NetSecure-IT

Next Generation Security

Whitepaper

NetSecure-IT Next Generation Security

Google Hacking

Whitepaper / Wild Wild Web

Unternehmensbezogen

Table of Contents:

Was ist Google-Hacking?	2
Möglichkeiten von Google-Hacking	2
Befehlssyntax Google, Yahoo und Bing	3,4
Kombinationen von Suchbegriffen	5,6
Ansprechbare Geräte (Peripherie) finden	7,8
Möglichkeiten der Ausnutzung	8
Möglichkeiten des Angriffs	8
Automatisierung	8
Möglichkeiten der Abwehrmaßnahmen	9
Fazit	9
Links	10

Google Hacking ist nicht etwa eine Methode, um die Suchmaschine selbst zu hacken. Es handelt sich hier um spezielle Suchbegriffe. Mittels dieser Parameter ist es möglich, die Verwundbarkeit von abgelegten Dateien und Verzeichnissen auf Webservern zu erforschen, Zugangsdaten aufzuspüren, Informationen zu sammeln, die Angriffe auf ein Netzwerk erleichtern können. Zudem können Informationen, die nicht für die Öffentlichkeit bestimmt sind abgerufen werden.

Google Hacking reiht sich in die Kategorien – wenn wir hier über Angreifer Motive sprechen – der Neugierde, Langeweile, Materielle Interessen, Spionage und Geltungssucht ein. Wobei die Neugierde unter Script-Kiddies wohl an einer der ersten Stellen zu zählen sein dürfte. Eine gute Nachricht gibt es, man hat es bei dieser Art von Angriff zu 98 % nicht direkt auf Sie abgesehen. Das war es aber dann auch schon mit den guten Nachrichten.

Die Möglichkeiten der Informationen, die dabei abgerufen werden können, sind groß. Unter anderem können:

- Zugangsdaten
- Passwort Dateien
- Kundeninformationen
- Persönliche Daten
- Vertrauliche Informationen
- Informationen über Webapplikationen / Fehlermeldungen
- Verzeichnisinhalte
- Administrationsschnittstellen (z.B. von Multifunktionsgeräten (Drucker, Faxgeräte, Scanner usw.)
- Intranet-Seiten, die nur für den internen Gebrauch konzipiert wurden und nie an die Öffentlichkeit gelangen sollten.
- Authentifikationen

Allerdings muss man die Syntax der Suchanfragen formulieren, wobei auch mehrere aufeinanderfolgende Anfragen akzeptiert werden, um die Suchanfrage in einem Möglichst genauen Kreis zu erhalten.

Nachfolgende Syntax, die Google und auch Yahoo sowie Bing bei der Detailsuche unterstützt.

Zu finden sind diese Befehle übrigens auf tausenden Internetseiten. Es ist also kein Geheimnis diese Befehle für sich zu verwenden. Allerdings ist es in der Bundesrepublik momentan noch so, dass die bloße Begutachtung, oder das Auffinden der Daten keine Straftat darstellt. Straftaten nach dem Hackerparagraph § 202c des deutschen Strafgesetzbuches (StGB) stellen sich erst dann ein, wenn mit den gefundenen Daten oder Zugängen Schindluder betrieben wird. In England ist es bereits eine Strafe, wenn bereits nach solchen Underground Informationen gesucht wird.

Folgende Parameter stellen die Hauptsyntax dar:

