Penetrante Suchmaschinen

Wo Angreifer über Suchmaschinen verwundbare Systeme oder vertrauliche Dateien finden, spricht man gemeinhin von "Google Hacking". Ein Überblick über typische "Verwundbarkeiten".

Von Sebastian Schreiber und Stefan Arbeiter, Tübingen

Längst nicht alles, was man über das Web erreichen kann, ist wirklich für die Öffentlichkeit bestimmt: Allzu oft lassen sich auch Passwortdateien, Konfigurationsseiten oder andere sensitive Daten und Services von jedermann abfragen. Die mehr oder weniger vollständige Erfassung des WWW durch Suchmaschinen liefert Hackern und Angreifern dazu umfassende Recherchemöglichkeiten. Der US-amerikanische Hacker und Sicherheitsspezialist Johnny Long hat hierfür die Bezeichnung "Google-Hacking" geprägt. Auch wenn derartige "globale Penetrationstests" gleichermaßen mit anderen Suchmaschinen funktionieren, hat sich der Begriff allgemein für die Verwendung einer Suchmaschine eingebürgert, um verwundbare Systeme, vertrauliche Daten oder kritische Anwendungen im Internet zu finden, ohne direkt auf die betroffenen Systeme zugreifen zu müssen.

Die Grundlage für diese Form des Hackens ist zum einen, dass Suchmaschinen gefundene Daten unabhängig davon, ob sie absichtlich oder versehentlich auf einem Webserver liegen, gleichgültig indizieren. Zum anderen "hilft" hierbei, dass sich heutzutage eine unüberschaubare Zahl von Produkten bequem über Browser benutzen und verwalten lassen (Web-Frontend/-GUI). Letztlich kommen auch dabei immer Webserver zum Einsatz und diese werden – sofern erreichbar – von den Suchmaschinen ebenfalls indiziert. Die einzige Kunst ist es also mit Geduld und Findigkeit passende Suchanfragen zu produzieren, um die genannten Seiten zu finden.

Ein konkretes und sehr klassisches Beispiel zu Beginn: Ein Angreifer sucht Sicherheitslücken, die er möglichst ohne großen Aufwand ausnutzen kann. Interessant sind dabei alle Dienste, die zur Verwaltung von Servern und Webseiten

dienen. Findet er einen solchen Dienst, der beispielsweise nicht mit einem Passwort geschützt ist, so benötigt er nur noch den entsprechenden Client, um darauf zuzugreifen, etwa Frontpage-Server-Extensions. Sehr alte Versionen der Extensions legten die gehashten Passworte der Anwender (und auch Administratoren) in verschiedenen Dateien auf dem jeweiligen Webserver ab. Der Fehler lag darin, dass ohne manuelle Änderungen die entsprechenden Verzeichnisse für jeden zugänglich waren – von den Extensions selbst wurden nach der Installation keine einschränkenden Berechtigungen gesetzt.

Daher können diese Passwortlisten sowohl von Hand als auch von Suchmaschinen gefunden werden. Die entsprechenden Dateien sind leicht zu identifizieren: Sie liegen in einem Verzeichnis /_vti_pvt/ und heißen zum Beispiel service.pwd, administrators.pwd oder passwords.pwd. Bei der Suche muss man lediglich darauf achten, dass man nicht auch Artikel und Konfigurationsanleitungen zu den Frontpage-Extensions findet. Um alle solchen Fundstellen auszuschließen und die Suche ausschließlich auf Dateiname und Pfad zu beschränken, gibt es den erweiterten Suchoperator "inurl:", der sowohl bei Google als auch Yahoo! verfügbar ist. Er beschränkt die Suche, wie der Name schon sagt, auf die Webadresse (Uniform Ressource Locator – URL).

Folgende Suchanfragen fördern die genannten Passwortlisten zutage:

```
inurl:/_vti_pvt/users.pwd
inurl:/_vti_pvt/administrators.pwd
inurl:/_vti_pvt/service.pwd
```

Sowohl die betroffenen Versionen der Frontpage-Extensions als auch die Tatsache, dass diese Dateien über Suchmaschinen leicht zu finden sind, sind knapp sechs Jahre alt. Trotzdem findet man auch heute noch derart konfigurierte Systeme online (vgl. Abb. 1).

