**Module 2: Footprinting and Reconnaissance**

**Terminology**

* **Footprinting:** collect information about a target network.
* **Passive Footprinting:** collect without direct interaction. Such as search engines, Top-level Domains (TLDs) and sub-domains of a target through web services, social networking sites, competitive intelligence, monitor website traffic of the target
* **Active Footprinting:** collect with direct interaction, such as Harvesting email lists, Whois lookup, extracting DNS info, Traceroute analysis, social engineering, extracting metadata of published documents and files, searching for digital files, querying published name servers of the target
* **Social Network Footprinting:** get information about the target.
* **Website Footprinting:** Information about the target through web pages.

**Methods**

* Examining the web page's source code
* Examining cookies
* Extracting metadata of web sites
* Monitoring website for updates
* Tracking email
* Email header analysis
* Competitive Intelligence Gathering
* Monitoring website traffic
* Tracking online reputation
* WHOIS
* IP geolocation
* DNS footprinting

**Information collected**

* Organization Information (phone numbers, employee details, etc...)
* Relations with other companies
* Network Information (Domains, IPs, etc...)
* System Information (OSes, passwords)

**Objectives of Footprinting:**

* **Know Security Posture**: know the security posture of the target organization
* **Reduce Focus Area**: reduce the attackers focus area to a specific range of IP, network, domain names, etc...
* **Identify Vulnerabilities**: identify vulnerabilities in the target system
* **Draw Network Map**: draw a map or outline the target organization's network infrastructure

**Footprinting Methodology**

* Search engines
* Web services,
* Social networking sites,
* Website footprinting,
* Emai footprinting,
* DNS footprinting,
* Network footpringting,
* Through social engineering

**Google Hacking**

**Operators (No spaces between the operator and the query):**

* **cache:** - Display the web page stored in the google cache
* **link:** - List of web pages that have links to the specified web page
* **related:** - List of web pages that are similar to a specified web page
* **info:** - Presents some information that google has about the particular page
* **site:** - Restrict the results to those websites in the given domain
* **allintitle:** - Restricts the result to those websites with all of the search keywords in the title
* **intitle:** - Restrict the results to documents containing the search keyword in the title
* **allinurl:** - Restrict the results to those with all of the search keywords in the URL
* **inurl:** - Restrict the results to documents containing the search keyword in the URL
* **location:** - Find information for a specific location
* **intext:** - Restrict the results to documents containing the search keyword in the content

**Google Hacking Database (GHDB):** An authoritative source for querying the ever-widing scope of the Google search engine.

**FTP Search Engines:** Search for files located on FTP servers that contain valuable info. Such as NAPALM FTP Indexer, Global FTP Search Engine, and FreewareWeb FTP File Search.

**IoT Search Engines:** Crawl the Internet for IoT devices that are publicly accessible, such as Shodan.

**SCADA:** Supervisory Control and Data Acquisition

**Findding a Company’s TLDS and Sub-domains**

* Sub-domains provide an insight into **different departments and business units**…
* **Sublist3r** python script can enumerates subdomains across multiple sources at once

**Search on Social Networking Sites and People Search Services**

* People search services: Such as Intelius ([www.intelius.com](http://www.intelius.com))
* Gather people and email info from LinkedIn, using **theHarvester**

**Determining the OS**

* Netcraft, SHODAN, CENSYS

**Competitive Intelligence Gathering:** The process of idenfifying, gathering, analylzing, verifying, and using info about your competitors from resources. **Non-interfering** and **subtle** in nature

**Website Footprinting:** The monitoring and analysis of the target org’s website for info

**Tracking Email Communications:** Monitor the delivery of emails

**WHOIS:** Whois databases are maintained by Regional Internet Registries and contain personal information of domain owner

whois uses TCP port 43.

**Example on Linux:**

whois danielgorbe.com

**DNS footprinting:** Reveal info about DNS zone data, including DNS domain names, computer names, IP address, and more.

**DNS record types:**

**A:** Points to a host's IP address

**MX:** Points to a domain's mail server

**NS:** Points to a host's name server

**CNAME:** Canonical naming allows aliases to a host

**SOA:** Indicate authority for domain

**SRV:** Service records

**PTR:** Maps IP address to a hostname

**RP:** Responsible person

**HINFO:** Host information record includes CPU type and OS

**TXT:** Unstructured text records

Example on Linux:

**dig danielgorbe.com**

**Traceroute:** Work on the concept of ICMP protocol and use the TTL field in the header of ICMP packets to discover…

Trace the path between you and your target computer.

**Examples**

**On Windows:** tracert 216.239.39.10

**On Linux:** tcptraceroute [www.google.com/](http://www.google.com/) traceroute [www.google.com](http://www.google.com) **(UDP Traceroute)**

**Footprinting Tools**

* **Maltego:** Be used to determine the relationships and real world links between people, groups, orgs, websites…
* **Recon-ng:** A web reconnaissance framework with independent modules and database interaction
* **FOCA (Fingerprinting Organizations with Collected Archives):** A tool used mainly to find metatada and hidden info in the document it scans.
* **OSRFramework:** Include applications related to username checking, DNS lookups, info leaks research, deep web search…
* **OSINT Framework:** An open source intelligence gathering framework that is focused on gathering info from free tools or resources.
* **Recon-Dog, BillCipher, theHarvester, Th3Inspector, Raccoon, Orb, PENTMENU**

**Countermeasures**

* Resctrict the employees’ access to …
* Configure web servers to avoid info leakage
* Educate employee to use pseudonyms on blogs, groups…
* Limit amount of info published