- Syntax: „**site:**" / Einschränkung der Suche auf Seiten, die sich unterhalb der als Argument eingesetzten Domäne befinden.
- Syntax: „**inurl:**" / Das Argument muss sich irgendwo in der URL befinden.
- Syntax: „**allinurl:**" / Alle hierauf folgenden Worte müssen sich in der URL befinden.
- Syntax: „**allintext:**" / Alle folgenden Worte müssen sich im Text befinden.
- Syntax: „**intext:**" / Das Argument muss sich im Text, also nicht etwa im Titel oder der URL befinden.
- Syntax „**intitle:**" / Das Argument muss sich im Titel der Seite befinden (innerhalb des Headers der HTML-Datei)
- Syntax „**allintitle:**" / Alle folgenden Worte müssen sich im Titel der Seite befinden.
- Syntax „**link:**" / Auf der gesuchten Seite muss sich ein Link auf das hier eingesetzte Argument befinden.
- Syntax „**related:**" / Findet Seiten mit ähnlichem Themenkontext, wie die Seite, deren URL als Parameter angegeben wird.
- Syntax „**filetyp:**" / Das gesuchte Dokument hat diese Dateiendung. Die Syntax „**ext:**" besitzt die gleiche Funktion.
- Syntax „**cache:**" / Als Argument wird eine URL eingesetzt. Das Suchergebnis liefert bei Vorhandensein die in den Google-Cache aufgenommene Version der Seite.
- Syntax „**define:**" / Sucht in Online-Enzyklopädien den Begriff, der als Argument angegeben wurde.
- Syntax „**numrange:**" / Es werden Zahlen von bis gesucht.

An sich stellen diese Parameter noch keinen Grund zur Besorgnis dar, da sie von jedermann für explizite Suchergebnisse verwendet werden können. Allerdings stellen die nachfolgenden Anfragen schon etwas mehr Ansprüche an Google. Erst mit Einsatz von zusätzlichen Code und Parametern können Daten und Zugänge gefunden werden, die nicht richtig oder falsch abgesichert wurden. Ein Auszug davon einmal auf der nächsten Seite.

Ganz oben auf der Liste steht noch immer der Syntax: „**index of/**" und heißt nichts anderes als dass dies Webserver Verzeichnisse sind, die per http auf Port 80 zugänglich sind und meist keine index-Datei besitzen. Man sieht quasi hinter die Kulisse eines Unternehmens oder einer privaten Site. Sind nun unbeabsichtigter Weise unzureichende Konfigurationen vorgenommen worden, sind die Verzeichnisse sichtbar und man kann darauf zugreifen.

Ein Beispiel nach Verzeichnissen ohne Index-Datei:

„**Intitle: „Index of:**". Interessanter wird es allerdings wenn weitere Parameter folgen, wie zum Beispiel „privat", „backup" oder „vertraulich". Allerdings muss man hierfür ein wenig Geduld aufbringen und nicht jede Syntax bringt heute noch brauchbares zu Tage, wenn es um Sammeln von nicht für die Öffentlichkeit bestimmte Daten oder Zugänge geht.

Tippen Sie mal folgende Syntax ein und recherchieren Sie einmal ein wenig: ***intitle:"index of" vertraege docx*** *(statt docx können Sie auch odt verwenden) / Die Endung odt brachte mich z.B. zur WikiLeaks Torrent Files Collection.*

Sie können das natürlich mit beliebigen Wörtern nach index of weiterverfolgen. Denken Sie jedoch daran. **Angucken erlaubt, Missbrauchen verboten.** Angucken erlaubt = in der Bundesrepublik noch nicht Strafbar. Es kann natürlich auch vorkommen, dass man in einen Honigtopf[1] tappt.

Weitere Syntax nachfolgend zum Ausprobieren. *(Der Hintergrund dessen ist jedoch nicht, dass dazu animiert werden soll, wie man an fremde ungeschützte Daten herankommt, sondern wie man sich davor schützen kann). Die weiterführenden Parameter dienen deshalb dem Verständnis zur Absicherung der eignen Konfiguration. Zudem gibt es etliche Seiten im Internet, die die Parameter führen. Im Anhang dieses Whitepapers sind einige Seiten aufgeführt.*

[1] Ein Honigtopf ist ein Dienst in Computer-Netzwerken, der die Aufgabe hat, falsche Fährten für Hacker zu legen um die Vorgänge besser verstehen zu können.