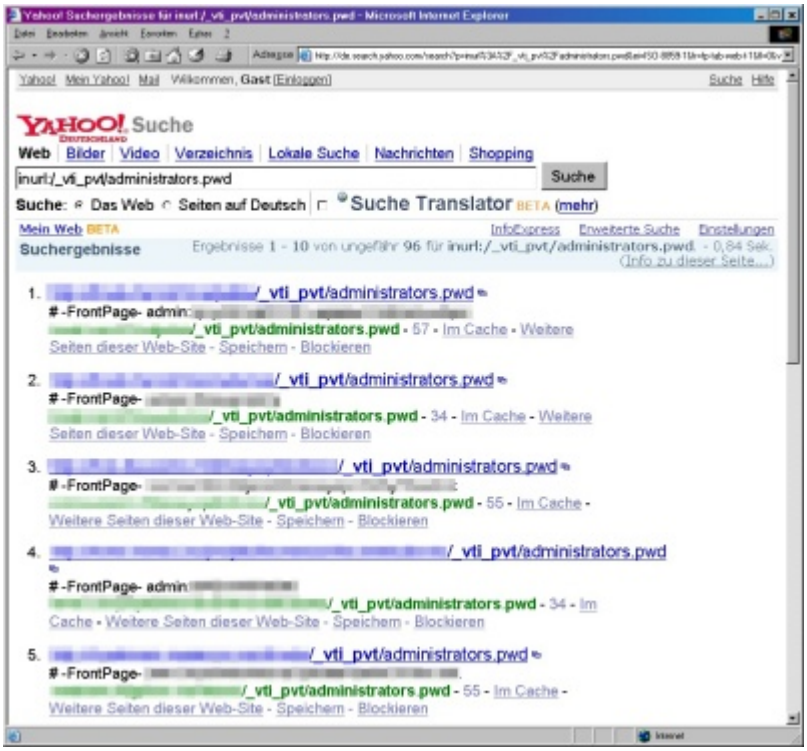


Abbildung 1: Auch für sechs Jahre alte Fehlkonfigurationen findet man im Internet noch immer Beispiele – im Bild: Suche nach Passwortdateien von Frontpage-Extensions.

Ein Angreifer muss anschließend nur noch einen Passwort-Knacker auf die Dateien ansetzen, um in den Besitz der Zugangsinformationen zu kommen (z. B. John the Ripper, www.openwall.com/john/). Zwar sind in diesem Fall das "Cracken" der Passwörter und ihre Verwendung illegal, jedoch stellt sich die Frage, ob die einfache Beschaffungsmethode für eine hohe Hemmschwelle bei potenziellen Tätern sorgt (das Betrachten der Suchergebnisse im eigenen Browser ist übrigens unproblematisch).

Ob man mithilfe eines "Google-Hacks" auch tatsächlich in das gefundene System eindringen kann, ist im konkreten Fall zu analysieren. Hat ein Angreifer jedoch ein System gefunden, auf dem tatsächlich noch sowohl die Extensions aktiv als auch die Passworte gültig sind, so muss er sich um die sonstige Sicherheit kaum Gedanken machen: Ein solches System wurde wahrscheinlich seit Jahren nicht aktualisiert.

Google-	**Bedeutung**

Keyword	
inurl:	sucht nach URLs, die den Such-String enthalten
allinurl:	sucht nach URLs , die sämtliche aufgelisteten Suchworte enthalten
intitle:	sucht nach Seiten, in deren Titel das Suchwort vorkommt
allintitle:	sucht nach Seiten, in deren Titel sämtliche Suchworte vorkommen
filetype:	sucht nach Dateien, die einen bestimmten Typ (Extension) haben, zum Beispiel filetype:xls für Excel-Dateien

Tabelle 1: Erweiterte Suchoperatoren ermöglichen in vielen Internet-Suchmaschinen gezielte Recherchen (z. B. in Google oder Yahoo!)

