Google Hacking

Ali Jahangiri

Sc.D, LPT, CEH, CHFI, CEI, ECSA, ISMS Lead Auditor, Security+,
CIW Security Analyst, MCSE: Security, MBCS CITP, MCSA, MCDBA, CCNA, A+

Abstract: Google hacking is the term used when a hacker tries to find vulnerable targets or sensitive data by using the Google search engine. In Google hacking hackers use search engine commands or complex search queries to locate sensitive data and vulnerable devices on the Internet.

Keywords: hacking, hack, Google, Google hack, hacking techniques, attack, ethical hacking, search engines, search engine hacking

What is Google Hacking?

Google hacking is the term used when a hacker tries to find vulnerable targets or sensitive data by using the Google search engine. In Google hacking hackers use search engine commands or complex search queries to locate sensitive data and vulnerable devices on the Internet.

Although Google hacking techniques are against Google terms of service[1] and Google blocks well-known Google hacking queries, nothing can stop hackers from crawling websites and launching Google queries.

Google hacking can be used to locate vulnerable web servers and websites which are listed in the Google search engine database. In other words, hackers can locate many thousands of vulnerable websites, web servers and online devices all around the world and select their targets randomly. This kind of attack is most commonly launched by applying Google hacking techniques to satisfy junior hackers.

It is obvious that the Google hacking procedure is based on certain keywords, which could be used effectively if they are used by some internal commands of the Google search engine. These commands can be used to help hackers narrow down their search to locate sensitive data or vulnerable devices.

Nevertheless, the success of Google hacking techniques depends on the existence of vulnerable sites, servers and devices. However, we should not ignore the power of the search engines in providing information about the targets to the hackers in the reconnaissance phase.

Beyond Vulnerability

Malicious hackers can use Google hacking techniques to identify vulnerable sites and web servers for known vulnerabilities. In addition, they can look for error pages with the help of technical

information or retrieve files and directories with sensitive contents such as databases, passwords, log files, login pages or online devices such as IP cameras and network storage.

Google Proxy

Hackers can use the *Google Translate* service (http://translate.google.com/translate_t) as a proxy server to visit a website or translate the contents of the website or URLs without leaving any footprints.

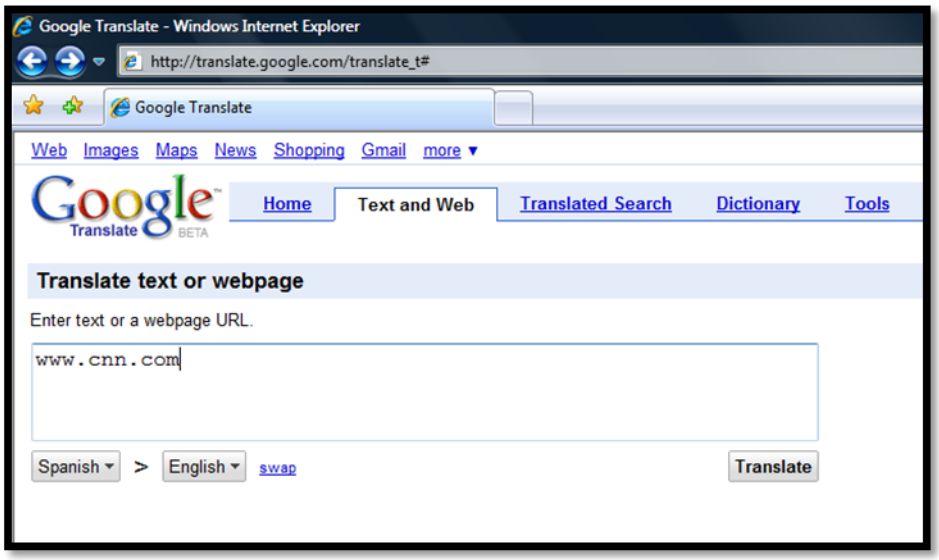

Figure 1: Google Translate Service.