Einige mögliche Google-Syntax -

Benutzung auf eigene Gefahr

intilte:index.of.etc

intitle:"index of" .sh_history

intitle:"index of" .bash_history

intitle:"index of" passwd

intitle:"index of" pwd.db

intitle:"index of" etc

intitle:"index of" etc/shadow

intitle:"index of" backup

intitle:"index of" spwd

intitle:"index of" master.passwd

intitle:"index of" htpasswd

intitle:"index of"password modified

intitle:"index.of.secure"

intitle:"index.of cgi-bin"

intitle:"index of" iissamples

allintitle:sensitive filetype:docx

allintitle:sensitive filetype:odt

allintitle:*.php?filename=*

Allintitle:*.php?page=*

intitle:tech-support inurl:show cisco

filetype:ini+ws_ftp+pwd

intitle:index.of ws_ftp.ini

allinurl:cgi-bin password

inurl:access.log filetype:log -cvs

filetype:log inurl:cache.log

filetype:log inurl:access.log TCP_HIT

filetype:log inurl:usergent.log

ext:log "software:microsoft internet information service *.*"

"index of" /"chat/logs"

filetype:log username putty

intitle:index.of.sh_history

filetype:inc intext:mysql_connect

inurl:admin.pwd filetype:pwd

adminpassword sysprep filetype:inf

intext:enc_User_Password=* ext:pcf ext:ini eudora.ini

"parent direktory"+proftpdpassword

filetype:inf sysprep

inurl:admin user password

intile:index.of.etc passwd

intitle:index.of passwd

intitle:index.of.htpasswd

intitle:index.of ".htpasswd" htpasswd.bak

filetype:xls username password email

intitle:index.of secring.pgp

intilte:index.of..etc hosts

intitle:index.of administrators.pwd

intitle:index.of passlist

Die Liste selbst bleibt unkommentiert!

Die verwendete obige Syntax führt oft auch zu Hackerportalen. Es wird daher empfohlen, das Sie eine virtuelle Umgebung, Testrechner (außerhalb, am besten in einer DMZ) oder einen „Browser in the Box" für Ihre Such-Tests verwenden, um mögliche Schäden gering zu halten bzw. abzuwenden.

*Die Liste stellt nur einen kleinen Teil möglicher Parameter dar und dient in erster Linie dem Verständnis für das Thema „**Google-Hacking**".*

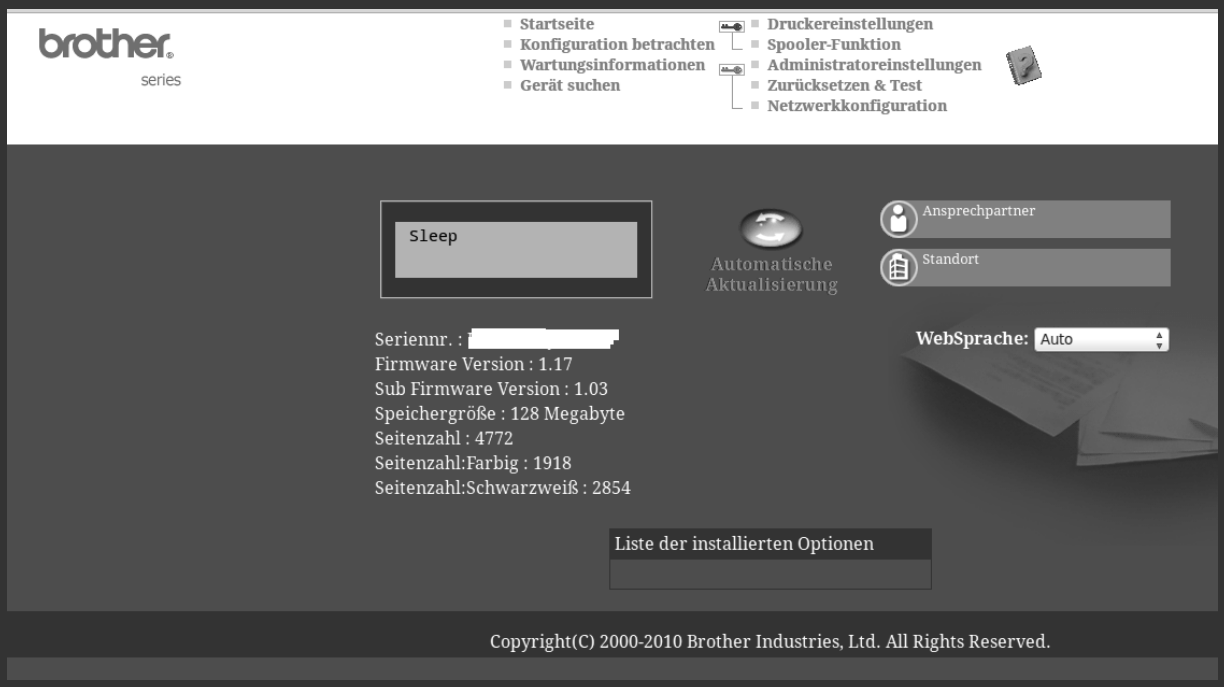
In jedem Fall ist es wichtig, dass Sie verstehen, worum es hier geht. Der wichtigste Aspekt ergibt sich aus der Maßnahme, dass mit den gefundenen Daten teils schwerwiegende Angriffe gefahren werden könnten. Umso wichtiger ist es zu wissen, wie Applikationen richtig gesichert werden können.

