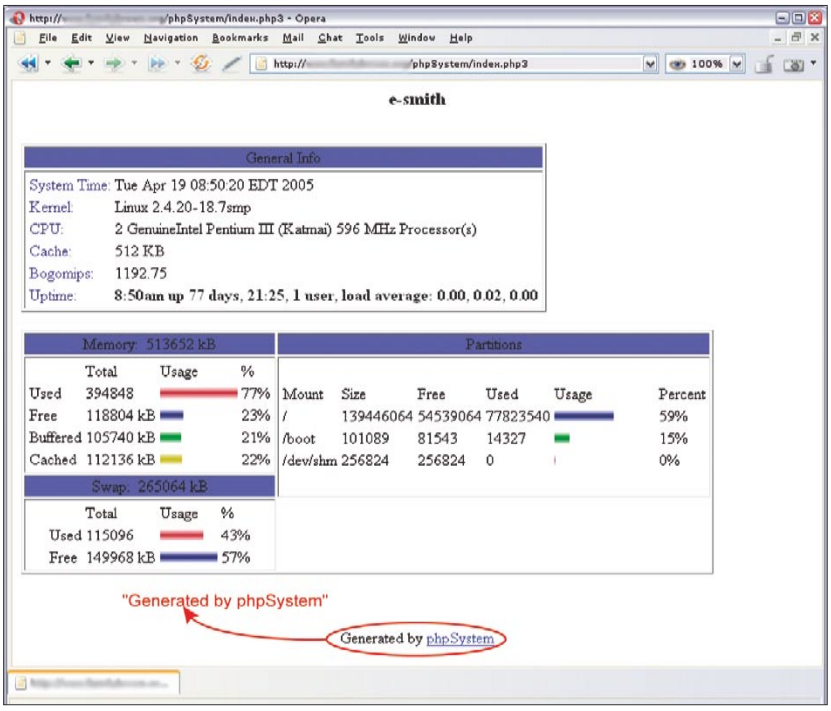
nięciu kodu SQL – patrz Artykuł *Ataki SQL Injection na PHP/MySQL* w *hakin9* 2/2005) uzyskanie dostępu do panelu administracyjnego. Wystarczy odnaleźć stronę logowania do panelu (patrz Rysunek 4) i zalogować się pozostawiając pole *username* puste, a w polu *password* wpisując `') OR ('a' = 'a`, lub odwrotnie – pole *password* zostawiając puste, a w polu *username* wpisując `? or 1=1 --`. Nasz przykładowy napastnik, aby znaleźć w sieci podatne witryny, może zadać wyszukiwarce Google jedno z następujących pytań: `intitle:`



Rysunek 3. *Podatna wersja programu WebJeff Filemanager*



Rysunek 4. *Advanced Guestbook – strona logowania*



Rysunek 5. *Statystyki phpSystem*

`Guestbook "Advanced Guestbook 2.2 Powered"` lub `"Advanced Guestbook 2.2" Username inurl:admin`.

Aby zapobiec działającemu w opisany sposób wyciekowi danych, administrator musi na bieżąco śledzić informacje o wszystkich programach, które wykorzystuje w utrzymywanych przez siebie serwisach i dokonywać aktualizacji w razie pojawienia się błędu w którymkolwiek z nich. Drugą rzeczą, o którą warto zadbać jest usunięcie banerów, nazw i numerów wersji programów ze wszystkich stron lub plików, w których występują.

Informacje o sieciach i systemach

Prawie każdy atak na system komputerowy jest poprzedzony rozpoznaniem celu. Zazwyczaj polega to na skanowaniu komputerów – próbie określenia działających usług, rodzaju systemu operacyjnego i wersji oprogramowania usługowego. Najczęściej wykorzystuje się do tego celu skanery typu *Nmap* lub *amap*, ale istnieje jeszcze inna możliwość. Wielu administratorów instaluje aplikacje WWW, które na bieżąco generują statystyki z pracy systemu, informują o zajętości dysków twardych, zawierają listy uruchomionych procesów lub nawet logi systemowe.

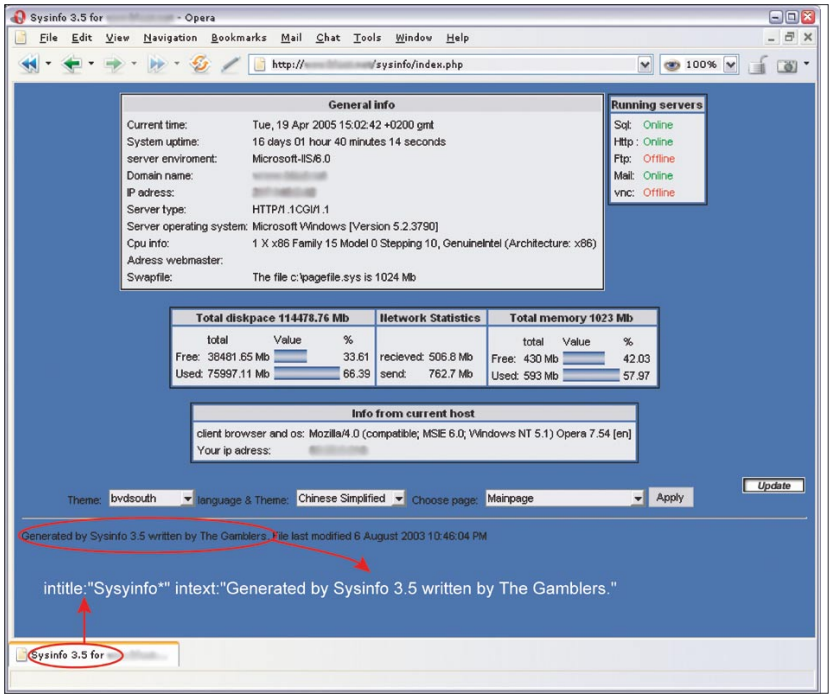
Dla włamywacza są to bardzo wartościowe informacje. Wystarczy, że zapyta Google o statystyki programu *phpSystem*: `"Generated by phpSystem"`, a otrzyma strony podobne do zaprezentowanej na Rysunku 5. Może również zapytać o strony generowane przez skrypt *Sysinfo*: `intitle:"Sysinfo * " intext: "Generated by Sysinfo * written by The Gamblers."`; które zawierają znacznie więcej informacji o systemie (Rysunek 6).