Offenherzige Hardware

Ebenso wie Software kann sich auch Hardware mit integrierten Webservern durch die URL verraten und damit Google-Hacking ermöglichen. Hier hat der Angreifer den Vorteil, dass derartige Webseiten meist vom Anwender nicht einmal verändert werden können: ihr Design ist statisch. Als Beispiel mögen die Webseiten der HP Jetdirect-Printserver dienen.

Gleich nach Aufruf der IP-Adresse eines derartigen Druckers erhält man die URL /hp/device/this.LCDispatcher, nach der man bequem suchen kann:

```
inurl:hp/device/this.LCDispatcher
```

Gesucht wird natürlich nicht nach den Druckern als solchen, sondern nach den Webservern, die in die Jetdirect-Systeme integriert sind. Falls diese über das Internet erreichbar sind, kann ein Angreifer in der Regel davon ausgehen, dass nicht nur die Drucker, sondern auch weitere Clients ebenfalls direkt übers Netz erreichbar sind (vgl. Abb 2 – wie für Netzwerk-Drucker typisch, gibt es jedoch nur recht wenige Suchergebnisse).

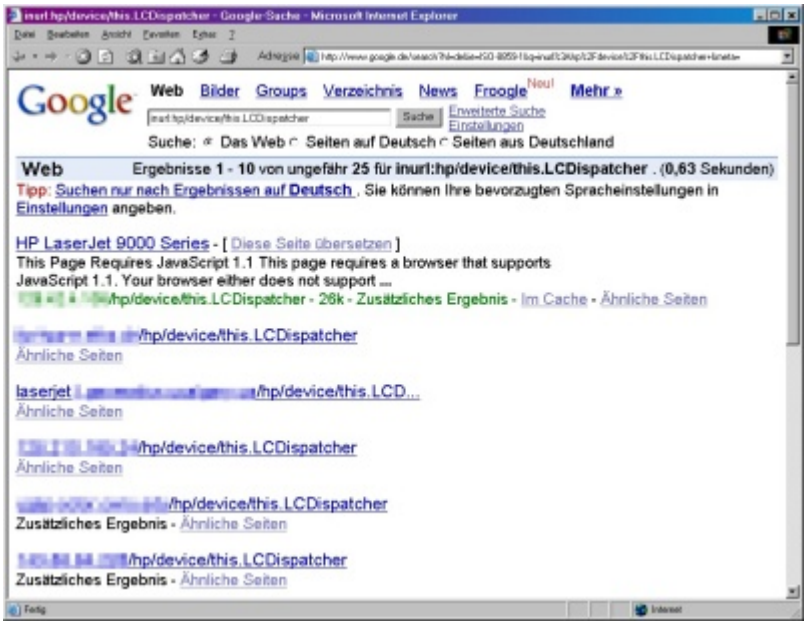


Abbildung 2: Auch Hardware stellt heutzutage häufig ihre Konfiguration über ein Web-Frontend bereit – im Bild: Suche nach Jetdirect-Printservern.

Würde man versuchen, diese Drucker mit einfachen Suchbegriffen zu erreichen, stieße man übrigens auf einige Schwierigkeiten:

- der Modellname ist je nach Druckertyp anders, obwohl die Webseiten des Jetdirect-Servers identisch sind,
- sowohl von HP als auch von Dritten werden viele Handbücher, Support-Seiten und Preislisten ins Internet gestellt, welche die entsprechenden Begriffe enthalten.

Die Verwendung der erweiterten Operatoren ist daher oft ein Muss. Umgekehrt liefern Drucker oder ähnliche Systeme, die in ihren Konfigurations-Servern banale Dateinamen nutzen, dementsprechend geringe Angriffsfläche für Recherchen über Suchmaschinen.

