Lord of the Bing

Taking Back Search Engine Hacking From Google and Bing

30 July 2010

Presented by:
Francis Brown and Rob Ragan
Stach & Liu, LLC
www.stachliu.com
Agenda

OVERVIEW

• Introduction

• Advanced Attacks
  • Google/Bing Hacking
  • Other OSINT Attack Techniques

• Advanced Defenses

• Future Directions
Goals

**DESIRED OUTCOME**

- *To understand* Google Hacking
  - Attacks and defenses
  - Advanced tools and techniques

- *To think differently* about exposures caused by publicly available sources

- To blow your mind!
Introduction/Background

GETTING UP TO SPEED
Open Source Intelligence

**SEARCHING PUBLIC SOURCES**

**OSINT** – is a form of intelligence collection management that involves finding, selecting, and acquiring information from *publicly available* sources and analyzing it to *produce actionable intelligence*. 
# Quick History

## Google Hacking Recap

<table>
<thead>
<tr>
<th>Dates</th>
<th>Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>2004</td>
<td>Google Hacking Database (GHDB) begins</td>
</tr>
<tr>
<td>May 2004</td>
<td>Foundstone SiteDigger v1 released</td>
</tr>
<tr>
<td>2005</td>
<td>Google Hacking v1 released by Johnny Long</td>
</tr>
<tr>
<td>Jan. 2005</td>
<td>Foundstone SiteDigger v2 released</td>
</tr>
<tr>
<td>Jan. 10, 2005</td>
<td>MSNPawn v1.0 released</td>
</tr>
<tr>
<td>Dec. 5, 2006</td>
<td>Google stops issuing Google SOAP API keys</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>
## Quick History

### Google Hacking Recap

<table>
<thead>
<tr>
<th>Dates</th>
<th>Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mar. 2007</td>
<td>Bing disables inurl: link: and linkdomain:</td>
</tr>
<tr>
<td>Nov. 2, 2007</td>
<td>Google Hacking v2 released</td>
</tr>
<tr>
<td>Mar. 2008</td>
<td>cDc Goolag - gui tool released</td>
</tr>
<tr>
<td>June 3, 2009</td>
<td>Bing goes online</td>
</tr>
<tr>
<td>Sept. 7, 2009</td>
<td>Google shuts down SOAP Search API</td>
</tr>
<tr>
<td>Nov. 2009</td>
<td>Binging tool released</td>
</tr>
<tr>
<td>Dec. 1, 2009</td>
<td>FoundStone SiteDigger v 3.0 released</td>
</tr>
<tr>
<td>2010</td>
<td>Googlag.org disappears</td>
</tr>
</tbody>
</table>
Threat Areas

WHAT YOU SHOULD KNOW
Our favorites are Google and Bing
Crawl and Index
Cache and RSS are forever
Query modifiers
  • site:target.com
  • related:target.com
  • filetype:xls
  • ip:69.63.184.142
Attack Targets

GOOGLE HACKING DATABASE

- Advisories and Vulnerabilities (215)
- Error Messages (58)
- Files containing juicy info (230)
- Files containing passwords (135)
- Files containing usernames (15)
- Footholds (21)
- Pages containing login portals (232)

- Pages containing network or vulnerability data (59)
- Sensitive Directories (61)
- Sensitive Online Shopping Info (9)
- Various Online Devices (201)
- Vulnerable Files (57)
- Vulnerable Servers (48)
- Web Server Detection (72)
Attack Targets

Examples

Error Messages

- filetype:asp + "[ODBC SQL"
- "Warning: mysql_query()" "invalid query"

Files containing passwords

- inurl:passlist.txt
Google Hacking Toolkit

STATE OF THE ART

• SiteDigger v3.0
  • Uses Google AJAX API
    • Not blocked by Google
    • But restricted to 64 results/query

• Binging
  • Uses Microsoft Bing search engine
  • Limited domain/ip profiling utils

• Gooscan, Goolag
  • Work still, but get blocked by Google bot detection
  • Download sites no longer around

STACH&LIU
Google Hacking Toolkit

FOUNDSTONE SITEDIGGER
Google Hacking Toolkit

B I N G I N G

Binging

Configuration
Search Query: site:stachliu.com
No of results: 50, default 50, max 1000

Action
Start
Stop

Filter
Unique Hosts
Unique URLs

Output
http://www.stachliu.com/
http://www.stachliu.com/collisions.html
http://www.stachliu.com/research_collisions.html
http://www.stachliu.com/md4coll.c
http://www.stachliu.com/md5coll.c
http://www.stachliu.com/research_publications.html
http://www.stachliu.com/company.html
http://www.stachliu.com/research.html
http://www.stachliu.com/training.html
NEW GOOGLE HACKING TOOLS

DEMO
New Toolkit

GoogleDiggity
- Uses Google AJAX API
  - Not blocked by Google bot detection
- Can Leverage Google custom search

BingDiggity
- Company/Webapp Profiling
  - Enumerate: URLs, IP-to-virtual hosts, etc.
- Bing Hacking Database (BHDB)
  - Regexs in Bing format
New Toolkit

GOOGLE DIGGITY

C:\GoogleDiggity\bin\Debug>GoogleDiggity.exe

==================================================================
| GoogleDiggity - Version 0.1
| Google Hacking Utility - Using Google AJAX API
| By: Fran Brown and Rob Ragan
| Stach & Liu, LLC - http://www.stachliu.com
|==================================================================

** Note: Google AJAX API won’t block you, but does limit you to 64 results per query.

Usage: GoogleDiggity.exe [-q Google_Search_Query] [-d Target_Domain] [-k Google_CustomSearchEngine_ID]

GoogleDiggity.exe [-f Google_Search_Query_File] [-d TargetDomain] [-k GoogleCustomSearchEngineID]

-q [Google search query]
   Optional, can alternatively load queries from file using the -f option.