Możliwości jest bardzo dużo (przykładowe zapytania o statystyki i informacje tworzone przez najpopularniejsze programy zawiera Tabela 4). Zdobycie tego typu informacji może zachęcić intruza do przeprowadzenia ataku na znaleziony system i pomóc mu w dobraniu odpowiednich narzędzi czy exploitów. Dlatego, jeśli korzysta-

my z programów umożliwiających monitorowanie zasoby naszych komputerów, musimy zadbać o to, aby dostęp do nich był chroniony i wymagał podania hasła.

Szukamy błędów

Komunikaty o błędach mogą być dla włamywacza niezwykle wartościowe – właśnie z tych informacji można otrzymać mnóstwo danych o systemie oraz konfiguracji i budowie baz danych. Przykładowo, aby odnaleźć błędy generowane przez bazę *Informix* wystarczy zadać wyszukiwarce następujące pytanie: `"A syntax error has occurred" filetype:ihtml`. W rezultacie włamywacz odnajdzie komunikaty zawierające informacje o konfiguracji bazy danych, układzie plików w systemie a czasem również hasła (patrz Rysunek 7). Aby zawęzić wyniki tylko do stron zawierają-



Rysunek 6. *Statystyki Sysinfo*

Tabela 4. *Programy tworzące statystyki pracy systemu*

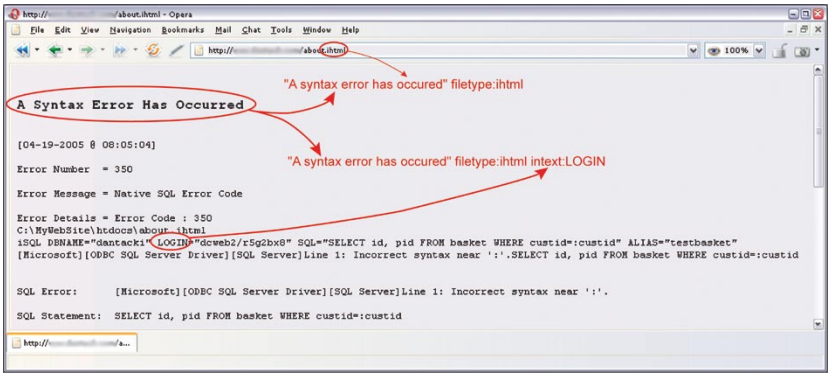
Pytanie	Rodzaj informacji
`"Generated by phpSystem"`	rodzaj i wersja systemu operacyjnego, konfiguracja sprzętowa, zalogowani użytkownicy, otwarte połączenia, zajętość pamięci i dysków twardych, punkty montowania
`"This summary was generated by wwwstat"`	statystyki pracy serwera WWW, układ plików w systemie
`"These statistics were produced by getstats"`	statystyki pracy serwera WWW, układ plików w systemie
`"This report was generated by WebLog"`	statystyki pracy serwera WWW, układ plików w systemie
`intext:"Tobias Oetiker" "traffic analysis"`	statystyki pracy systemu w postaci wykresów MRTG, konfiguracja sieci
`intitle:"Apache::Status" (inurl:server-status | inurl:status.html | inurl:apache.html)`	wersja serwera, rodzaj systemu operacyjnego, lista procesów potomnych i aktualne połączenia
`intitle:"ASP Stats Generator *.*" "ASP Stats Generator" "2003-2004 weppos"`	aktywność serwera WWW, dużo informacji o odwiedzających
`intitle:"Multimon UPS status page"`	statystyki pracy urządzeń UPS
`intitle:"statistics of" "advanced web statistics"`	statystyki pracy serwera WWW, informacje o odwiedzających
`intitle:"System Statistics" +"System and Network Information Center"`	statystyki pracy systemu w postaci wykresów MRTG, konfiguracja sprzętowa, działające usługi
`intitle:"Usage Statistics for" "Generated by Webalizer"`	statystyki pracy serwera WWW, informacje o odwiedzających, układ plików w systemie
`intitle:"Web Server Statistics for ****"`	statystyki pracy serwera WWW, informacje o odwiedzających
`inurl:"/axs/ax-admin.pl" -script`	statystyki pracy serwera WWW, informacje o odwiedzających
`inurl:"/cricket/grapher.cgi"`	wykresy MRTG z pracy interfejsów sieciowych
`inurl:server-info "Apache Server Information"`	wersja i konfiguracja serwera WWW, rodzaj systemu operacyjnego, układ plików w systemie
`"Output produced by SysWatch *"`	rodzaj i wersja systemu operacyjnego, zalogowani użytkownicy, zajętość pamięci i dysków twardych, punkty montowania, uruchomione procesy, logi systemowe

cych hasła, można nieco zmodyfikować pytanie: `"A syntax error has occurred" filetype:ihtml intext:LOGIN`.

