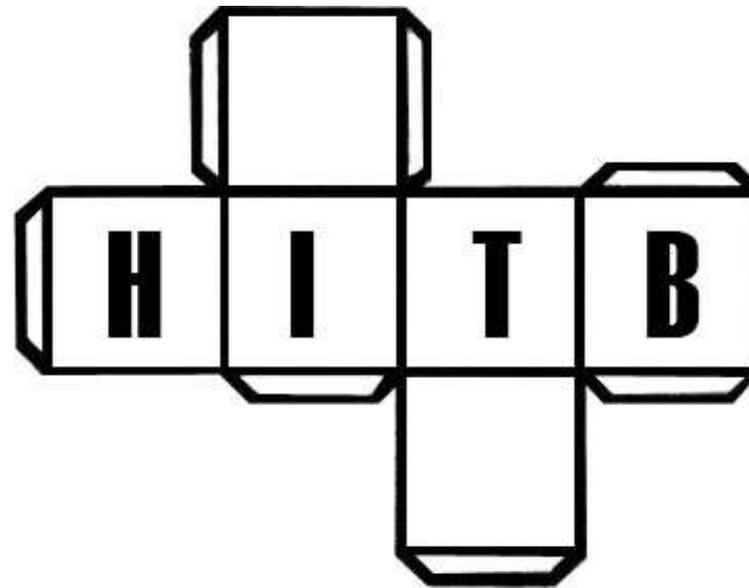


Advanced Information Gathering



2004

Gareth Davies – NSS MSC Sdn Bhd
gareth.davies@mynetsec.com



Outline

- Information Gathering
- What you can Harvest
- The power of modern search engines
- Using Google fully
- The Google API & Google based apps
- Some juicy Google hacks
- Recommendations

Info Gathering (I)

- Generally the first step of a pen test
- Very important for profiling a company
- Leveraging public resources to find private information
- Fully utilized by 'Blackhat' hackers
- By linking information from different sources a complete picture can be built

Info Gathering (II)

- The main thing to note here is **STEALTH**
- These activities are passive and non-intrusive
- Utilizing the memory of the net:
 - Google Cache
 - archive.org
 - Newsgroups (used to be Dejanews)

Example of DNS leakage

```
C:\WINDOWS\System32\cmd.exe - nslookup
> set type=any
> server NS5.ZONEEDIT.COM
Default Server:  NS5.ZONEEDIT.COM
Address:  65.125.228.92

> ls -d hackinthebox.org
[NS5.ZONEEDIT.COM]
hackinthebox.org.      SOA      ns5.zoneedit.com soacontact.zoneedit.com.
(1040609788 14400 7200 950400 7200)
hackinthebox.org.      NS       ns5.zoneedit.com
hackinthebox.org.      NS       ns14.zoneedit.com
hackinthebox.org.      A       203.115.193.176
hackinthebox.org.      MX       0      mail.hackinthebox.org
wap                     A       203.115.193.176
ircweb                  A       69.90.127.43
monitor                 A       210.187.14.98
video                   A       211.24.141.190
ipnews                  A       210.187.14.98
l33tdawg                A       210.187.14.98
l33tblogz               A       210.187.14.98
nakl                    A       210.187.14.98
mail                    A       210.187.14.98
forum                   A       203.115.193.176
www                     A       203.115.193.176
shivandguido            A       210.187.14.98
gateway                 A       210.187.14.98
irc                     A       61.6.39.100
conference              A       203.115.193.176
stats                   A       203.115.193.176
moddef                  A       210.187.14.98
cann                    A       210.187.14.98
hackinthebox.org.      SOA      ns5.zoneedit.com soacontact.zoneedit.com.
(1040609788 14400 7200 950400 7200)
>
```

Info Gathering (IV)

- Such information gathering is now even easier with tools such as
 - dnsreports.com
 - whois.sc
 - nqt.php
 - network-tools.com
 - netcraft.com
 - ip-plus.net/tools/dns_check_set.en.html

Info Gathering (V)

whois.sc reverse IP lookup

Reverse IP Lookup

[Member Area](#) > [Reverse IP Lookup](#)

hackinthebox.org = 203.115.193.176

7 domains found on 203.115.193.176
Search is limited to first 3 results.

Website
www.Atti2de.com
www.CherieIew.com
www.Hackinthebox.org
4 more domains found...

