Chapter 2

Footprinting Tools and Techniques

Dr. Ahmed Redha
PSU

Learning Objective

- Identify common information-gathering tools and techniques.

Key Concepts

- Information gathering—reasons, tools, and techniques
- Footprinting as a data-gathering technique
- Footprinting countermeasures
- Google hacking

What is footprinting and reconnaissance?
Collecting general information about a target
without directly interacting withit (mostly passive)

What is Footprinting?

First phase of hacking, where the hacker collects publicly information about the target system or network

•Footprinting is the process of collecting as much information as possible about a target network, for identifying various ways to intrude into an organization's network system.

•Footprinting is the first step of any attack on information systems; attacker gathers publicly available sensitive information, using which he/she performs social engineering, system and network attacks, etc. that leads to huge financial loss and loss of business reputation.

موجودة بالعلن

the goal

if they ask: "why is footprinting important?"
list these

Why Footprinting?

- **Know Security Posture**: Footprinting allows attackers to know the external security posture of the target organization.
  understand how secure the organization looks from outside
- **Reduce Focus Area**: It reduces attacker's focus area to specific range of IP address, networks, domain names, remote access, etc.
  Narrow down IP ranges, domains, systems
- **Identify Vulnerabilities**: It allows attacker to identify vulnerabilities in the target systems in order to select appropriate exploits.
  Find weak points to exploit
- **Draw Network Map**: It allows attackers to draw a map or outline the target organization's network infrastructure to know about the actual environment that they are going to break.
  to understand the network structure

- There are a lot of places you, as an ethical hacker, can get information about your targets, though.
- **Open-source intelligence (OSINT)** is the term that describes identifying information about your target using freely available sources.

from

what is OSINT

Key word

examples:
- Company websites
- Google
- Social media
- Public database

Why OSINT?

The Information-Gathering Process

1. Gathering information from general resources (such as Google or the organizations' website)
2. Determining the network's logical and physical dimensions
3. Identifying active computers and devices
4. Finding open ports, active services, and access points
5. Detecting operating systems
6. Researching known vulnerabilities of running software

The Information-Gathering Process (Cont.)

Footprinting activities:

- Examine company's website
- Identify key employees
- Analyze open positions and job requests
- Assess affiliate, parent, or sister companies
- Find technologies and software used by the organization

.....

The Information-Gathering Process (Cont.)

Footprinting activities (cont.):

- Determine network address and range
- Review network range to determine whether the organization is the owner or if the systems are hosted by someone else
- Look for employee postings, blogs, and other leaked information
- Review collected data

The Information on a Company Website

how company info. gets leaked

- Employee names and email addresses
- Branch office locations
- Technologies the organization uses

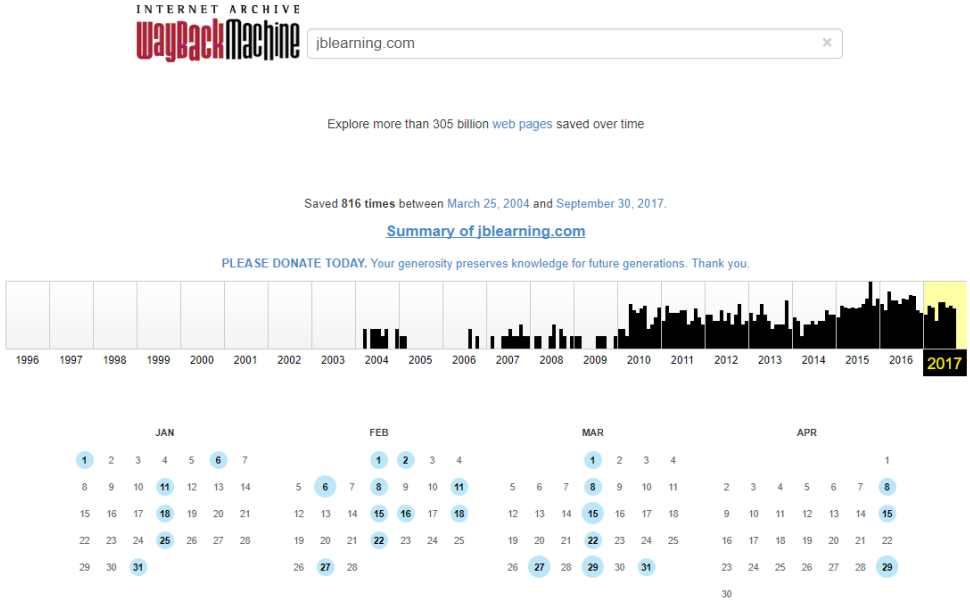
The Information on a Company Website (Cont.)

```
<html>
      <head>
            <title>Company Webpage</title>
      </head>
      <body>
            <! -- This webpage prompts for the password to log
            on to the database server HAL9000 -->
      </body>
</html>
```

their database server

The Information on a Company Website (Cont.)