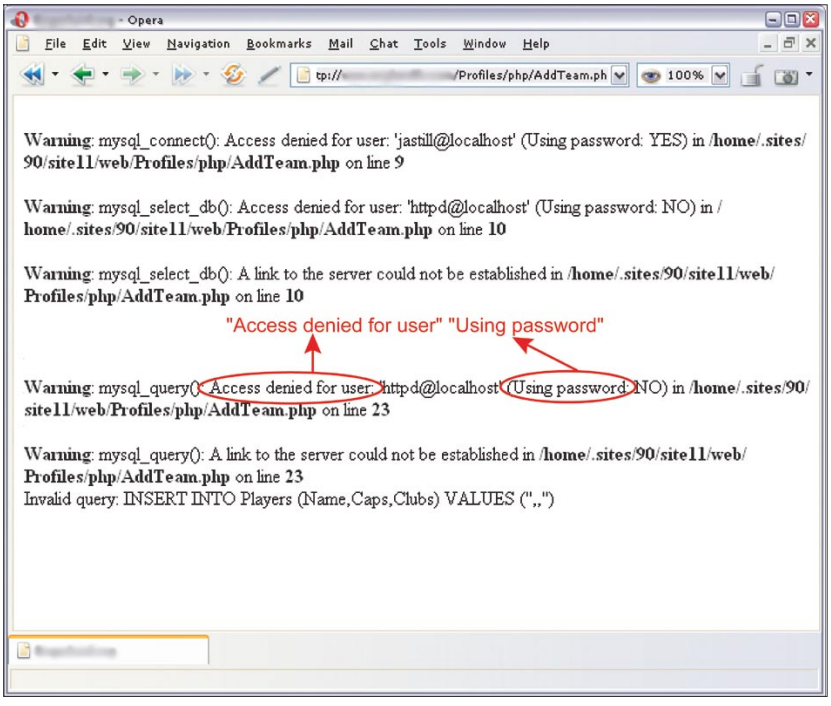
Równie ciekawe informacje można uzyskać z błędów bazy danych *MySQL*. Widać to choćby na przykładzie zapytania `"Access denied for user" "Using password"` – Rysunek 8 przedstawia jedną ze stron odnalezionych tym sposobem. Inne przykładowe pytania wykorzystujące takie błędy znajdują się w Tabeli 5.

Jedynym sposobem ochrony naszych systemów przed publicznym informowaniem o błędach jest przede wszystkim szybkie usuwanie nieprawidłowości oraz, jeśli mamy taką możliwość, skonfigurowanie oprogramowania w taki sposób, aby informacje o błędach były zapisywane w przeznaczonych specjalnie do tego celu plikach, a nie wysyłane na strony dostępne użytkownikom.

Należy przy tym pamiętać, że nawet jeśli błędy będziemy usuwać dosyć szybko (a tym samym powodować, że strony wskazywane przez Google będą już nieaktualne), to intruz może obejrzeć kopię strony przechowywaną przez *cache* wyszukiwarki Google. Wystarczy, że na liście z wynikami kliknie w odnośnik do kopii witryny. Na szczęście, ze względu na ogromną ilość zasobów internetowych, kopie są przechowywane w *cache* przez krótki okres.



Rysunek 7. *Błąd bazy danych Informix*



Rysunek 8. *Błąd bazy MySQL*

Szukamy haseł

W sieci można znaleźć mnóstwo haseł do wszelkiego rodzaju zasobów – kont pocztowych, serwerów FTP czy nawet kont shellowych. Wynika to głównie z niewiedzy użytkowników nieświadomie umieszczających hasła w publicznie dostępnych miejscach, ale też z niedbalstwa producentów oprogramowania, którzy albo nieodpowiednio chronią dane użytkowników, albo nie informują ich o konieczności modyfikacji standardowej konfiguracji swoich produktów.

Rozważmy przykład *WS_FTP*, dobrze znanego i powszechnie używanego klienta FTP, który podobnie jak większość oprogramowania użytkowego umożliwia zapamiętywanie haseł do kont. *WS_FTP* zapisuje swoją konfigurację i informacje o kontach użytkownika w pliku *WS_FTP.ini*. Niestety nie wszyscy zdajemy sobie sprawę z tego, że każdy, kto uzyska dostęp do konfiguracji klienta FTP będzie miał jednocześnie dostęp do naszych zasobów. Co prawda hasła przechowywane w pliku *WS_FTP.ini* są zaszyfrowane, ale nie jest to wystarczające zabezpieczenie – mając plik konfiguracyjny, włamywacz może skorzystać z narzędzi pozwalających hasła odszyfrować lub po prostu zainstalować program *WS_FTP* i uruchomić go z naszą konfiguracją. A w jaki sposób włamywacz może dotrzeć do tysięcy plików konfiguracyjnych klienta WS_FTP? Oczywiście przez Google. Dzięki pytaniom `"Index of/" "Parent Directory" "WS_FTP.ini"` lub `filetype:ini WS_FTP PWD` otrzyma mnóstwo odnośników do interesujących go danych, które w swojej niewiedzy sami wkładamy mu w ręce (Rysunek 9).