Google Cash

Google copies the content of a website in its database. This function helps users to access the content of the website if the site is not available. However, a hacker can use this function to access and visit a targeted website without leaving any footprint and in complete anonymity.

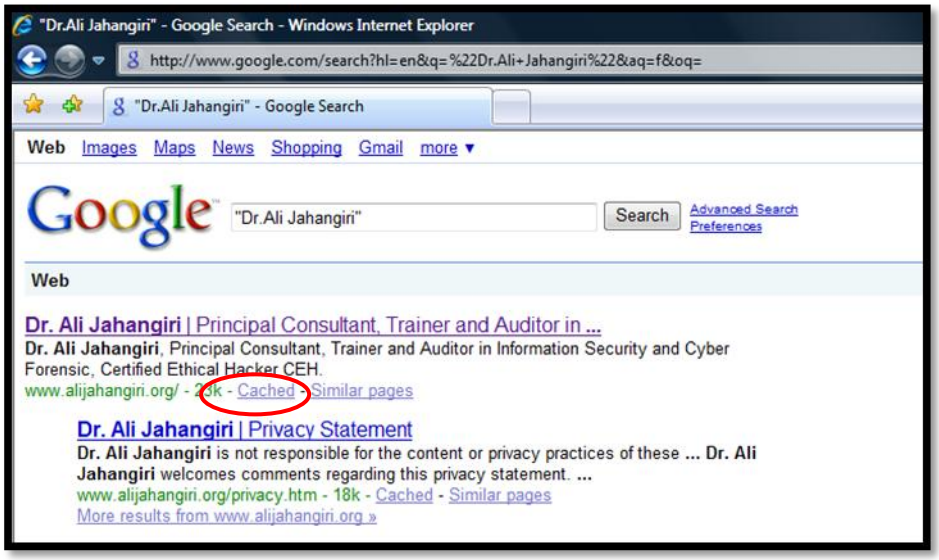

Figure 2: The red cycle indicates the link to access the Cached page.

Directory Listings

Web server applications such as Apache and IIS provide facilities that a user can browse and navigate website directories by clicking on the directory name and links such as *Parent Directories*. The directories and their content can be listed if directory listing or directory browsing are enabled by the administrator. This vulnerability gives an unauthorized access to the files and it may help hackers to gain access to the information which can help them to hack a website or a web server or download its contents.

Directory listings make the parent directory links available to browse directories and files. Hackers can locate the sensitive information and files just by simple browsing. In Google it is easy to find websites or web servers with enabled directory listings because the title of the pages start with the "index of" phrase so we can use *index of* in the search box to find the directory listings-enabled website. If we want to get better result from our search we can use this combination in the search box *intitle:index.of* or we can use *intitle:index.of "Parent Directory"*.