The Information on a Company Website (Cont.)



The Information on a Company Website (Cont.)

- Job postings give valuable clues into how its infrastructure is organized
- **Example:**
  - Advanced knowledge of Microsoft Windows 8, Windows 10, Server 2008, Server 2012, Server 2016, Microsoft Office 2010, Office 2013, Office 2016, Microsoft SQL Server, Microsoft IIS, Visual Basic
  - Relevant experience/knowledge: Cisco PIX; Check Point Firewall helpful but not necessary
  - VMWare, SAP, and other data-gathering systems
  - Knowledge of Active Directory
  - Experience with Symantec Security Suite

Footprinting using using Advanced Google Hacking Techniques

- Goal is to locate useful information using techniques already provided by the search engine, in new ways
  - Can use Yahoo and Bing as well

Google Advanced Search Operators

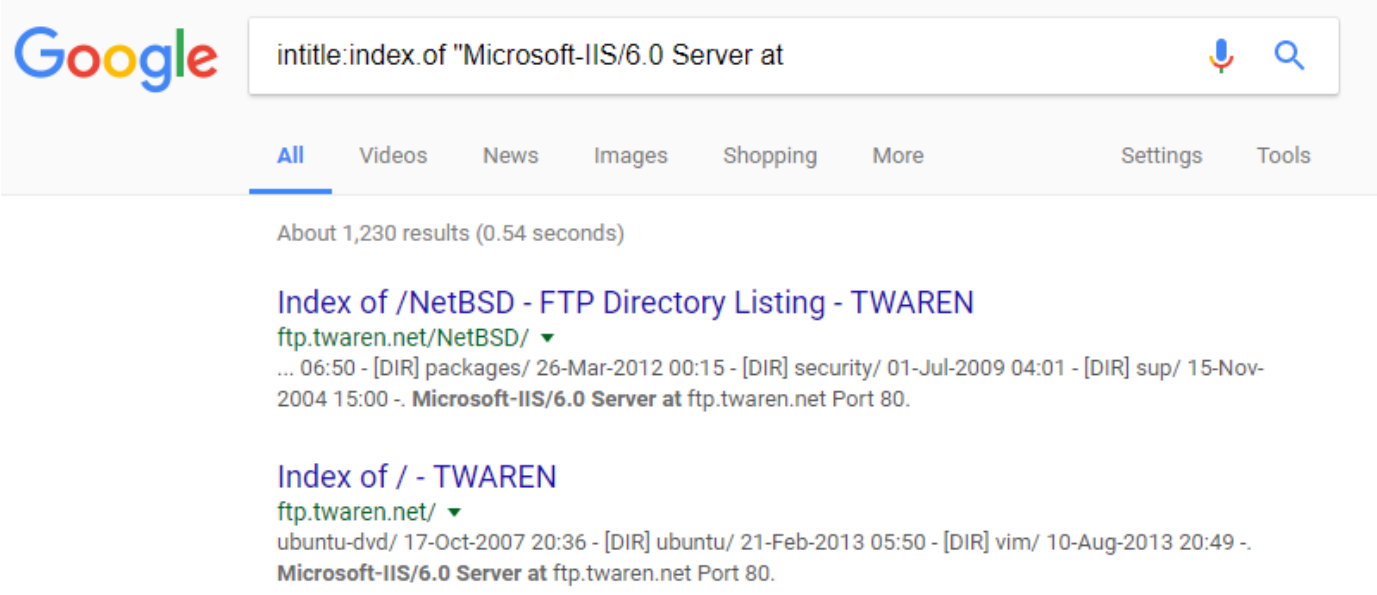
- **`intitle`** instructs Google to search for a term within the title of a document
  - `intitle:"index of" .bash_history`
  - `intitle:"index of" etc/shadow`
  - `intitle:"index.of" finances.xls`
  - `intitle:"index of" htpasswd`
  - `intitle:"Index of" inurl:maillog`

Google Hacking (Cont.)