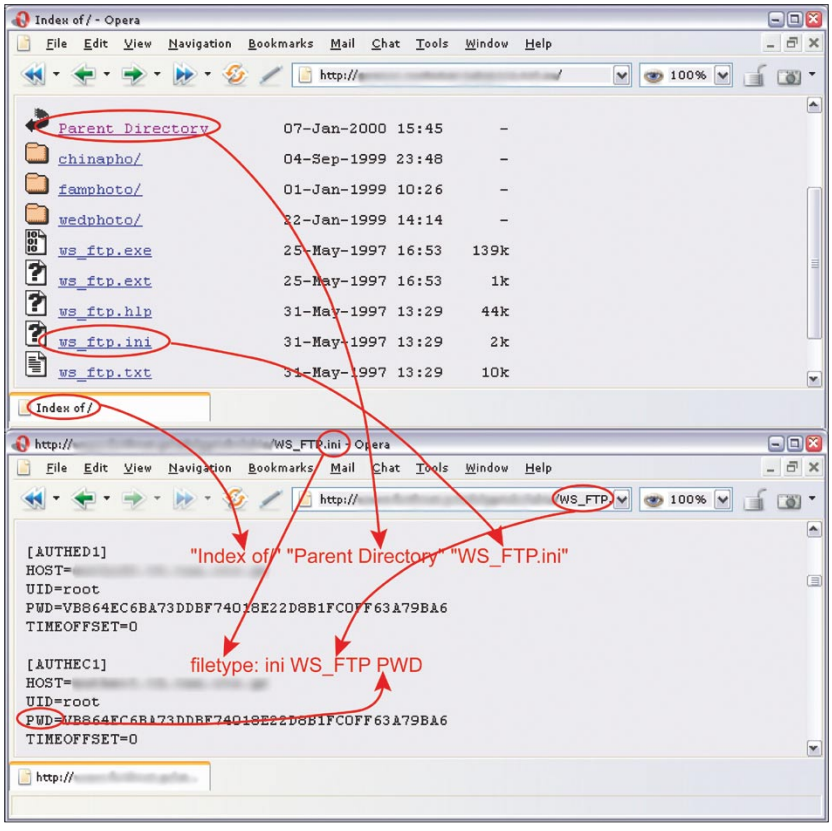
Inny przykład to aplikacja webowa o nazwie *DUclassified*, która umożliwia dodawanie i zarządzanie reklamami w serwisach internetowych. W standardowej konfiguracji tego programu nazwy użytkowników, hasła i inne dane są przechowywane w pliku *duclassified.mdb*, który znajduje się w niechronionym przed odczytem podkatalogu *_private*. Wystarczy zatem zna-

Tabela 5. *Komunikaty o błędach*

Pytanie	Rezultat
`"A syntax error has occurred" filetype:ihtml`	Błędy bazy Informix – mogą zawierać nazwy funkcji, nazwy plików, informacje o układzie plików, fragmenty kodu SQL oraz hasła
`"Access denied for user" "Using password"`	błędy autoryzacji – mogą zawierać nazwy użytkownika, nazwy funkcji, informacje o układzie plików i fragmenty kodu SQL
`"The script whose uid is " "is not allowed to access"`	błędy PHP związane z kontrolą dostępu – mogą zawierać nazwy plików, nazwy funkcji i informacje o układzie plików
`"ORA-00921: unexpected end of SQL command"`	błędy bazy Oracle – mogą zawierać nazwy plików, nazwy funkcji i informacje o układzie plików
`"error found handling the request" cocoon filetype:xml`	błędy programu *Cocoon* – mogą zawierać numer wersji *Cocoon*, nazwy plików, nazwy funkcji i informacje o układzie plików
`"Invision Power Board Database Error"`	błędy forum dyskusyjnego *Invision Power Board* – mogą zawierać nazwy funkcji, nazwy plików, informacje o układzie plików w systemie oraz fragmenty kodu SQL
`"Warning: mysql_query()" "invalid query"`	błędy bazy MySQL – mogą zawierać nazwy użytkowników, nazwy funkcji, nazwy plików i informacje o układzie plików
`"Error Message : Error loading required libraries."`	błędy skryptów CGI – mogą zawierać informacje o rodzaju systemu operacyjnego i wersji oprogramowania, nazwy użytkowników, nazwy plików oraz informacje o układzie plików w systemie
`"#mysql dump" filetype:sql`	błędy bazy MySQL – mogą zawierać informacje o strukturze i zawartości bazy danych

leźć serwis korzystający z *DUclassifield* o przykładowym adresie *http://<host>/duClassified/* i zmienić go na *http://<host>/duClassified/_private/duclassified.mdb*, aby otrzymać plik z hasłami i tym samym uzyskać nieograniczony dostęp do aplikacji (pokazuje to Rysunek 10). Natomiast w znalezieniu witryn, które korzystają z omawianej aplikacji może pomóc następujące pytanie zadane w Google: `"Powered by DUclassified" -site:duware.com` (aby uniknąć wyników dotyczących witryny producenta). Co ciekawe, producent *DUclassified* – firma DUware – stworzyła kilka innych aplikacji, które również są podatne na podobne nadużycia.