-f [Google Hack queries file]
   Optional, text file contain Google search queries, one per line. GHDB and Phish.txt files provided (e.g. "Error Messages.txt", "Files containing passwords.txt", etc.).

-d [Domain to search]
   Optional, example: whatever.com

-k [Google Custom Search Engine ID]
   Optional, this can also be set in the configuration file. Note, to use the Google CSE ID in the config file, specify the -k option without any params. For more info, go to http://www.google.com/cse/

Examples:
GoogleDiggity.exe -q test -d stachliu.com
GoogleDiggity.exe -f "Error Messages.txt" -k 006113852371342063645:ytae6riqi4
GoogleDiggity.exe -f "Error Messages.txt" -k
New Toolkit

BINGDIGGITY

C:\BingDiggity\bin\Debug>BingDiggity.exe /?

Usage:
BingDiggity.exe [-q Bing_Search_Query] [-d Target_Domain] [-m Max_Search_Results] [-k Bing_2.0_App_ID]
BingDiggity.exe [-f Bing_Search_Query_File] [-d TargetDomain] [-k Bing_2.0_App_ID]
BingDiggity.exe [-I IP_Address] [-r IP_Address_Range] [-m Max_Search_Results] [-k Bing_2.0_App_ID]

-q [Bing search query]
  Optional, can alternatively load queries from file using the -f option.

-f [Bing Hack queries file]
  Optional, text file contain Bing search queries, one per line.

-d [Domain to search]
  Optional, example: whatever.com

-l [IP address to search]
  Optional, example: 10.1.1.1

-r [IP address range to search]
  Optional, example: 10.1.1.1-10.1.1.15

-m [Max # of results to return per query]
  Optional, note Bing 2.0 doesnt limit your # of queries per day.

-k [Bing 2.0 App ID]
  Required, this can also alternatively be set in the configuration file.
  To obtain a Bing 2.0 App ID, go to http://www.bing.com/dev

Examples:
BingDiggity.exe -q "mySQL error with query"
BingDiggity.exe -d stachliu.com -k 2136a233a5df3f0c9b87c73b9946pfjygbvfd32
BingDiggity.exe -r test -d stachliu.com
BingDiggity.exe -l 10.1.1.1
BingDiggity.exe -r 10.1.1.1-10.1.1.15
Defenses

GOOGLE/BING HACKING DEFENSES

• “Google Hack yourself” organization
  • Employ tools and techniques used by hackers

• Policy and Legal Restrictions

• Regularly update your robots.txt

• Data Loss Prevention/Extrusion Prevention Systems
  • Free Tools: OpenDLP, Senf

• Social Sentry
  • Service to monitor employee FaceBook and Twitter for $2-$8 per employee
    (MySpace, YouTube, and LinkedIn support by summer)
Google Apps Explosion

SO MANY APPLICATIONS TO ABUSE
Google PhoneBook

S P E A R  P H I S H I N G

Residential Phonebook

<table>
<thead>
<tr>
<th>Name</th>
<th>Phone 1</th>
<th>Phone 2</th>
<th>Address</th>
<th>Map</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eric Schmidt</td>
<td>(952) 403-9689</td>
<td></td>
<td>1761 Countryside Dr, Shakopee, MN 55379-4451</td>
<td>Map</td>
</tr>
<tr>
<td>Eric Schmidt</td>
<td>(815) 522-6299</td>
<td>(636) 942-3494</td>
<td>413 Prairie St, Kirkland, IL 80146-0000</td>
<td>Map</td>
</tr>
<tr>
<td>Eric Schmidt</td>
<td>(618) 644-9277</td>
<td></td>
<td>84U4 Lipizzaner Dr, Imperial, MO 63052-4142</td>
<td>Map</td>
</tr>
<tr>
<td>Eric Schmidt</td>
<td>(715) 324-5232</td>
<td></td>
<td>2260 Steinkenig School R, Highland, IL 62248-4006</td>
<td>Map</td>
</tr>
<tr>
<td>Eric Schmidt</td>
<td>(360) 854-0973</td>
<td></td>
<td>N17082 Roth Ln, Pembine, WI 54158-0000</td>
<td>Map</td>
</tr>
<tr>
<td>Eric Schmidt</td>
<td>(650) 964-4017</td>
<td></td>
<td>Mountain View, CA 90403-0000</td>
<td>Map</td>
</tr>
<tr>
<td>Eric Schmidt</td>
<td>(207) 848-2221</td>
<td></td>
<td>41 Hardwood Dr, Hermon, ME 04401-0253</td>
<td>Map</td>
</tr>
</tbody>
</table>
Google Code Search