- **`filetype`** query allows the search to look for a particular term only within a specific filetype

- `filetype:bak inurl:"htaccess|passwd|shadow|htusers"`
- `filetype:conf slapd.conf`
- `filetype:ctt "msn"`
- `filetype:mdb inurl:"account|users|admin|administrators|passwd|password"`
- `filetype:xls inurl:"email.xls"`

Google Hacking (Cont.)



Google Hacking (Cont.)

- **`Inurl`** string is used to search within a site's uniform resource locator (URL)
- Useful if attacker has some knowledge of URL strings or of standard URL strings used by different types of applications and systems
  - `inurl:admin filetype:db`
  - `inurl:admin inurl:backup intitle:index.of`
  - `inurl:"auth_user_file.txt"`
  - `inurl:"/axs/ax-admin.pl" -script`
  - `inurl:"/cricket/grapher.cgi"`

Try this: inurl:"view.shtml" "Network Camera"

Advanced Operators at a Glance

Advanced operators can be combined in some cases.

In other cases, mixing should be avoided.

Some operators can only be used to search specific areas of Google, as these columns show.

Operator	Purpose	Mixes with other operators?	Can be used alone?	Does search work in			
				Web	Images	Groups	News
intitle	Search page title	yes	yes	yes	yes	yes	yes
allintitle	Search page title	no	yes	yes	yes	yes	yes
inurl	Search URL	yes	yes	yes	yes	not really	like intitle
allinurl	Search URL	no	yes	yes	yes	yes	like intitle
filetype	Search specific files	yes	no	yes	yes	no	not really
allintext	Search text of page only	not really	yes	yes	yes	yes	yes
site	Search specific site	yes	yes	yes	yes	no	not really
link	Search for links to pages	no	yes	yes	no	no	not really
inanchor	Search link anchor text	yes	yes	yes	yes	not really	yes
numrange	Locate number	yes	yes	yes	no	no	not really
daterange	Search in date range	yes	no	yes	not really	not really	not really
author	Group author search	yes	yes	no	no	yes	not really
group	Group name search	not really	yes	no	no	yes	not really
insubject	Group subject search	yes	yes	like intitle	like intitle	yes	like intitle
msgid	Group msgid search	no	yes	not really	not really	yes	not really

Advanced Search Operators	Description
site :	Search for the result in the given domain
related :	Search for Similar web pages
cache :	Display the web pages stored in Cache
link :	List the websites having a link to a specific web page
allintext :	Search for websites containing a specific keyword
intext :	Search for documents containing a specific keyword
allintitle :	Search for websites containing a specific keyword in the title
intitle :	Search for documents containing a specific keyword in the title
allinurl :	Search for websites containing a specific keyword in URL
inurl :	Search for documents containing a specific keyword in URL

Exercises

1. Use the site: operator to search for armchairs on IKEA's site, www.ikea.com.

2. Use the Advanced Search form to find the page whose title is "Detect Plagiarism."

3. Find all pages on *google.com* but not on *answers.google.com* nor on *directory.google.com* whose titles include the words "FAQ" or "help."

4. Use the link: operator to see who links to *googleguide.com*, your company's website, or your favorite website.

5. Find pages whose titles include surfing that are not about surfing the World Wide Web.

6. How can you search for [google help] on Google Guide, www.googleguide.com, and on the UC Berkeley library website, www.lib.berkeley.edu?

Google Hacking Database (GHDB)

Google hacking, Google Dorking is a combination of computer hacking techniques that find the security holes within an organization's network and systems using Google search and other applications powered by Google. Google Hacking popularized by Johnny Long. He categorized the queries in a database known as Google Hacking Database (GHDB). This categorized database of queries is designed to uncover the information. This information might be sensitive and not publically available. Google hacking is used to speed up searches.

Google hacking database provide the updated information that is useful for exploitation such as footholds, sensitive directories, vulnerable files, error messages and much more.

- it's a collection of known Google hacking queiries
- it's design to uncover the info. that is sensitive and not Public
- it's used to speed up searches
- it provides updated information

used to:

- Find exposed data
- Password files
- Admin Panels

Google Hacking Database (Cont..)

- Google Hacking Database (GHDB): www.offensive-security.com/community-projects/google-hacking-database/

Footholds (57)
Examples of queries that can help an attacker gain a foothold into a web server

Sensitive Directories (123)
Googles collection of web sites sharing sensitive directories. The files contained in here will vary from sensitive to über-secret!