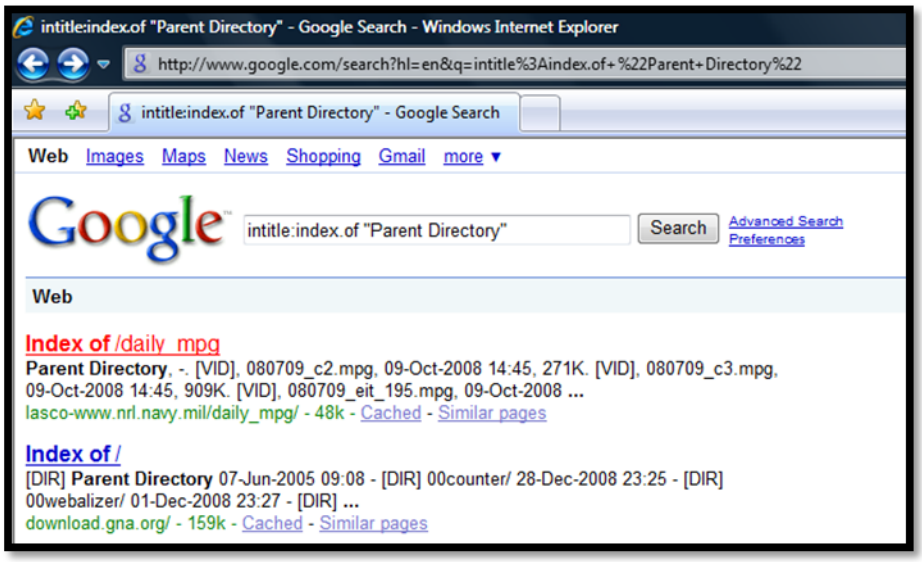

Figure 3: The result of using *intitle:index.of "Parent Directory"*.

It is obvious that with the first command we used the Google search engine to search in its database for the websites which have been listed with the title of "Index of". In the second command we used Google to search for sites with the directory listings and with the keyword which is often found in the directory listings.

Specific Directory

Hackers can locate specific directories by using the directory name in their search queries. For instance to locate an "admin" directory in addition to directory listings, the hacker can use these commands: *intitle:index.of.admin* or *intitle:index.of inurl:admin*.

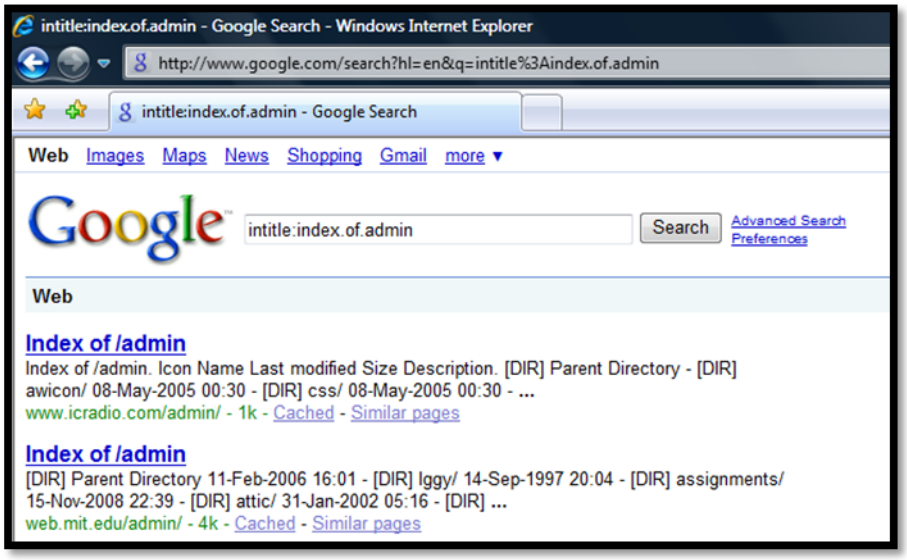

Figure 4: The result of using *intitle:index.of.admin*.

Specific File

It is possible to search for a certain file by directory listings. For instance, to search for the password.mdb file, this search query can be used: *intitle:index.of password.mdb* .