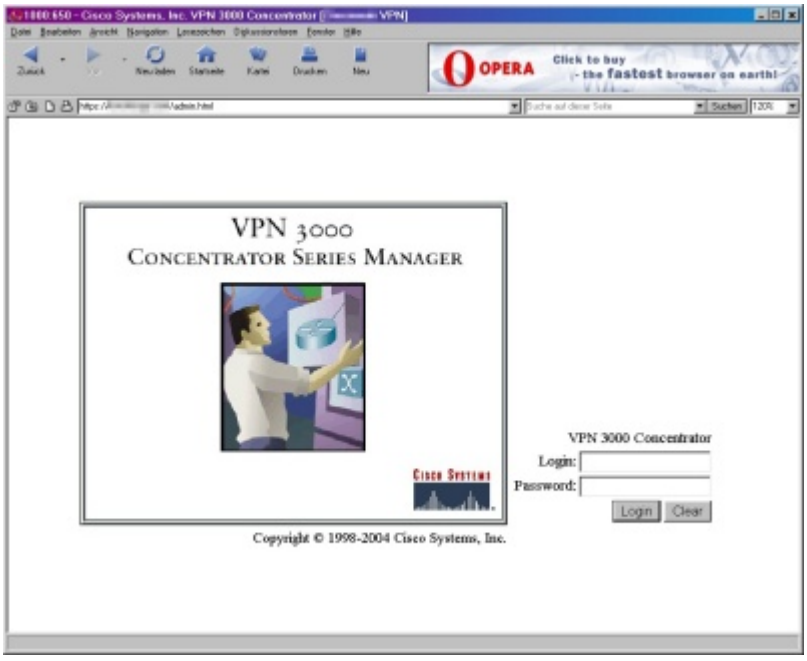
Schlecht getitelt

Neben der URL ist ein weiteres, schnell zum Ziel führendes Seitenelement häufig der Title-Tag. Ebenso wie bei der URL erlauben viele webbasierte Anwendungen und Webserver in Geräten nicht, den Seitentitel zu modifizieren. Ist er genau auf ein Produkt beschränkt, so kann man ihn jedoch auch via Suchmaschine recherchieren.

Kennt ein Angreifer zum Beispiel die Webseiten von Ciscos VPN Concentrator (vgl. Abb 3), die einen festen Titel enthalten, so kann er mit dem erweiterten Operator "intitle:" problemlos danach suchen:

```
intitle:"Cisco Systems, Inc. VPN 3000 Concentrator"
```

Derartige Login-Portale sind für Angreifer dahingehend interessant, dass sie jeweils nur so sicher sind wie die im Einsatz befindlichen Passworte. In der Regel versuchen (nicht auf ein konkretes Ziel hin agierende) Angreifer dort jedoch nur, eine Anzahl von Trivial-Passworten und Benutzernamen (admin usw.) zu testen – derartige Versuche sind allerdings strafbar, sofern sie von Erfolg gekrönt sind – (erfolglose) Versuche sind jedoch nicht strafbewehrt (§ 202a StGB).



VPN -Concentrator-Manager Login-Seite]">

Abbildung 3: Viele Systeme haben in ihren Konfigurationsseiten unveränderliche und signifikante Titel-Einträge, die sich via Suchmaschine leicht finden lassen – im Bild: "Cisco Systems, Inc. VPN 3000 Concentrator".

Da sich die Operatoren inurl: und intitle: auch kombinieren lassen, ergeben sich für eine große Menge an Anwendungen und Geräten signifikante Kombinationen. Dass

Software zu Administrationszwecken häufig nicht auf den Standard-HTTP- oder HTTPS-Ports 80 beziehungsweise 443 läuft, schafft keine Abhilfe: Der Port ist ebenfalls Teil der URL und kann in die Suche einbezogen werden. Als Beispiel mag hier der populäre VNC-Server dienen, der den Java-Client-Zugriff über den TCP-Port 5800 abwickelt (vgl. Abb. 4) – mittels der folgenden Anfrage finden:

```
intitle:VNC inurl:5800
```

VNC-Desktop Login-Seite]">
Abbildung 4: Da Suchmaschinen Port-Nummern als Teil der URL ansehen, lassen sich auch Konfigurationsserver auf Nicht-HTTP(S)-Ports finden – im Bild: Java-Steuerung des VNC-Servers.