See up to 1000 domains for all IP addresses with [Silver Membership](#)

Enter the IP address (or hostname) of a webserver:

IP Address:

What you can harvest (I)

- Like any kind of hacking, passive information gathering is about thinking outside the box
- Utilizing the many links between information sources is a key
- Picking out useful info is the skill, this activity is akin to modern age dumpster diving

What you can harvest (II)

- Think outside the lines:
 - Check job databases for vacancies
 - Discloses types of technology used
 - Trawl newsgroups for technical postings
 - Sometimes can reveal whole topology
 - Locate company registration details
 - Can give away physical locations
 - Find out personal details about employees
 - May be used in social engineering attacks

What you can harvest (III)

Example usenet posting

All messages from thread

From: [netadmin REMOVE @██████████](#) (netadmin REMOVE @██████████)
Subject: Universal Password Enabling
Newsgroups: [novell.support.native-file-access](#)
Date: 2004-07-20 08:10:39 PST

I have a question regarding Universal Password which I am told needs to be enabled so that Macs can login to our Netware Servers. My network contains 3 Netware Servers 2 are Netware 5.1 SP6 and 1 Netware 6.5 SP1B. Here is how my Tree is setup

```
Tree   - OU#1  contains NW 5.1 Server is Replica Master w/RW of OU#2 & OU#3
        - OU#2  contains NW 6.5 Server is Replica Master w/RW of OU#1 & OU#3
        - OU#3  contains NW 5.1 Server is Replica Master w/RW of OU#2 & OU#1
```

My Mac needs to login to OU#2, so can I enable Universal Password for only OU#2 and not affect OU#1 and OU#3 in anyway. My concern is that I am not completely running Netware 6.5 on every server so I just want to make sure I don't create problems for my other OUs since those servers are in different locations across USA. Also, if I enable Universal Password, can I disable it and be back to square 1 without any after affects?

Thanks

What you can harvest (IV)

Example job posting

System Administrators - Windows NT/2000 (Singapore - City [\[Map\]](#))

Requirements:

Minimum 3-5 years of experience required.

SKILLS REQUIRED:

- Windows 2000 and Active Directory
- Exposure to BMC Patrol or Veritas backup (At least 1 year)
- Preferably a MCSE with at least 5 years of Systems Integration Experience in the areas of NT/Windows 2000, Exchange 2000
- Experience in the deployment of Windows2000 servers and desktops and implementation of email platform (preferably Exchange)
- Should be familiar with the Compaq Server range
- Success attributes include initiative, the ability to perform under pressure, willing to learn, interpersonal and communication skills as well as analytical problem-solving abilities

DESIRED SKILLS:

- Microsoft Clustering Technology
- Microsoft Monitoring Management
- Work experience in the areas of SAN/NAS and wireless platforms

The Power of Search Engines

- Search engines have evolved hugely since the beginning
- They now have MEMORY
 - Google cache
- Advanced search operands now exist
 - filetype:
 - inurl:
 - intitle:

The Power of SEs (II)

- Google is the MIGHT
 - Masses of info, often unrelated
- Teoma is the REFINE
 - Find exactly what you want
- Fast is the DEPTH
 - Locate some obscure parts of the web
- Don't be reliant on one engine

Reference: searchlore.org

The Power of SEs (III)

- Only 3% of people use the advanced features of Google..
- People tend to get locked in to 1 search engine when there are so many
- Each engine has different strengths, learn to utilize them all
- searchlore.org is a great place to learn about the different engines

Using Google Fully (I)

- People often overlook the breadth of Google:
- Google Groups (usenet archive)
- Google images (searchable images)
- Google News (aggregated news)

All of these resources are useful when
information gathering

Using Google Fully (II)

- The most useful advanced operators:
 - site: (restrict search to a single site)
 - [all]intitle: (looks for words in the title)
 - [all]inurl: (look for words within the url)
 - cache: (view the Google cache)
 - filetype: (locate a certain file type)
 - link: (lists links to a given site)
 - related: (view related sites)

Using Google Fully (III)

- Other operators:
 - [+] Essential words (e.g. +word)
 - [-] Words to exclude (e.g. -exclude)
 - [~] Similar to, will include synonyms
 - [.] Single letter wildcard
 - [""] Put around multiple words to search for an exact phrase (e.g. +"my phrase")