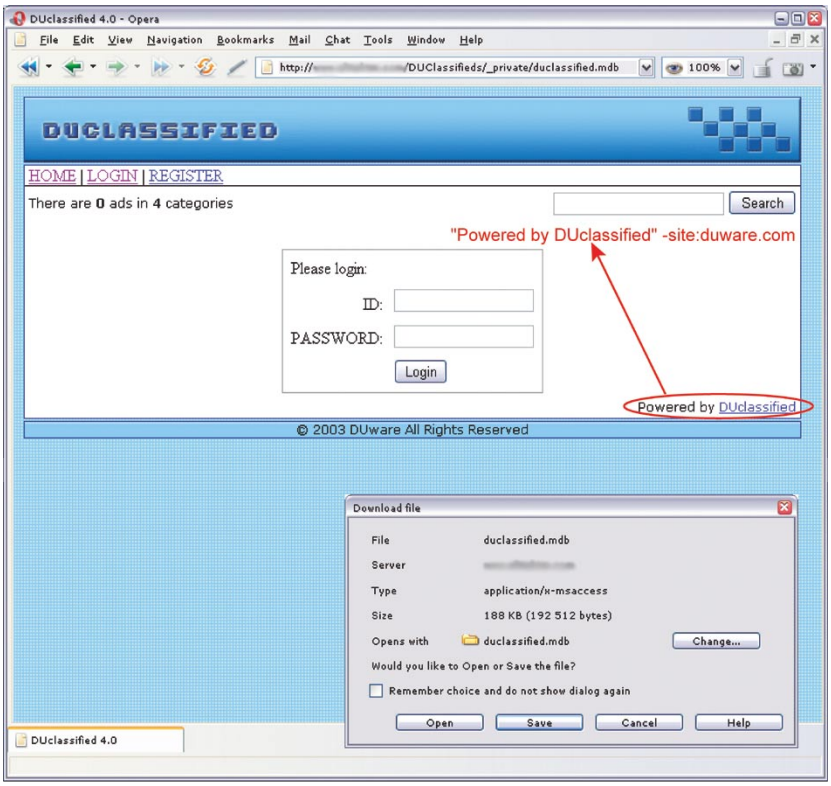
Teoretycznie wszyscy wiemy, że nie należy przylepiać haseł do monitora lub ukrywać ich pod klawiaturą. Tymczasem sporo ludzi zapisuje hasła w plikach i umieszcza je w swoich katalogach domowych, które, wbrew oczekiwaniom, są osiągalne z Internetu. W dodat-



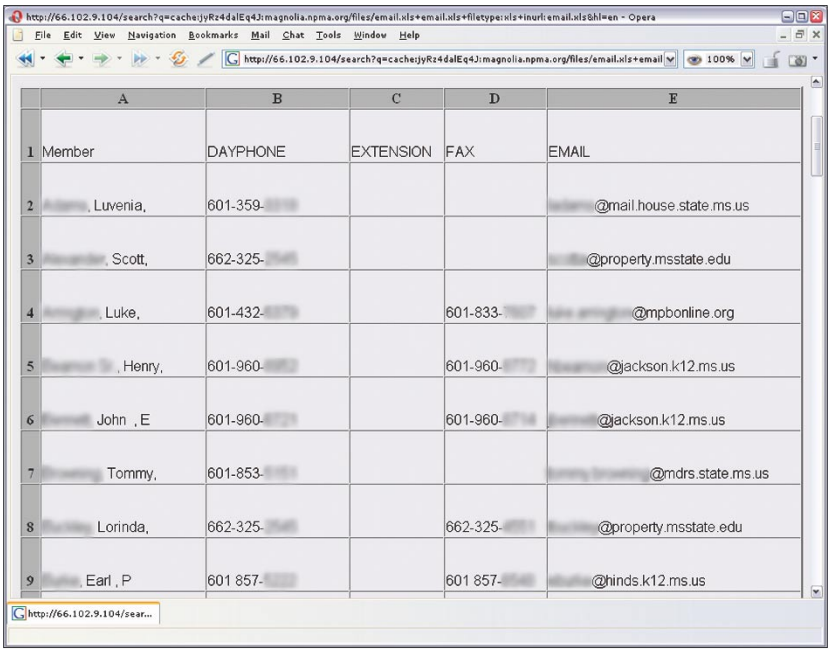
Rysunek 9. *Plik konfiguracyjny programu WS_FTP*

