Who you calling a Dork? – Using Google to find vulnerable and exploited web servers

Aaron Goldstein

November 20, 2014
About Me

• Currently handle Cyber Security and Threat Intel for Amgen (Biotechnology)
• Previous Experience-
  – Incident Response and Forensics Consulting
  – Responded to over 100+ incidents for large and small companies, including Fortune 500, Universities, Medical, Financial, Gov’t, etc.
• Over 8 years experience in the DFIR field
• (Ethical) Hacker by heart
What we will cover today

- What is Google Dorking
- Legitimate / Nefarious uses for using Google
- How can this info be used by Pen Testers / Vuln assessments
- Manual examples
- Automated tools
- How to protect your own systems
What is Google Dorking?

• Straight from the source:

Google Dorking

Google Dorking is a term that refers to the practice of applying advanced search techniques and specialized search engine parameters to discover confidential information from companies and individuals that wouldn't typically show up during a normal web search.
The process of using google indexing service to find (potentially sensitive) information

– Can be completely legitimate and useful.
– Can also be used for evil

*source: Infosec Institute
• Quick and easy searching across multiple domains

• Limit your searches to only items of importance

• Combine multiple searches into one query
**Dorking Database**

Welcome to the google hacking database

We call them 'googledorks': Inept or foolish people as revealed by Google. Whatever you call these fools, you've found the center of the Google Hacking Universe!

---

**Search Google Dorks**

<table>
<thead>
<tr>
<th>Category:</th>
<th>Free text search:</th>
</tr>
</thead>
<tbody>
<tr>
<td>All</td>
<td></td>
</tr>
</tbody>
</table>

---

**Latest Google Hacking Entries**

<table>
<thead>
<tr>
<th>Date</th>
<th>Title</th>
<th>Category</th>
</tr>
</thead>
<tbody>
<tr>
<td>2014-11-18</td>
<td>ext:txt inurl:gov intext:&quot;Content-Type: text/...&quot;</td>
<td>Files containing juicy info</td>
</tr>
<tr>
<td>2014-11-17</td>
<td>ext:msg OR ext:mail OR ext:site:gov OR site:edu</td>
<td>Files containing juicy info</td>
</tr>
<tr>
<td>2014-11-03</td>
<td>inurl:CHANGELOG.txt intext:drupal intext:&quot;SA-...&quot;</td>
<td>Vulnerable Servers</td>
</tr>
<tr>
<td>2014-11-03</td>
<td>inurl:robots.txt intext:CHANGELOG.txt intext:disal...</td>
<td>Vulnerable Servers</td>
</tr>
<tr>
<td>2014-10-21</td>
<td>filetype:log intext:org.apache.hadoop.hdfs</td>
<td>Files containing juicy info</td>
</tr>
<tr>
<td>2014-10-15</td>
<td>inurl:cgi-bin/mailgraph.cgi</td>
<td>Various Online Devices</td>
</tr>
<tr>
<td>2014-10-14</td>
<td>inurl:logon.html &quot;CSC08&quot;</td>
<td>Pages containing login portals</td>
</tr>
<tr>
<td>2014-10-09</td>
<td>(intext:mail AND intext:samAccountName) AND (filet...</td>
<td>Files containing juicy info</td>
</tr>
<tr>
<td>2014-10-09</td>
<td>intext:5baa61e4c9b3f3f0d68225b66cf8331b7ee8f488 AN...</td>
<td>Files containing juicy info</td>
</tr>
<tr>
<td>2014-10-09</td>
<td>intitle:FRITZBox inurl:login.lua</td>
<td>Pages containing login portals</td>
</tr>
</tbody>
</table>
Questionable uses for Dorking

Quickly and easily source pirated material

Finding torrents

Grabbing pdf books
• Easy to find misconfigured applications and servers for information gathering and password harvesting

---

"/activity-monitor - GitHub
https://github.com/.../sftp-config.json
The tab key will cycle through the settings when first created. // Visit http://wbond.net/sublime_packages/sftp/settings for help. // sftp, ftp or ftps. "type": "sftp",
"save_before_upload": true,
"upload_on_save": true,
"confirm_sync":true,
"confirm_overwrite_newer":false,
"sync_down_on_open": false,
"sync_same_age": true,
"host": "...".
"user": "apiuser",
"password": "...",
"port": "22",
"remote_path": "/var/www/api",
"file_permissions": "664",
"dir_permissions": "775".

"/sftp-config.json at master ... - GitHub
https://github.com/.../sftp-config.json
Apr 20, 2012 - Contribute to development by creating an account on GitHub.

"/sftp-config.json at master · GitHub
https://github.com/.../sftp-config.json
Mar 6, 2014 - The tab key will cycle through the settings when first created. // Visit http://wbond.net/sublime_packages/sftp/settings for help. // sftp, ftp or ftps.

---

"site:github.com inurl:sftp-config.json
About 5,960 results (0.31 seconds)

"/activity-monitor - GitHub
https://github.com/.../sftp-config.json
The tab key will cycle through the settings when first created. // Visit http://wbond.net/sublime_packages/sftp/settings for help. // sftp, ftp or ftps. "type": "sftp",
"save_before_upload": true,
"upload_on_save": true,
"confirm_sync":true,
"confirm_overwrite_newer":false,
"sync_down_on_open": false,
"sync_same_age": true,
"host": "...".
"user": "apiuser",
"password": "...",
"port": "22",
"remote_path": "/var/www/api",
"file_permissions": "664",
"dir_permissions": "775".

"/sftp-config.json at master ... - GitHub
https://github.com/.../sftp-config.json
Apr 20, 2012 - Contribute to development by creating an account on GitHub.
• Finding vulnerable servers (like weak SSL)
  – Like Heartbleed and Shellshock
Finders Keepers

• Find already exploited web servers
  – Why work hard to exploit a server, when you can hijack an existing one?
• Find People and their hotel reservations (creepy)
• As you can see, this is all very easy to do, but if you have many targets we need to work smarter not harder
• There are many tools to assist in this
V3N0M Automated Dorking – FOSS – GPL v2
- “Largest and most powerful d0rker online”
- 18k+d0rks searched over 13 Engines at once
• the dumpster – python script for dorking
  – Older, but allows proxying (important!)
• Dork Searcher – FOSS (Windows)
  – Requires Proxy!
Basic OPSEC

- *I am not responsible or liable for what you do with this information!*
- *Potentially accessing sensitive / protected information. Be careful! This might be (likely is) considered illegal.*
- Better hide your tracks!
- Best option - **TOR**
  - Setup SOCKS5 proxy and route traffic through 127.0.0.1:9050
- Alternative option – **Proxies**
  - If you can’t find a proxy, how about the help of google (potentially insecure!)
    - intitle:"glype proxy“
    - intitle:"PHPProxy“

**Remember!**

*Don’t be evil*
• All of these tools provide the capability to ensure you and your clients / customers / friends aren’t leaking critical information

• Robots.txt – add an exclusion file to restrict indexing locations
  – Ex: **User-agent: * Disallow: /*

• Use “noindex” page meta tags
  – `<meta name="robots" content="noindex" />

• Password Protect sensitive areas

• Use “nofollow” page meta tags
  – `<meta name="robots" content="noindex" /`
List of tools and resources

- https://github.com/v3n0m-Scanner/V3n0M-Scanner
- https://github.com/tunnelshade/thedumpster
- http://sourceforge.net/projects/dorksearcher/
- http://www.exploit-db.com/google-dorks
• Q & A
• Comments

• Vă mulțumesc pentru timpul acordat!