The Google API and Apps

- Information about the Google web API can be found here:

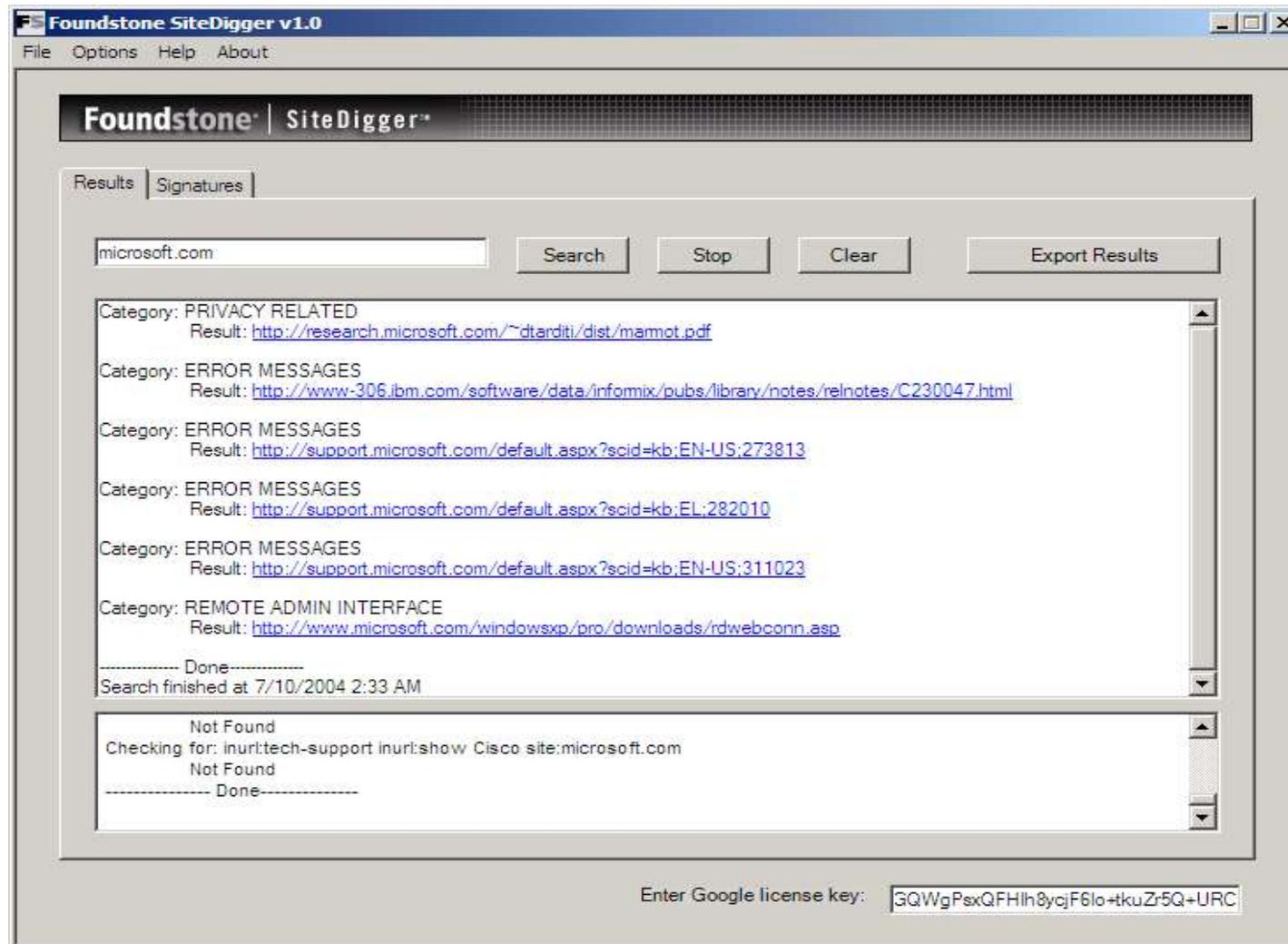
<http://www.google.com/apis/>

- Google is offering a BETA service utilising WSDL and SOAP which allows developers to create applications that can call Google information directly
- Queries are limited to 1000 per day

API and Apps (II)