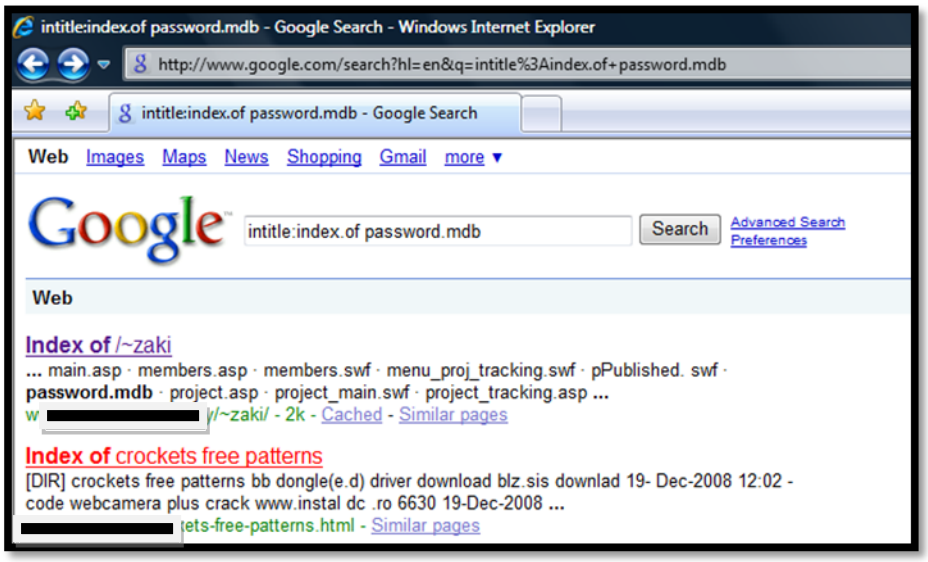

Figure 5: The result of using *intitle:index.of.password.mdb.*

Specific File Extension

Google lets users search its database for a specific file extension by using the *filetype:* command. For instance, if you want to search for pdf files, then you can use the query *filetype:pdf* in the search box.

Server Information

It is possible to use Google hacking techniques to determine the version of the web server application along with directory listings. This kind of information is vital to an attacker because it will

help him or her to use the best way to attack the web server. For instance, hackers can use the search query *intitle:index.of "server at"* to find the web sites with vulnerable directory listings which are operated by an Apache server.

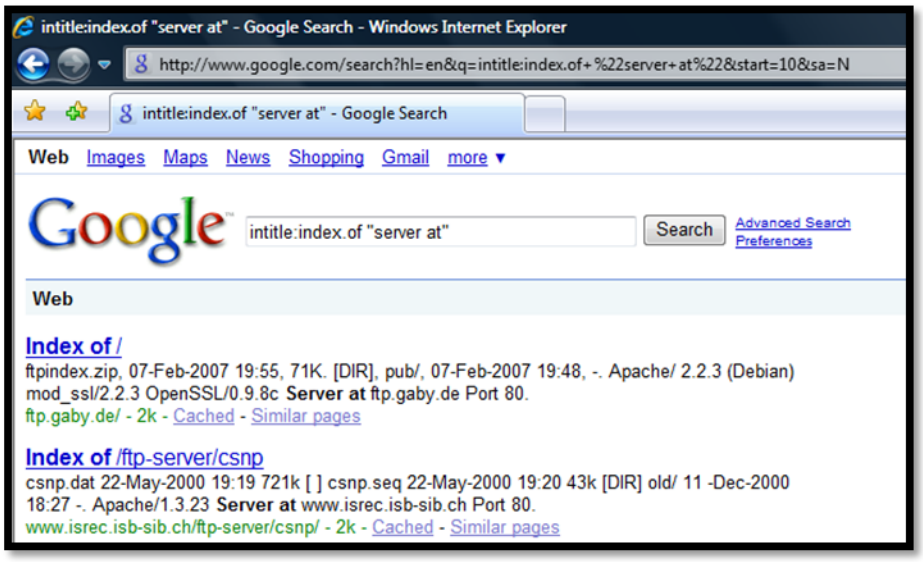

Figure 6: The result of *intitle:index.of "server at".*

Different versions of Microsoft IIS servers have wide usage all around the world. It would be easy to find the servers which are operated by Microsoft IIS 6.0 servers, which are listed in the Google database by using the query *"Microsoft IIS/6.0 server at"* on the Google search engine.

Error Pages

The error pages and warning pages are informative for hackers because these pages could be used to determine the vulnerability of the target. Most of the time hackers use the error messages as keywords or search phrase to find their targets. For instance, if you use *"Syntax error in query expression " –the* in the Google search box, you can find the websites which have this error message as an Access error message; this message can display path names, function names and filenames which are helpful for the hackers.

