Google Hacking against Privacy

Emin İslam Tatlı
tatli@th.informatik.uni-mannheim.de

Department of Computer Science, University of Mannheim
(on leave to the University of Weimar)

Fidis Third International Summer School Karlstad-Sweden,
6-10 August 2007

Outline

1. Google Hacking
2. Privacy Searches
   - Identification Data
   - Sensitive Data
   - Confidential Data
   - Secret Data
3. Countermeasures
4. Future Work

Motivation

- Google has the index size over 20 billion entries
    - try to search `-"fgkdfgjisdfgjsiod"`
- Hackers use google to search vulnerabilities
    - called Google Hacking
    - vulnerable servers, files and applications, files containing usernames-passwords, sensitive directories, online devices, etc.
    - Google Hacking Database [1] ⇒ 1423 entries in 14 groups (by July 2007)
- What about Private Data?
- In this talk, we find out many private data with google

Advanced Search Parameters

- [all]inurl
- [all]intext
- [all]intitle
- site
- ext, filetype
- symbols: - . * |

Examples of Google Hacking I

Unauthenticated programs

```
"PHP Version" intitle:phpinfo inurl:info.php
```

Applications containing SQL injection & path modification vulnerabilities

- `"advanced guestbook * powered" inurl:addentry.php`
- `intitle:"View Img" inurl:viewimg.php`

Security Scanner Reports

```
"Assessment Report" "nessus" filetype:pdf
```

Examples of Google Hacking II

Database applications&error files

- `"Welcome to phpmyadmin ***" "running on * as root@*" intitle:phpmyadmin`
- `"mysql error with query"`

Online Devices

- `inurl:"hp/device/this.LCDispatcher"`
- `intitle:liveapplet inurl:LvAppl`
- `"Please wait....." intitle:"SWW link"`

Privacy Searches

1. Identification Data
2. Sensitive Data
3. Confidential Data
4. Secret Data

Identification Data I

Data related to the personal identity of Users

Name, address, phone, etc.

- `allintext:name email phone address intext:"thomas fischer" ext:pdf`
- `Twiki inurl:"view/Main" "thomas fischer"`

Curriculum Vitae

- `intitle:CV OR intitle:Lebenslauf "thomas fischer"`
- `intitle:CV OR intitle:Lebenslauf ext:pdf OR ext:doc`

Identification Data II

Usernames

- `intitle:"Usage Statistics for" intext:"Total Unique Usernames"`

Sensitive Data I

Data which is normally public but whose reveal may disturb its owner

Postings in Forums and Mailinglists

- `inurl:"search.php?search_author=thomas"`
- `inurl:pipermail "thomas fischer"`

Sensitive Directories

- `intitle:"index of" inurl:"backup"`

Sensitive Data II

Web 2.0

- `"thomas fischer" site:blogspot.com`
- `"thomas" site:flickr.com`
- `"thomas" site:youtube.com`

Confidential Data I

Data that is expected to stay confidential against unauthorized access

Chat Logs

- `"session start" "session ident" thomas ext:txt`

Private Emails

- `"index of" inbox.dbx`
- `"To parent directory" inurl:"Identities"`

Confidential Data II

Confidential Directories and Files

- `"index of" (private | secure | geheim | gizli)`
- `"robots.txt" "User-agent" ext:txt`
- `"This document is private | confidential | secret" ext:doc | ext:pdf | ext:xls`
- `intitle:"index of" "jpg | png | bmp" inurl:personal | inurl:private`

Online Webcams

- `intitle:"Live View / - AXIS" | inurl:view/view.shtml`

Secret Data I

Non-public Data

Usernames and Passwords

- `"create table" "insert into" "pass|passwd|password" (ext:sql|ext:dump|ext:dmp|ext:txt)`
- `"your password * is" (ext:csv | ext:doc | ext:txt)`

Secret Keys

- `"index of" slave_datatrans OR from_master`

Secret Data II

Private Keys

- `"BEGIN (DSA|RSA)" ext:key`
- `"index of" "secring.gpg"`

Encrypted Messages

- `-"public|pubring|pubkeysignature|pgp|and|or|release" ext:gpg`
- `-intext:"and" (ext:enc | ext:axx)`
- `"ciphervalue" ext:xml`

Privacy Countermeasures I

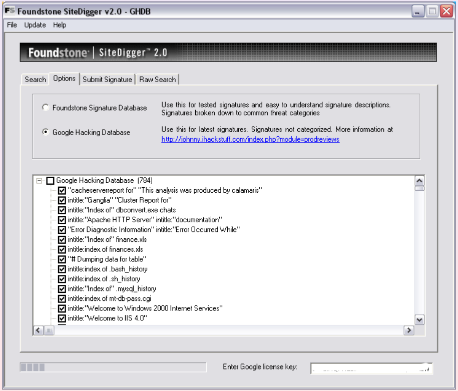
- User-self protection
  - Do not make any sensitive data like documents containing your address, phone numbers, backup directories, secret data like passwords, private emails, etc. online accessible to the public.
  - Provide only required amount of personal information for the Wiki-similar systems.
  - Use more pseudonyms over Internet
  - Considering forum postings and group mails, try to stay anonymous for certain email contents
  - Do not let private media get shared over Web2.0 services
  - Activate authentication mechanisms for your online devices

Privacy Countermeasures II

- System-wide protection
  - Use automatic tools to check your system (e.g. gooscan, sitedigger, goolink)
  - Use Robot Exclusion Standart (robots.txt)
  - Be aware of database backups containing usernames and passwords
  - Install and manage Google Honeypot [2]

Sitedigger [4]

- free from Foundstone company
- supports both GHD and Foundstone's own hacking database
- for a given host, all entries in the database are queried



Foundstone SiteDigger v2.0 - GHDB
File
Update
Help
Foundstone | SiteDigger™ 2.0
Search
Options
Submit Signature
Raw Search

○ Foundstone Signature Database — Use this for tested signatures and easy to understand signature descriptions. Signatures broken down to common threat categories

● Google Hacking Database — Use this for latest signatures. Signatures not categorized. More information at http://johnny.ihackstuff.com/index.php?module=prodreviews

□ Google Hacking Database (784)

- ☑ "cacheserverreport for" "This analysis was produced by calamaris"
- ☑ intitle:"Ganglia" "Cluster Report for"
- ☑ intitle:"Index of" dbconvert.exe chats
- ☑ intitle:"Apache HTTP Server" intitle:"documentation"
- ☑ "Error Diagnostic Information" intitle:"Error Occurred While"
- ☑ intitle:"Index of" finance.xls
- ☑ intitle:index.of finances.xls
- ☑ "# Dumping data for table"
- ☑ intitle:index.of .bash_history
- ☑ intitle:index.of .sh_history
- ☑ intitle:"Index of" .mysql_history
- ☑ intitle:index.of mt-db-pass.cgi
- ☑ intitle:"Welcome to Windows 2000 Internet Services"
- ☑ intitle:"Welcome to IIS 4.0"

Enter Google license key:

Future Work

We are implementing the tool for automatic searches of private data via Google

Conclusion

- Search engines index our private data and make public
- User privacy is in danger
- We need to take the required privacy countermeasures and protect our privacy

References

- Google Hacking Database. http://johnny.ihackstuff.com
- Google Hack Honeypot Project. http://ghh.sourceforge.net
- Goolink- Security Scanner. www.ghacks.net/2005/11/23/goolink-scanner-beta-preview/
- SiteDigger v2.0 - Information Gathering Tool. http://www.foundstone.com
- Gooscan - Google Security Scanner. http://johnny.ihackstuff.com