Ein einfaches Beispiel:

Viele benutzen für Ihren Upload ein Ftp-Client. Wir nehmen hier das Beispiel von WS_FTP. Immer dann, wenn Sie auf Ihren FTP-Server zugreifen, werden dort Login Daten abgelegt. Ist dieses Verzeichnis nicht geschützt, so kann darauf zugegriffen werden. Oft sind die Passwörter verschlüsselt. Um den Klartext zu erhalten müssen Sie nur noch einen Passwortcracker verwenden. **Sie machen sich jedoch strafbar, wenn Sie die gewonnenen Informationen gegen Zustimmung und Wissen des Betreibers verwenden.** Auf FTP Server sollte übrigens nur verschlüsselt (SFTP) zugegriffen werden.

Eine andere oft unterschätze Art von Sicherheitslücke kann eine Peripherie darstellen. Drucker, Multifunktionsgeräte, Webcam usw. Viele Druckermodelle bieten eine Konfiguration über http Port 80 an. Dumm ist es dann, wenn die Standard Benutzer und Passwörter nicht geändert wurden. Diese finden Sie meist bei den Herstellern auf deren Internetseite.

Ein weiteres Beispiel:



Auf Seite 7 sehen Sie ein Interface eines Druckers, welcher über http angesprochen werden kann.

Unter Gerät suchen fand ich noch eine ganze Reihe weiterer Geräte, darunter auch Faxe, die es zuließen, sämtliche Listen von eingegangenen sowie ausgegangenen Listen einzusehen. Zu bedauern waren die Administrator- und Druckereinstellungen, die mit den Default Werten des Herstellers belegt waren. Insbesondere bei Brother wird man schnell fündig, wenn man nach *„Brother „Geräte Name" Passwort"* sucht. Was mit Zugang der Administrator Einstellungen angestellt werden kann, können Sie sich nun selbst ausmalen.

Sie fragen sich sicherlich nun, wie das alles in die Suchmaschine kommen kann. Nun, Google benutzt einen Webcrawler auch Spider, Robot oder Searchbot genannt. Diese durchkämmen das Netz in regelmäßigen Abständen. Wie beim Internetsurfen gelangt der Crawler über Hyperlinks von einer Seite zur anderen. Dabei werden die Funde in der Datenbank der Suchmaschine gespeichert. Über die „robots.txt" und in bestimmten Meta-Tags im HTML Header *(teils schon überflüssig)* kann man dem Spider mitteilen, welche Seiten er indizieren soll und somit in seine Datenbank aufzunehmen oder welche er nicht aufnehmen soll. Fehlen diese Einträge oder sind die Angaben falsch konfiguriert, stehen den Neugierigen und in böser Absicht gestellte Zeitgenossen Tür und Tor auf, Ihre Daten abzugreifen.