Auch hier liefern die Seiten der Systeme zusätzliche Informationen: Zum einen kann man über den Titel "VNC Desktop", "Ultr@VNC Desktop" oder "VNC Viewer for Java" die verwendete Version unterscheiden. Zum anderen verlangen die meisten VNC-Server in der Praxis nur eine Passwort-Eingabe, aber keinen Benutzernamen – es handelt sich dabei in der Regel um die "Free Edition" von RealVNC, die nicht mehrbenutzerfähig ist. Die Sicherheit eines solchen Systems hängt also ausschließlich von der Komplexität eines einzelnen Passworts ab! Der Einsatz von VNC zur Fernwartung mag zwar einfach sein, die Suche danach aber ebenso...

Ebenso wie sich über intitle: VNC-Server unterscheiden lassen, so ist dies auch bei Geräten möglich: Im einfachsten Fall steht im Titel unmittelbar die Produktversion. Dies ist beispielsweise bei Webcams der Firma Axis der Fall. Die Suche

```
intitle:"Live View / - AXIS"
```

führt zu einer ganzen Reihe von Axis Network Cameras (vgl.

Abb. 5).

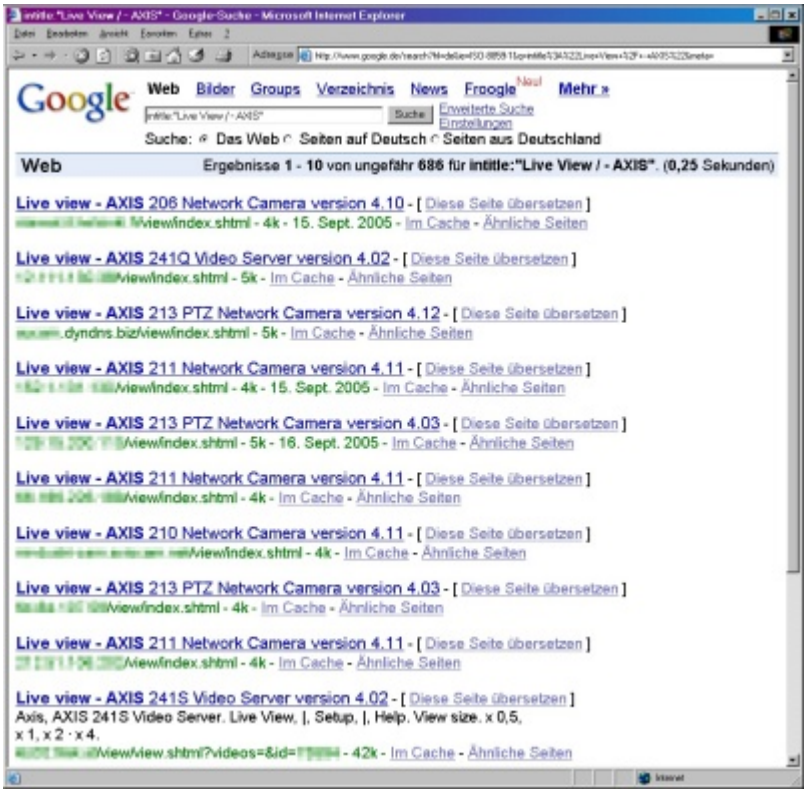


Abbildung 5: Viele Systeme verraten bereits im Titel eine Menge über ihren Softwarestand oder die vorhandene Produktversion – im Bild: Suche nach Axis Webcams.

Anhand der Titel kann ein Google-Hacker auch selektiv nach unterschiedlichen Modellen suchen:

- `intitle:"Live View / - AXIS 205"` bezeichnet eines der kleineren Modelle,
- `intitle:"Live View / - AXIS 241Q"` liefert hingegen Video-Server, die bis zu vier Kameras bedienen können.