VULNS IN OPEN SOURCE CODE

- Regex search for vulnerabilities in public code
- Example: SQL Injection in ASP querystring
  - `select.*from.*request\.QUERYSTRING`
DEMO

GOOGLE CODE SEARCH HACKING
SHODAN Computer Search Engine

- Scans and probes the Internet for open HTTP ports and indexes the headers returned in the response
- Profile a target without directly probing their systems
- Discover specific network appliances
- Easily find vulnerable systems!
Target NAS Appliances

SHODAN
Computer Search Engine

"Server: NAShttpd"

Top countries matching your search

Spain
European Union
United Kingdom
Italy

80.14.108.13
Added on 28.01.2010

Default user is ‘admin’

NAS device located

Server: NAShttpd
Target SCADA

CRITICAL INFRASTRUCTURE SECURITY

• Supervisory control and data acquisition
Target SCADA

CRITICAL INFRASTRUCTURE SECURITY

• SHODAN: Target Aquired!
Black Hat SEO
SEARCH ENGINE OPTIMIZATION

• Why use real news events?
• Black hats make their own fake news
• Faux celebrity sex tape anyone?
• Send to college students
• It works!
• Other scammers imitate what works
Google Trends
BLACK HAT SEO RECON

Fake lyrics websites setup to appear in Google search results, infect you with malware upon clicking.

Top Google searches over past 6 years:
- Lada Gaga, Rihanna lyrics sites used to foist Java exploit

As expected, virus writers now are actively exploiting a zero-day Sun vulnerability through drive
Defenses

Blackhat SEO Defenses

• Google SafeBrowsing plugin
• Microsoft SmartScreen Filter
• No-script and Ad-block browser plugins
• Install software security updates
• Stick to reputable sites!
  • Google results aren’t safe.
Metadata Attacks

DATA ABOUT DATA

• It’s everywhere!
  • In documents (doc, xls, pdf)
  • In images

• What can be data mined?
  • Usernames, emails
  • File paths
  • Operating systems, software versions
  • Printers
  • Network information
  • Device information
FOCA

AUTOMETADATA MINING

- Automated doc search via Google/Bing
- Specify domains to target
- Automated download and analysis of docs

STACH&LIU
Defenses

**METADATA MINING DEFENSES**

- Implement a policy to review files for sensitive metadata before they’re released
- Run metadata extraction tools on your resources
- Utilize metadata cleaning tools
- Digital Rights Management (DRM) tools
Advanced Defenses

PROTECT YOUR NECK
Existing Defenses

"H A C K Y O U R S E L F"

✔ Tools exist

✗ Convenient

✗ Real-time updates

✗ Multi-engine results

✗ Historical archived data

✗ Multi-domain searching
Advanced Defenses

NEW HOT SIZZLE

Stach & Liu now proudly presents:

• Google Hacking Alerts
• Bing Hacking Alerts
Google Hacking Alerts

- All GHDB/FSDB regexs using Google alerts
- Real-time vuln updates to 1623 hack queries via RSS
- Organized and available via Google reader importable file
Advanced Defenses

GOOGLE HACKING ALERTS

Several thousand GHDB vulnerability alerts generated in a day

James Bond needs help!
mySQL error page snippet conveniently provided in RSS summary

James Bond 007 :: MI6 - The Home Of James Bond

mySQL error with query SELECT c.item as itemid, c.number as commentid, c.body as body, c.user as user, c.email as userd, c.email as email, ...

Advanced Defenses

Bing Hacking Alerts

- Bing searches with regexs from BHDB
- Leverage `&format=rss` directive to turn into update feeds
Future Direction

PREDICTIONS
Future Directions

PREDICTIONS

Data Explosion
- More data indexed, searchable
- Real-time, streaming updates
- Faster, more robust search interfaces

Google Involvement
- Filtering of search results
- Better GH detection and tool blocking

Renewed Tool Dev
- Google Ajax API based
- Bing/Yahoo/other engines
  - Search engine aggregators
- Google Code and Other Open Source Repositories
  - MS CodePlex, SourceForge, …
- More automation in tools
  - Real-time detection and exploitation
  - Google worms
Future Directions

REAL-TIME UPDATES

Google

Web > Updates > Hide options

Results 1 - 10 of about 4 for obama. (0.58 seconds)

New results will appear below as they become available. Pause

Top links

Obama to discuss Supreme Court pick with party leaders - CNN.com
President Obama is expected to meet with key Republican and Democratic leaders Wednesday to discuss a replacement for retiring Supreme ...

Obama falters on immigration reform promises
from #usnews #news
reform, Obama's priorities shift - latimes.com - latimes.com

Filtered news - Twitter - 1 minute ago

Real-time updates!
Questions?
Ask us something
We’ll try to answer it.

For more info:
Email: contact@stachliu.com
Stach & Liu, LLC
www.stachliu.com
Thank You