Tabela 6. *Hasła – przykładowe zapytania w Google*

Pytanie	Rezultat
`"http://*:*@www" site`	hasła do strony *site*, zapisane w postaci `"http://username:password@www..."`
`filetype:bak inurl:"htaccess|passwd|shadow|htusers"`	kopie zapasowe plików, w których mogą znajdować się informacje o nazwach użytkowników i hasła
`filetype:mdb inurl:"account|users|admin|administrators|passwd|password"`	pliki typu *mdb*, które mogą zawierać informacje o hasłach
`intitle:"Index of" pwd.db`	pliki *pwd.db* mogą zawierać nazwy użytkowników i zakodowane hasła
`inurl:admin inurl:backup intitle:index.of`	katalogi zawierające w nazwie słowa *admin* i *backup*
`"Index of/" "Parent Directory" "WS_FTP.ini" filetype:ini WS_FTP PWD`	pliki konfiguracyjne programu WS_FTP, które mogą zawierać hasła do serwerów FTP
`ext:pwd inurl:(service|authors|administrators|users) "# -FrontPage-"`	pliki zawierające hasła programu *Microsoft FrontPage*
`filetype:sql ("passwd values ****" | "password values ****" | "pass values ****" )`	pliki zawierające kod SQL i hasła dodawane do bazy danych
`intitle:index.of trillian.ini`	pliki konfiguracyjne komunikatora *Trillian*
`eggdrop filetype:user user`	pliki konfiguracyjne ircbota *Eggdrop*
`filetype:conf slapd.conf`	pliki konfiguracyjne aplikacji *OpenLDAP*
`inurl:"wvdial.conf" intext:"password"`	pliki konfiguracyjne programu *WV Dial*
`ext:ini eudora.ini`	pliki konfiguracyjne programu pocztowego *Eudora*
`filetype:mdb inurl:users.mdb`	pliki *Microsoft Access*, które mogą zawierać informacje o kontach
`intext:"powered by Web Wiz Journal"`	serwisy WWW korzystające z aplikacji *Web Wiz Journal*, która w standardowej konfiguracji umożliwia pobranie pliku zawierającego hasła; zamiast domyślnego adresu *http://<host>/journal/* należy wpisać *http://<host>/journal/journal.mdb*
`"Powered by DUclassified" -site:duware.com "Powered by DUcalendar" -site:duware.com "Powered by DUdirectory" -site:duware.com "Powered by DUclassmate" -site:duware.com "Powered by DUdownload" -site:duware.com "Powered by DUpaypal" -site:duware.com "Powered by DUforum" -site:duware.com intitle:dupics inurl:(add.asp | default.asp | view.asp | voting.asp) -site:duware.com`	serwisy WWW, korzystające z aplikacji *DUclassified*, *DUcalendar*, *DUdirectory*, *DUclassmate*, *DUdownload*, *DUpaypal*, *DUforum* lub *DUpics*, które w standardowej konfiguracji umożliwiają pobranie pliku zawierającego hasła; zamiast domyślnego adresu (dla *DUclassified*) *http://<host>/duClassified/* należy wpisać *http://<host>/duClassified/_private/duclassified.mdb*
`intext:"BiTBOARD v2.0" "BiTSHiFTERS Bulletin Board"`	serwisy WWW korzystające z aplikacji *Bitboard2*, która w standardowej konfiguracji umożliwia pobranie pliku zawierającego hasła; zamiast domyślnego adresu *http://<host>/forum/forum.php* należy wpisać *http://<host>/forum/admin/data_passwd.dat*



Rysunek 10. *Standardowo skonfigurowany program DUclassified*



	A	B	C	D	E
1	Member	DAYPHONE	EXTENSION	FAX	EMAIL
2	, Luvenia,	601-359-			@mail.house.state.ms.us
3	, Scott,	662-325-			@property.msstate.edu
4	, Luke,	601-432-		601-833-	@mpbonline.org
5	, Henry,	601-960-		601-960-	@jackson.k12.ms.us
6	John , E	601-960-		601-960-	@jackson.k12.ms.us
7	Tommy,	601-853-			@mdrs.state.ms.us
8	Lorinda,	662-325-		662-325-	@property.msstate.edu
9	Earl , P	601 857-		601 857-	@hinds.k12.ms.us

Rysunek 11. *Elektroniczna książka adresowa zdobyta dzięki Google*

W Sieci

- *http://johnny.ihackstuff.com* – największe repozytorium informacji o *Google hacking*,
- *http://insecure.org/nmap/* – skaner sieciowy *Nmap*,
- *http://thc.org/thc-amap/* – skaner *amap*.

ku wielu z nich piastuje funkcje administratorów sieci lub podobne, przez co pliki te osiągają pokaźne rozmiary. Trudno podać konkretną zasadę szukania takich danych, ale dobrze sprawdzają się kombinacje słów *account*, *users*, *admin*, *administrators*, *passwd*, *password* itp. W połączeniu z typami plików *.xls*, *.txt*, *.doc*, *.mdb* i *.pdf*. Warto również zwrócić uwagę na katalogi zawierające w nazwie słowa *admin*, *backup* lub podobne: `inurl:admin intitle:index.of`. Przykładowe pytania o dane związane z hasłami można znaleźć w Tabeli 6.

Aby utrudnić intruzom dostęp do naszych haseł, musimy przede wszystkim myśleć o tym, gdzie i po co je wpisujemy, jak są przechowywane oraz co się z nimi dzieje. Jeśli opiekujemy się serwisem internetowym, powinniśmy przeanalizować konfigurację wykorzystywanych aplikacji, odnaleźć słabo chronione lub wrażliwe dane i odpowiednio je zabezpieczyć.