- Foundstone SiteDigger™ :
 - Uses the Google API
 - Requires Google API licence key (free)
 - Suited to targeting a specific domain or organisation
 - Over 100 pre-defined queries
 - Uses XML so you can add more queries
 - Has auto-update for definitions

API and Apps (III)



API and Apps (IV)

- Buyukada Athena::
 - Doesn't rely on the Google API
 - Suited to finding general misconfigurations
 - Uses extensible XML format which allows support for engines other than Google
 - Developed after SiteDigger was found too limited

API and Apps (V)

The screenshot shows the Athena web browser interface. The title bar reads "Athena". The menu bar includes "File" and "Help". The address bar shows the current URL: `http://www.google.com/search?q="robots.txt"++"Disallow:" filetype:txt+&btnG=Google+Se`. The search engine is set to "Default (Google.com)". The search results are displayed in a list format, showing the following entries:

- "Network Host Assessment Report" "Internet Scanner"
"These statistics were produced by getstats"
- "robots.txt" + "Disallow:" filetype:txt**
- "Thank you for your order" +receipt
- "This file was generated by Nessus"

The query description is: "The query looks for the robots.txt file that contains "disallow" tag telling a web crawler where not to look !".

The search results are displayed in a list format, showing the following entries:

- Web** Results 1 - 10 of about 4,500 for "robots.txt" "Disallow:" filetype:txt . (0.32 seconds)
- [# These are the links for charset converter on the fly # You'll ...](#)
... gif **Disallow:** /photo/main.htm **Disallow:** /photo/nikola/ **Disallow:** /photo/nina/ **Disallow:** /photo/photos/ **Disallow:** /photo/robots.txt **Disallow:** /photo/rollover ...
fly.srk.fer.hr/robots.txt - 4k - [Cached](#) - [Similar pages](#)
- [# robots.txt for http://www.whitehouse.gov/ User-agent: * **Disallow** ...](#)
robots.txt for http://www.whitehouse.gov/ User-agent: * **Disallow:** /cgi-bin **Disallow:** /search **Disallow:** /query.html **Disallow:** /help **Disallow:** /360pics/iraq ...
www.whitehouse.gov/robots.txt - 74k - [Cached](#) - [Similar pages](#)
- [# robots, scram User-agent: * **Disallow:** /cgi-bin **Disallow** ...](#)
robots, scram User-agent: * **Disallow:** /cgi-bin **Disallow:** /TRANSCRIPTS **Disallow:** /development **Disallow:** /third **Disallow:** /beta **Disallow:** /java **Disallow** ...
www.cnn.com/robots.txt - 4k - [Cached](#) - [Similar pages](#)

v1.0 (Acropolis Now)

Juicy Google Hacks

- The number one source for Google Hacks is Johnny Long's site:

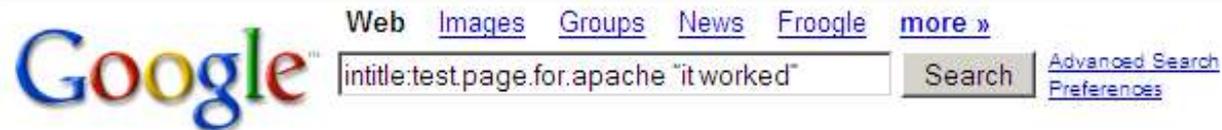
<http://johnny.ihackstuff.com>

- Most of the following info comes from his Google Hacks database

Juicy Google Hacks (II)

- Default Server pages
 - Shows sloppy administration
 - `intitle:test.page.for.apache` "it worked"
 - `allintitle:Netscape FastTrack Server Home Page`
 - `intitle: "Welcome to Windows 2000 Internet Services"`

Juicy Google Hacks (III)



Web

Results 1 - 10 of about 94 for intitle:test.page.for.apache "it worked".