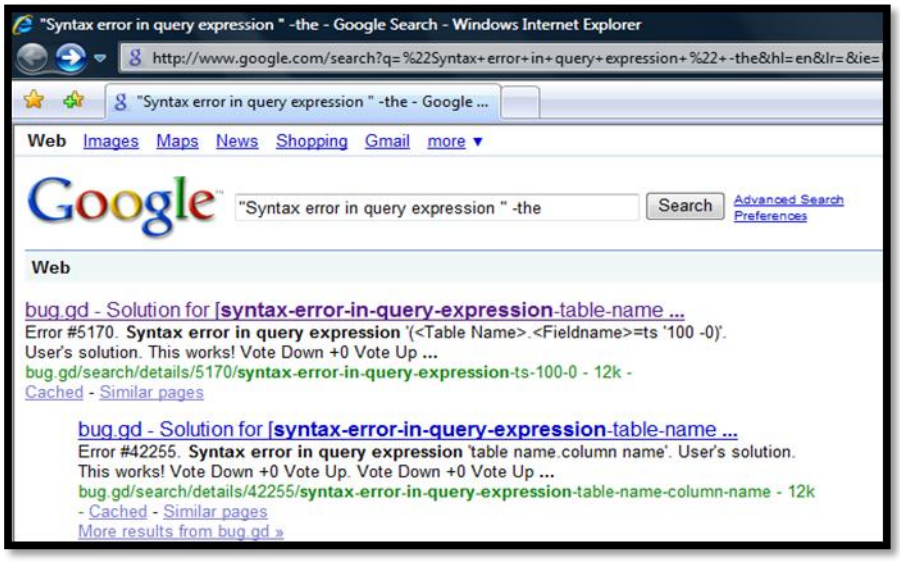

Figure 7: The result of "Syntax error in query expression " –the.

Hackers may use Google to locate vulnerable servers by searching for the error pages of web servers such as IIS. The queries *intitle:"the page cannot be found"*and *"internet information services"* can be used to search for IIS servers that present error 404.

Default Pages

Default pages are major sources of information about targets for hackers. They use Google to find live servers which are on the default page; most of the time, these servers have default configurations with many vulnerabilities.

Login Pages

The login pages can be use for brute force attacks and gain unauthorized access to the target. In addition, the login pages can be useful to provide information about the target server. For instance, if we use the search query *allinurl:"exchange/logon.asp"* in the Google search box, we can find the login page of the Microsoft Outlook Web Access.

For the typical login page in the web applications or portals which have been programmed by ASP, you can use inurl:login.asp or inurl:/admin/login.asp.