Dane osobowe i dokumenty poufne

Zarówno w Polsce czy Unii Europejskiej, jak i w Stanach Zjednoczonych istnieją odpowiednie regulacje prawne, które mają za zadanie ochraniać naszą prywatność. Niestety zdarza się, że wszelkiego rodzaju poufne dokumenty zawierające nasze dane są umieszczane w publicznie dostępnych miejscach lub przesyłane przez sieć bez właściwego zabezpieczenia. Wystarczy, że intruz uzyska dostęp do poczty elektronicznej zawierającej nasze Curriculum Vitae wysłane podczas poszukiwania pracy, a pozna nasz adres, numer telefonu, datę urodzenia, przebieg edukacji, wiedzę i doświadczenie.

W Internecie można znaleźć mnóstwo dokumentów tego typu. Aby je odszukać, należy zadać następujące pytanie: `intitle: "curriculum vitae" "phone * * *" "address *" "e-mail"`. Łatwo również znaleźć dane teleadresowe w postaci list nazwisk, numerów telefonów i adresów e-mail (Rysu-

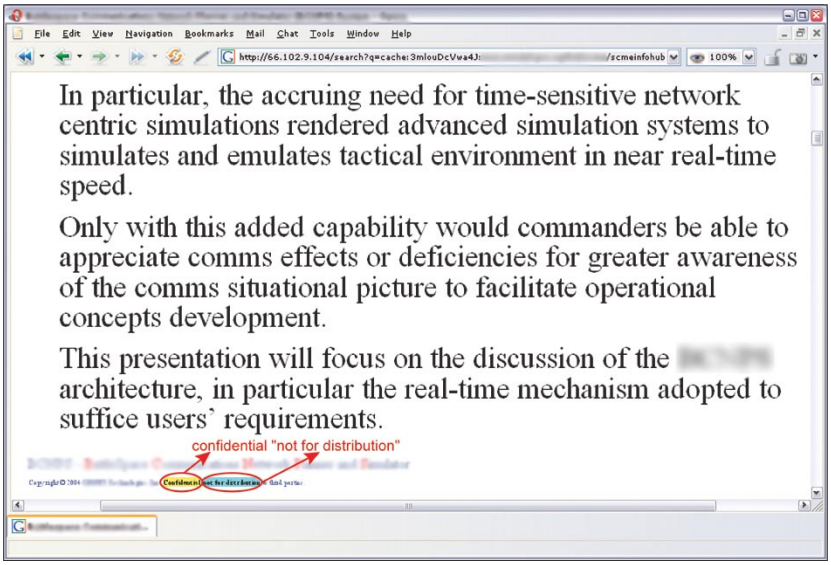
Tabela 7. *Szukanie danych osobowych i poufnych dokumentów*

Pytanie	Rezultat
`filetype:xls inurl:"email.xls"`	pliki *email.xls*, które mogą zawierać dane teleadresowe
`"phone * * *" "address *" "e-mail" intitle: "curriculum vitae"`	dokumenty CV
`"not for distribution" confidential`	dokumenty opatrzone klauzulą *confidential*
`buddylist.blt`	listy kontaktów komunikatora *AIM*
`intitle:index.of mystuff.xml`	listy kontaktów komunikatora *Trillian*
`filetype:ctt "msn"`	listy kontaktów *MSN*
`filetype:QDF QDF`	baza danych programu finansowego *Quicken*
`intitle:index.of finances.xls`	pliki *finances.xls*, które mogą zawierać informacje o kontach bankowych, zestawienia finansowe i numery kart kredytowych
`intitle:"Index Of" -inurl:maillog maillog size`	pliki *maillog*, które mogą zawierać wiadomości e-mail
`"Network Vulnerability Assessment Report" "Host Vulnerability Summary Report" filetype:pdf "Assessment Report" "This file was generated by Nessus"`	raporty z badania bezpieczeństwa sieci, testów penetracyjnych itp.

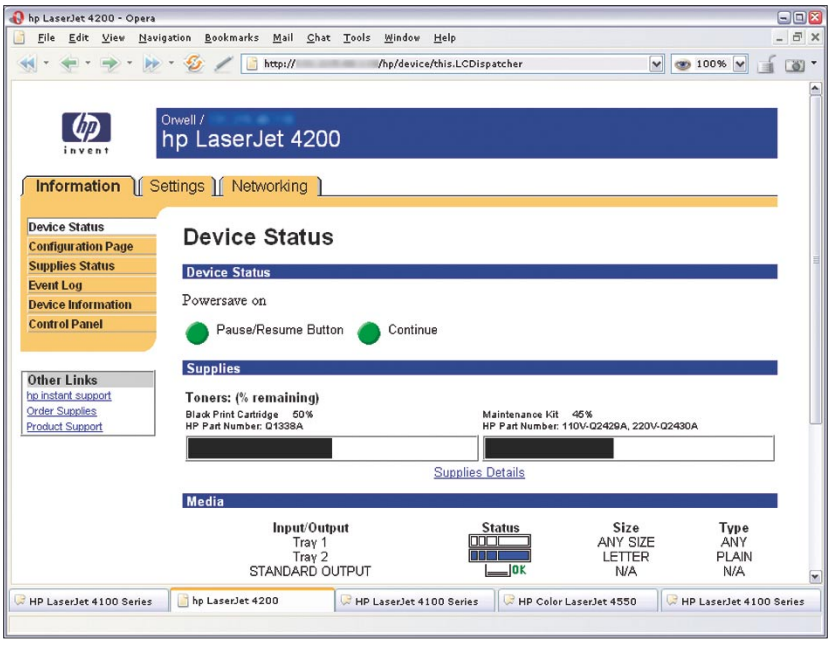
Tabela 8. *Ciągi charakterystyczne dla urządzeń sieciowych*