Vulnerable Files (62)
HUNDREDS of vulnerable files that Google can find on websites.

Vulnerable Servers (83)
These searches reveal servers with specific vulnerabilities. These are found in a different way than the searches found in the "Vulnerable Files" section.

Error Messages (94)
Really verbose error messages that say WAY too much!

Network or Vulnerability Data (70)
These pages contain such things as firewall logs, honeypot logs, network information, IDS logs... All sorts of fun stuff!

Various Online Devices (317)
This category contains things like printers, video cameras, and all sorts of cool things found on the web with Google.

Web Server Detection (80)
These links demonstrate Googles awesome ability to profile web servers.

Files Containing Usernames (17)
These files contain usernames, but no passwords... Still, Google finding usernames on a web site.

Files Containing Passwords (200)
PASSWORDS!!! Google found PASSWORDS!

Sensitive Online Shopping Info (11)
Examples of queries that can reveal online shopping infomation like customer data, suppliers, orders, credit card numbers, credit card info, etc

Files Containing Juicy Info (374)
No usernames or passwords, but interesting stuff none the less.

Pages Containing Login Portals (383)
These are login pages for various services. Consider them the front door of a websites more sensitive functions.

Advisories and Vulnerabilities (1996)
These searches locate vulnerable servers. These searches are often generated from various security advisory posts, and in many cases are product or version-specific.

Exploring Domain Information Leakage

- Manual registrar query → checks domain registration manually
- Automatic registrar query
- Whois
- Nslookup
- Internet Assigned Numbers Authority (IANA)
- Determining a network range

Manual Registrar Query

REGIONAL INTERNET REGISTRY	REGION OF CONTROL
ARIN	North and South America
APNIC	Asia and Pacific
RIPE	Europe, Middle East, and parts of Africa
LACNIC	Latin America and the Caribbean
AFRINIC	Africa

Automatic Registrar Query

- http://www.betterwhois.com
- http://geektools.com
- www.all-nettools.com
- www.smartwhois.com
- www.dnsstuff.com
- http://whois.domaintools.com

Whois

Shows the domain owner, location and contact

Queries databases that hold registration information

Designed to interrogate the Internet DNS and return domain ownership, addresses, location, phone number, etc.

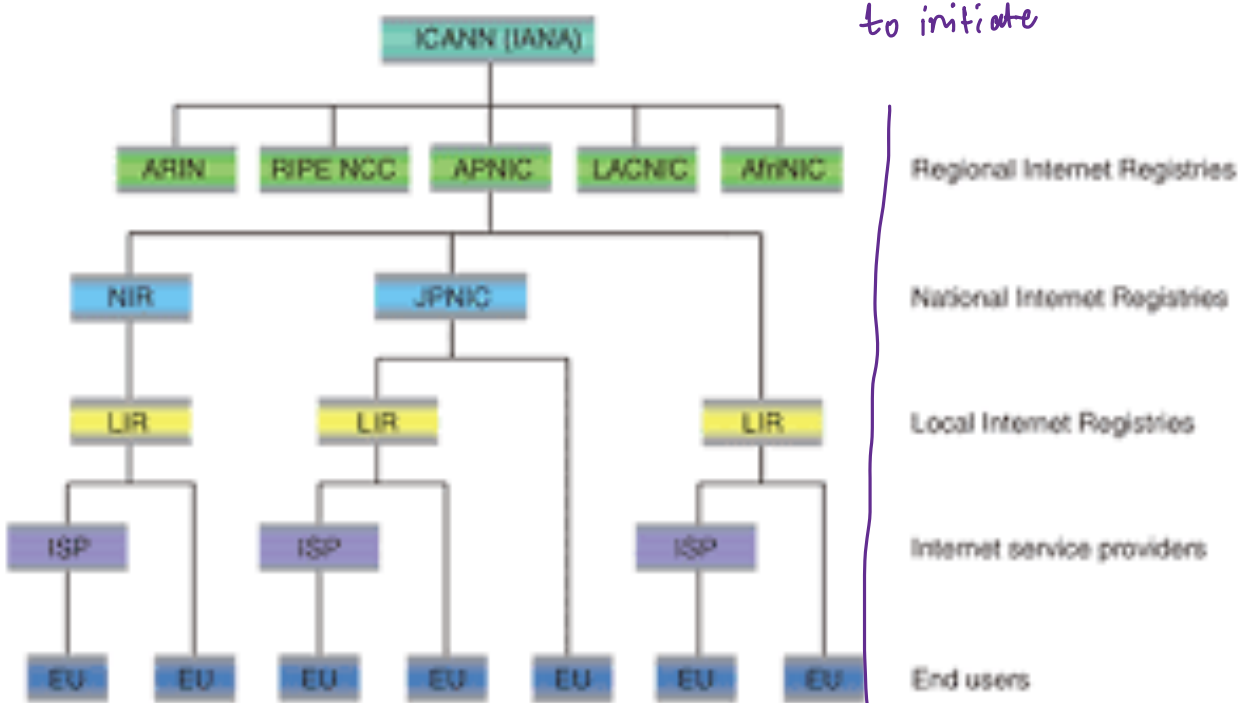
Accessibility of tool depends on operating system