Zudem können natürlich auch – bei mangelnder Trennschärfe der Titel oder URLs – beliebige Inhalte der Webseiten für eine Unterscheidung von Textbeiträgen oder zur Differenzierung verschiedener Versionen oder Produktserien dienlich sein. So erhält ein Angreifer schon im ersten Schritt gegebenenfalls auch Hinweise auf mögliche Verwundbarkeiten beziehungsweise anzuwendende Exploits für konkrete Versionen.

Gegenmaßnahmen und Ausblick

Als Schutzmaßnahme sollte eigentlich bereits die korrekte Konfiguration von Firewalls genügen: Denn wenn "alles, was nicht erlaubt ist, ist verboten" (default deny) gilt, dann müsste das an sich auch bedeuten "alles, was nicht explizit im Internet veröffentlicht werden soll, darf nicht erreichbar sein". Auffallend viele Einträge von Google-Hacks in der Google-Hacking-Database [1] führen dennoch zu Anwendungen und Geräten, die ein einfacher Paketfilter vor Entdeckung schützen könnte. Man steht also vor einem alten Problem: Es wird in vielen Fällen entweder nicht ausreichend gefiltert oder aus Bequemlichkeit stehen Anwendungen zur Fernwartung oder -nutzung bereit, die womöglich besser "im Verborgenen" blieben. Diese ließen sich zwar auch manuell finden; durch die weitreichende Indexierung des Webs durch Suchmaschine erreicht dies aber eine neue Dimension – noch nie war das Auffinden von Zielen so einfach wie heute.

Weitere "Anwendungsmöglichkeiten" bietet Google-Hacking übrigens auch im Bereich Vertrieb und Lizenzierung: Der Hersteller eines Produkts mit Webschnittstelle könnte seine Vertriebsmannschaft nach eigenen Produkten suchen lassen, um zum Beispiel veraltete Versionen zu ermitteln. Anschließend kann man dann telefonisch den Kauf eines Upgrades empfehlen. In ähnlicher Weise könnte unter Umständen auch festgestellt werden, ob der Betreiber überhaupt eine Lizenz für das Produkt hat.

Im Bereich von Sicherheitstests, die typischerweise nur eine begrenzte Zahl von IP-Adressen umfassen, erscheint der Einsatz von suchmaschinengestütztem Hacking übrigens nicht sinnvoll: Jede Form von Webserver kann im Rahmen eines "lokalen" Penetration-Tests auch direkt gefunden, identifiziert und analysiert werden. Interessanter sind hier Vorführungen zur Sensibilisierung, wobei auch angepasste Such-Strings auf der Basis innerhalb der Organisation verfügbarer Anwendungen Verwendung finden können. Sehr schnell wird dann jedem Zuschauer klar, welch fatale Folgen der unsachgemäße Umgang mit Systemen und Informationen haben kann.

Was sich mit Suchmaschinen finden lässt, ist allerdings trotz allem nur die Spitze des Eisbergs: Suchmaschinen führen – glücklicherweise – keine Portscans nach Webservern und anderen offenen Ports durch und können nur diejenigen Webseiten finden, auf die Hyperlinks verweisen. Tatsächlich ist die Menge der vertraulichen Informationen, die ungeschützt im Internet zu finden sind, weitaus größer. Und Penetrationstests beweisen immer wieder, dass es allzu oft mit einfachsten Mitteln möglich ist, auf solche Informationen

zuzugreifen.

Sebastian Schreiber ist Gründer und Geschäftsführer der SySS GmbH, Stefan Arbeiter ist Penetration Tester bei SySS (www.syss.de).

Literatur

[1]
Google Hacking Database (GHDB),
http://johnny.ihackstuff.com/index.php?module=prodreviews

[2]
Johnny Long, Ed Skoudis, Alrik van Eijkelenborg, Google Hacking for Penetration Testers, Syngress Media, ca. 34 €, ISBN 1-931836-36-1 *oder auf Deutsch:* Google Hacking, Mitp-Verlag, 32 €, ISBN 3-8266-1578-6