Pytanie	Urządzenie
`"Copyright (c) Tektronix, Inc." "printer status"`	drukarki PhaserLink
`inurl:"printer/main.html" intext:"settings"`	drukarki Brother HL
`intitle:"Dell Laser Printer" ews`	drukarki Della z technologią EWS
`intext:centreware inurl:status`	drukarki Xerox Phaser 4500/6250/8200/8400
`inurl:hp/device/this.LCDispatcher`	drukarki HP
`intitle:liveapplet inurl:LvAppl`	kamery Canon Webview
`intitle:"EvoCam" inurl:"webcam.html"`	kamery Evocam
`inurl:"ViewerFrame?Mode="`	kamery Panasonic Network Camera
`(intext:"MOBOTIX M1" | intext:"MOBOTIX M10") intext: "Open Menu" Shift-Reload`	kamery Mobotix
`inurl:indexFrame.shtml Axis`	kamery Axis
`SNC-RZ30 HOME`	kamery Sony SNC-RZ30
`intitle:"my webcamXP server!" inurl:":8080"`	kamery dostępne przez aplikację *WebcamXP Server*
`allintitle:Brains, Corp. camera`	kamery dostępne przez aplikację *mmEye*
`intitle:"active webcam page"`	kamery z interfejsem USB

nek 11). Wynika to z faktu, że prawie wszyscy użytkownicy Internetu tworzą różnego rodzaju elektroniczne książki adresowe – mają one niewielkie znaczenie dla przeciętnego intruza, ale już wprawiony socjotechnik będzie potrafił wykorzystać zawarte w nich dane, zwłaszcza jeśli dotyczą kontaktów w obrębie jednej firmy. Dosyć dobrze w tym przypadku sprawdza się na przykład pytanie: `filetype: xls inurl:"email.xls"`; które wyszuka arkusze kalkulacyjne o nazwie *email.xls*.



Rysunek 12. *Zastrzeżony dokument odnaleziony przez wyszukiwarkę*



Rysunek 13. *Odnaleziona przez Google strona konfiguracyjna drukarki HP*

Podobnie wygląda sytuacja z komunikatorami sieciowymi i zapisywanymi w nich listami kontaktów – po zdobyciu takiego zestawienia intruz będzie mógł próbować podszyć się pod naszych przyjaciół. Co ciekawe, dosyć dużo danych osobowych można znaleźć we wszelkiego rodzaju dokumentach urzędowych – raportach policyjnych, pismach sądowych czy nawet kartach przebiegu choroby.

W Sieci można również odnaleźć dokumenty, którym nadano jakąś klauzulę tajności i które tym samym zawierają zastrzeżone informacje. Mogą to być plany projektowe, dokumentacja techniczna, różne ankiety, raporty, prezentacje i całe mnóstwo innych wewnętrznych dokumentów firmowych. Można je znaleźć, gdyż bardzo często zawierają wyraz *confidential*, frazę *Not for distribution* lub podobne (patrz Rysunek 12). Tabela 7 zawiera kilka przykładowych pytań o dokumenty mogące zawierać dane osobowe i informacje poufne.

Tak jak w przypadku haseł, aby uniknąć ujawniania naszych prywatnych informacji możemy jedynie zachować ostrożność i panować nad publikowanymi danymi. Firmy i instytucje powinny (a w wielu przypadkach nawet muszą) opracować i wdrożyć odpowiednie regulaminy, procedury oraz zasady postępowania określające wewnętrzny obieg informacji, odpowiedzialność i konsekwencje za ich nieprzestrzeganie.

Urządzenia sieciowe

Wielu administratorów nie traktuje poważnie bezpieczeństwa takich urządzeń, jak drukarki sieciowe czy kamery webowe. Tymczasem źle zabezpieczona drukarka może być przyczółkiem, który włamywacz zdobywa w pierwszej kolejności, a potem wykorzystuje do przeprowadzania ataków na pozostałe systemy w sieci lub poza nią. Kamery internetowe oczywiście nie są aż tak niebezpieczne, więc można je traktować jako rozrywkę, jednak nietrudno wyobrazić sobie sytuację, kiedy takie dane miałyby znaczenie (szpiegostwo przemysłowe, napad rabunkowy). Pytania o drukarki i kamery zawiera Tabela 8, zaś Rysunek 13 prezentuje znalezioną w sieci stronę konfiguracyjną drukarki.

Poufne w indeksie

Dzisiejszy świat jest światem informacji, w którym wiedza jest cennym towarem – jej posiadanie przekłada się na realne korzyści, takie jak pieniądze, zaufanie klientów lub przewaga nad konkurencją. Informacja sama w sobie przybiera bardzo zróżnicowaną postać, coraz częściej elektroniczną.

Włamywaczy interesuje specyficzny rodzaj informacji – dane handlowe, wewnętrzne dokumenty firmowe, plany projektowe, dane teleadresowe, numery kart płatniczych, hasła... Wszystko można odnaleźć pod adresem *http://www.google.com*, w wyszukiwarce i jednocześnie zbiorze odnośników do wszystkiego, co może być dostępne w Sieci dla zwykłego użytkownika. Trzeba tylko pogooglać.■