Nslookup

- **Nslookup** is a command-line tool used for querying Domain Name System (DNS) servers to obtain domain name or IP address mapping, or other DNS-related information.
- The term "nslookup" stands for "name server lookup." This tool is available on most operating systems, including Windows, macOS, and various Unix-like systems.

Nslookup (Cont..)

```
nslookup
> set type=mx
> cisco.com
    Server: x.x.x.x
    Address: x.x.x.x#53
Non-authoritative answer:
    cisco.com mail exchanger = 10 smtp3.cisco.com.
    cisco.com mail exchanger = 10 smtp4.cisco.com.
    cisco.com mail exchanger = 10 smtp1.cisco.com.
    cisco.com mail exchanger = 10 smtp2.cisco.com.
Authoritative answers can be found from:
    cisco.com nameserver = ns1.cisco.com.
    cisco.com nameserver = ns2.cisco.com.
    cisco.com nameserver = ns3.cisco.com.
    cisco.com nameserver = ns4.cisco.com.
    ns1.cisco.com internet address = 216.239.32.10
    ns2.cisco.com internet address = 216.239.34.10
    ns3.cisco.com internet address = 216.239.36.10
    ns4.cisco.com internet address = 216.239.38.10
```



who's responsible to initiate

until the target

Determining a Network Range

- American Registry for Internet Numbers (ARIN)
  - North America and parts of the Caribbean
- Réseaux IP Européens Network Coordination Centre (RIPE NCC)
  - Europe, the Middle East, and Central Asia
- Asia-Pacific Network Information Centre (APNIC)
  - Asia and the Pacific Rim
- Latin American and Caribbean Internet Addresses Registry (LACNIC)
  - Latin America and parts of the Caribbean region
- African Network Information Centre (AfrNIC)
  - Africa

Traceroute

- Software program that helps to determine the path a data packet traverses to get to a specific IP address
- One of the easiest ways to identify the path to a targeted website
- In Windows operating systems, the command is `tracert`
- In UNIX/Linux, the command is `traceroute`
- Traceroute displays list of routers on a path to a network destination by using time-to-live (TTL) time-outs and Internet Control Message Protocol (ICMP) error messages

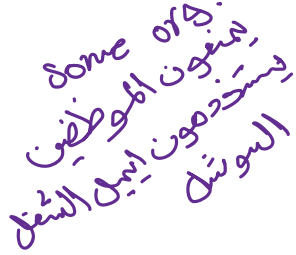
Traceroute (Cont.)

```
C:\tracert www.cisco.com
Tracing route to arin.net [202.131.95.30]
1 1 ms 1 ms 1 ms 192.168.123.254
2 12 ms 15 ms 11 ms adsl-69-151-223-254-dsl.hstntx.swbellnet
  [69.151.223.254]
3 12 ms 12 ms 12 ms 151.164.244.193
4 11 ms 11 ms 11 ms bb1-g14-0.hstntx.sbcglobal.net
  [151.164.92.204]
5 48 ms 51 ms 48 ms 151.164.98.61
6 48 ms 48 ms 48 ms gi1–1.wil04.net.reach.com [206.223.123.11]
7 49 ms 50 ms 48 ms i-0-0-0.wil-core02.bi.reach.com
  [202.84.251.233]
8 196 ms 195 ms 196 ms i-15-0.sydp-core02.bx.reach.com
  [202.84.140.37]
9 204 ms 202 ms 203 ms unknown.net.reach.com [134.159.131.110]
10 197 ms 197 ms 200 ms ssg550-1-r1-1.network.netregistry.net
   [202.124.240.66]
11 200 ms 227 ms 197 ms forward.planetdomain.com
   [202.131.95.30]
```