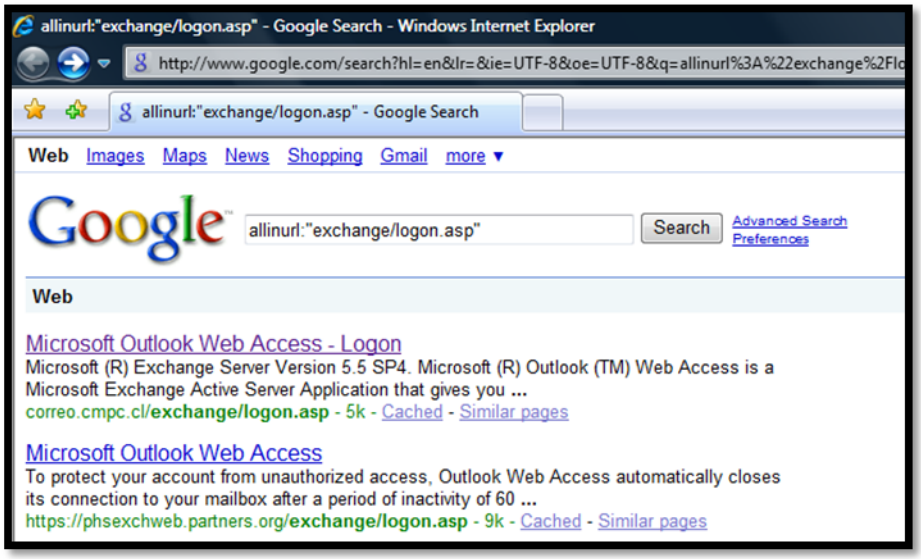

Figure 8: The result of *allinurl:"exchange/logon.asp".*

Locating CGI-BIN

Common Gateway Interface (CGI) is a standard protocol for interfacing external application software with web servers. Hackers can use Google to locate the CGI-BIN applications or pages to target. For instance, the search query *inurl:/cgi-bin/login.cgi* locates the login pages base on CGI-BIN.

Online Devices

It is possible to create special search phrases to locate online devices such as IP cameras, network storage and printers with Google. In this technique hackers use the default pages or the application names which vendors used for hardware and that have been supplied by vendors.

For instance, if you want to locate AXIS Network cameras then you can apply the search phrase *inurl:indexFrame.shtml Axis* to find online AXIS cameras. Here is another example: to locate online Linksys network storage with the GigaDrive Utility, you can use the search phrase *intitle:"GigaDrive Utility"* in the Google Search box.

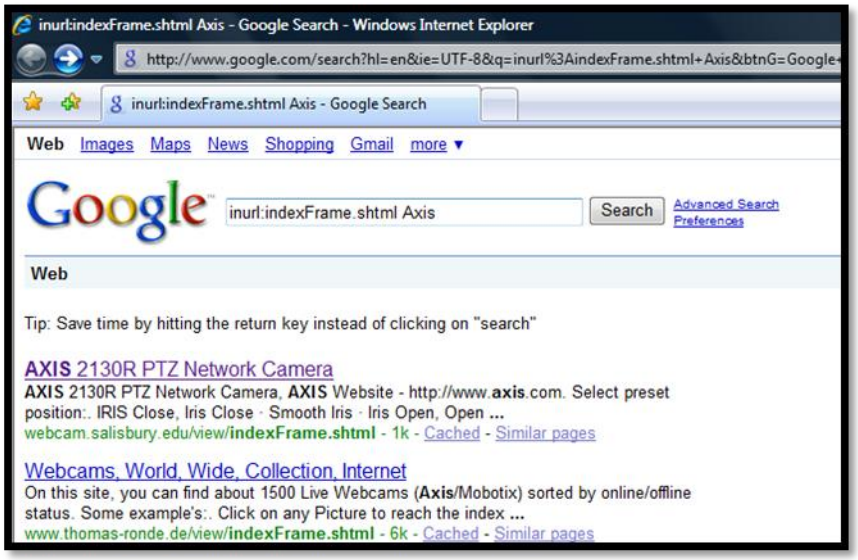

Figure 9: The result of *inurl:indexFrame.shtml Axis*.

Google Hacking Database

There is an unofficial website (http://johnny.ihackstuff.com/ghdb.php) which acts as a database for hacking of Google. This database has been used since its creation in 2004 by the Google hacking community.

You would be able to develop your own Google hacking database by studying the behaviour of the equipment and identifying the pages, page titles and files which can be called and accessed by user and which will be listed in Google.

Disclaimer:

- This document is to educate, introduce and demonstrate Google hacking. You should not use the information which has been presented in this document for illegal or malicious attacks and you should not use the described techniques in an attempt to compromise any computer system.
- Ali Jahangiri operates a policy of continuous development. The information which this document contains reflects his understanding at the time when presented. Ali Jahangiri reserves the right to revise this document or withdraw it at any time without prior notice and states no obligation to update the data included in this document.
- The contents of this document are provided "as is". No warranties of any kind, either express or implied, including, but not limited to, the implied warranties of solutions and instructions for a particular purpose, are made in relation to the accuracy, reliability or contents of this document.
- Under no circumstances shall Ali Jahangiri be responsible for any loss of data or income or any special, incidental, consequential or indirect damages howsoever caused.

[1] http://www.google.com/accounts/TOS