[Test Page for Apache Installation on Web Site](#)

It Worked! The Apache Web Server is Installed on this Web Site! If you can see this page, then the people who own this domain have ...
[mis.cc.ntu.edu.tw/](#) - 2k - [Cached](#) - [Similar pages](#)

[Test Page for Apache](#)

It Worked! If you can see this, then your Apache installation was successful. You may now add content to this directory and replace this page. ...
[www.iaa.com.au/](#) - 1k - [Cached](#) - [Similar pages](#)

[Test Page for Apache Installation on Web Site](#)

Good God! **It Worked!** The Apache Web Server is Installed on this Web Site! ...and its secured. If you can see this page, then the ...
[https://ls.berkeley.edu/](#) - 3k - [Cached](#) - [Similar pages](#)

[Test Page for Apache Installation on Web Site](#)

It Worked! Apache_SSL is Operational! If you can see this page, then the people who own this domain have just installed the Apache ...
[https://www.procurement.virginia.edu/](#) - 2k - [Cached](#) - [Similar pages](#)

Juicy Google Hacks (IV)

- Passwords
 - "Index of" htpasswd / passwd
 - allinurl: admin mdb
 - "config.php"
 - auth_user_file.txt
 - filetype:dat "password.dat"
 - filetype:ini ws_ftp pwd

Juicy Google Hacks (V)



Web

Results 1 - 99 of about 135 for filetype:pwd service

- [FrontPage- ekendall:bYld1Sr73NLKo louisa:5zm94d7cdDFiQ](#)
-FrontPage- ekendall:bYld1Sr73NLKo louisa:5zm94d7cdDFiQ
www.heyerlist.org/garderobe/_vti_pvt/service.pwd - 1k - [Cached](#) - [Similar pages](#)
- [FrontPage- virtex:JrzLYaUcqW8mM nedda:fzdam45JQb5NI](#)
-FrontPage- virtex:JrzLYaUcqW8mM nedda:fzdam45JQb5NI
www.wappi.com/suonerie/_vti_pvt/service.pwd - 1k - [Cached](#) - [Similar pages](#)
- [FrontPage- grahaale:yLSF8Eqk/cQs ftpdch:Zh4nBb7KWKsxl rineerdo ...](#)
-FrontPage- grahaale:yLSF8Eqk/cQs ftpdch:Zh4nBb7KWKsxl rineerdo:caskSSqUyjjzQ
spykecwi:VRhzwct3oVPQ
eclipse.cps.k12.va.us/Schools/DCHS/_vti_pvt/service.pwd - 1k - [Supplemental Result](#) - [Cached](#) - [Similar pages](#)

[FrontPage- grahaale:5XLzoNL12VeNE ftpbrp:Ed8A/1lcpwfqc](#)
-FrontPage- grahaale:5XLzoNL12VeNE ftpbrp:Ed8A/1lcpwfqc
eclipse.cps.k12.va.us/Schools/BRP/_vti_pvt/service.pwd - 1k - [Supplemental Result](#) - [Cached](#) - [Similar pages](#)

[FrontPage- admin:YbV1JnafKRmnQ](#)
-FrontPage- admin:YbV1JnafKRmnQ
librarv.thinkouest.org/C007492F/_vti_pvt/service.pwd - 1k - [Cached](#) - [Similar pages](#)

Sponsored |

[Fortis Escrow Services](#)
A trusted third party to
and minimise financial
www.fortis-escrow.com

[Drivteknik mg ab](#)
Vi levererar kompletta
transmissionslösningar
www.drivteknikmg.se

[See your messa](#)

Juicy Google Hacks (VI)

- There are currently 540 Google Hacks in the database
- Shown were just a few common examples to outline the amount of information available
- Play around, be inventive

This is just the beginning

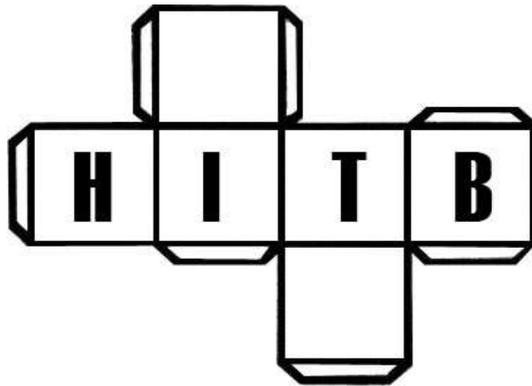
Recommendations

- Disable Directory Browsing
- Do not put sensitive information in web browseable directories
- Don't rely on security through obscurity
- Conduct these tests on your own domains and fix any rogue findings

Conclusion

- Information gathering is important and is being used
- There is no way to know people are doing this
- Be aware of what you have available on the web
- Learn and understand the discussed techniques and tools

Questions?



www.mynetsec.com