Tracking an Organization's Employees

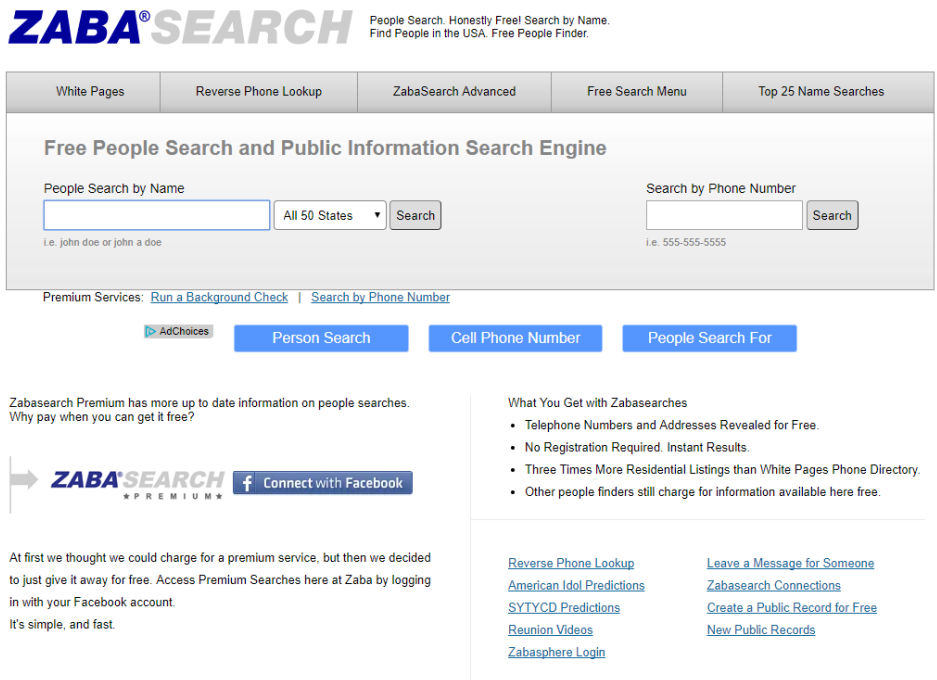
- Information that can be uncovered online:
  - Posted photographs, videos, or information
  - Posted content about personal activities, political or activist affiliations, and beliefs
  - Posting derogatory information about previous employers, coworkers, or clients
  - Discriminatory comments or fabricated qualifications

Tracking an Organization's Employees (Cont.)



- These sites can be examined for names, email addresses, addresses, phone numbers, photographs, and so on:
  - Blogs
  - Personal pages on a social networking site, such as Facebook, YouTube, LinkedIn, Instagram, Twitter, sucks domains
  - Organizational social medial outlet pages

Tracking an Organization's Employees (Cont.)



Using Social Networks

- Facebook, LinkedIn, Twitter, etc.
- Be aware of potentially malicious URLs
  - Example: Clicking a video can be linked to malicious site that steals victims' credentials or other information
  - Website could install malicious software onto the user's system
- Innocent-looking items (trivia items, fantasy sports, etc.) may distribute or publish user information
- Profile information valuable to ethical and nonethical hackers

تقول احفظوا اربعة واقروا الباقي

Using Basic Countermeasures

- Restrict the employees to access social networking sites from organization's network
- Configure web servers to avoid information leakage
- Educate employees to use pseudonyms on blogs, groups, and forums
- Do not reveal critical information in press releases, annual reports, product catalogues, etc.
- Limit the amount of information that you are publishing on the website/Internet
- Use footprinting techniques to discover and remove any sensitive information publicly available
- Prevent search engines from caching a web page and use anonymous registration services
- Enforce security policies to regulate the information that employees can reveal to third parties
- Set apart internal and external DNS or use split DNS, and restrict zone transfer to authorized servers
- Disable directory listings in the web servers
- Educate employees about various social engineering tricks and risks
- Opt for privacy services on Whois Lookup database
- Avoid domain-level cross-linking for the critical assets
- Encrypt and password protect sensitive information

Summary

- Information gathering—reasons, tools, and techniques
- Footprinting as a data-gathering technique
- Footprinting countermeasures
- Google hacking