Ein weiteres Feature in Bezug auf Google-Hacking stellt den **Google-Cache** dar. Dieser Cache enthält eine Art Schnappschuss von Seiten und Inhalten, die sich im Google-Index befinden. Haben Sie ausversehen vertrauliche Inhalte hinterlegt, diese jedoch bereits entfernt, kann immer noch ein Snapshot im Google-Cache vorhanden sein. Wie lange allerdings ein Snapshot im Cache bleibt kann nicht ganz genau beantwortet werden. Je öfter eine Seite besucht wird, umso schneller aktualisiert auch der Cache. Sonst bleibt nur abzuwarten bis der Spider die Seite erneut gecrawlt hat, Dauer ca. max. 60 Stunden bei Google und Yahoo.

Durch **Vulnerability-Scanner** ist es möglich die ganze Suche noch zu automatisieren. Im Allgemeinen wird durch einen „Vulnerability Scan" ein IT-System größtenteils automatisiert und auf Anfälligkeit gegenüber bestimmten Angriffsmustern überprüft. In diesem Zusammenhang kann davon ausgegangen werden, dass Google-Hacking unter diese Rubrik des Scannens fällt. Automatisierte Scanner sind beispielweise GooScann, Sitedigger oder Wikto.

Abwehrmöglichkeiten und betrieblicher Mehrwert durch Google-Hacking ist davon abhängig, welche Maßnahmen bereits in einem Betrieb ergriffen worden sind und / oder die Sicherheitszustände bereits erfasst worden sind.

Als präventive Abwehrmaßnahme gilt der Einsatz von einer ordentlich konfigurierten „***robots.txt***". Weiter sind professionelle Tools mit entsprechenden Security Mechanismen für Web-Publishing unabdingbar. Ein Provider mit Sitz in de und Zertifizierung ist von Vorteil. Als weitere Barriere empfiehlt sich z.B. Schutz vor SPAM Bots die „***htaccess***" Datei, da diese die „robots.txt" übergehen.

Haben Sie dagegen einen Webserver in Eigenregie in Ihrem Betrieb, so ist es ratsam durch regelmäßige Tests sicher zu stellen, dass keine Schlupflöcher vorhanden sind. Datenbanken, Tabellenkalkulationen, E-Mail Listen und Telefonbücher gehören nicht ungesichert auf einen Web Sever. Wenn Sie pdf Dateien auf Ihrem Webserver lagern haben oder Word bzw. Libre Office / Open Office Dokumente, tragen Sie in die „robots.txt" den Zusatz „Disallow:/*.pdf$, Disallow:/*.docx$ oder Disallow:/*.odt$" usw. ein.

Testen Sie Ihren Webspace auf Schwachstellen bevor es andere tun!

Zusammenfassung:

Google Hacking kann also dazu benutz werden, Angriffsziele auszumachen und /oder auch zu benutzen. Was in der Vergangenheit anfängliche Neugierde war, kann nun unter Umständen zu schwerwiegenden Konsequenzen führen.

Sicherheitstest sollten deshalb unbedingt „Google-Hacking" als ergänzender Vulnerability-Scan mit aufgenommen werden um mögliche Verwundbarkeiten aufzudecken. Der Mehrwert für ein Unternehmen liegt auf der Hand, da durch diese Methode spezielle Funde gemacht werden können, an die sonst wohl kaum einer denkt.

Google-Hacking stellt somit eine sinnvolle Ergänzung zu bestehenden Sicherheitstests dar. Die Kosten für ein solches Konzept sind relativ niedrig. Die angesprochenen Scanner sind meist kostenlos und die Zeit für einen solchen Scan ist relativ gering gegenüber anderen Sicherheitstests.

Anhang /

Sites mit Parametern für Google-Hacking & Exploit Database

http://www.thehackerslibrary.com/?p=849

http://www.exploit-db.com/google-dorks/

http://hackerlib.blogspot.de/2009/12/chapter-6.html

http://johnny.ihackstuff.com/index.php?module=prodreviews

http://it.toolbox.com/blogs/managing-infosec/google-hacking-master-list-28302

Weitere finden Sie, wenn Sie recherchieren.

Ein Whitepaper von http://netSecure-IT.de
E-Mail: info@netSecure-IT.de