Pentest +
Information Gathering

Author: Joseph Lee
Email: joseph@ripplesoftware.ca
Mobile: 778-725-3206

Information Gathering General Terms

- **Passive intelligence gathering**
  - Does not engage the target IT infrastructure, people or physical locations
- **Active intelligence gathering**
  - Engages the target IT infrastructure, people or physical locations
- **Foot-printing / mapping**
  - A listing of all IT infrastructure, networks, physical locations, etc. that an organization has
  - **Provided** for **white-box** test
  - **Not provided** for **black-box** test
  - Standards for foot-printing
- **OSINT – Open Source Intelligence**
  - Enumeration
  - Packet crafting
  - Packet capture
  - Packet Inspection
  - Code analysis

Source of Information

Location and Organizational Data

- Maps / GPS / satellite images
- Publicly available images
- Reverse image search
- Social Engineering engagements
- In-person security control testing
  - Entrances / exits
  - Locations of external / internal cameras
  - Locations of offices
  - Locations of restricted areas
  - Badge / Entry access systems
  - Guards
  - Fences
- Dumpster diving

- Paper records
- Property ownership records
- Tax records

OSINT Data Sources

- DNA registrars
- Web-searches
- Security centric databases / websites (Shodan, Censys)
- Social media
- Corporate tax filings
- Any other publicly availably information
- Corporate employees, email addresses, phone numbers
- Social Media

Electronic Documents

- **Exif metadata from documents**
  - **Exiftool**
    - https://exiftool.org/
    - Read or edit metadata
    - For MacOs, Linux, and Windows
    - Platform independent Perl library
    - Linux installation
      - sudo apt-get install libimage-exiftool-perl
    - Can retrieve data
      - GPS location
      - Date / time document was created or modified
      - Image / Camera metadata
- **Employee Information**
  - LinkedIn searches
  - Enrich data found on **LinkedIn / Facebook / GlassDoor / UpLead / hunter.io** or other corporate profile data providers
  - Names and positions
  - Phone numbers
  - Email addresses
  - Company reviews may provide information about the company / people, etc.
- **FOCA – Fingerprinting Organizations with Collected Archives**
  - https://github.com/ElevenPaths/FOCA
  - Scans search engines for document files
  - Scans found files for metadata
  - Scans for software used to create document, email addresses, operating systems, passwords
- **Google Dorking**
  - Search for **PDF files** or other **.doc/.docx, .xls, .xlsx** files available from on

internet which may contain useful proprietary information or classified

- **EDGAR – Electronic Data Gathering Analysis and Retrieval**
  - https://www.sec.gov/edgar.shtml
  - https://www.sec.gov/edgar/searchedgar/companysearch.html
  - SEC Filings
  - Corporate addresses
  - Employee names

IT Infrastructure and Network Mapping

- **Domains**
  - Ownership and organizational information provided by **WHOIS** service
  - Ownership information may be redacted for privacy
  - **gTLD –** General top level domains
  - **ccTLD –** Country code top level domains
  - Central authority is **IANA – Internet Assigned Numbers Authority**
    - https://www.iana.org
    - Africa **– AFRINIC**
      - https://www.afrinic.net
    - Asian / Pacific **– APNIC**
      - https://www.apnic.net
    - North America / Atlantic **– ARIN**
      - https://ws.arin.net
    - Latin America and Caribbean **– LACNIC**
      - https://www.lacnic.net
    - Europe / Russia / Middle East / Central Asia **– RIPE**
      - https:/www.ripe.net
- **Virtual Hosts on same IP**
  - Looking for **other domains on the same server** as the target may provide alternative vectors for gaining access to primary target
  - https://pentest-tools.com/information-gathering/find-virtual-hosts
  - https://hackertarget.com/server-info/
- **Zone transfers (AXFR)**
  - A protocol used to **transfer DNS information** between DNS servers (master / slave)
  - AXFR reveal resource records including subdomain names
  - DNS zone may be **sensitive from an operational security** aspect because information included can be used to discover information about an organization
  - Provide a larger attack surface
  - Usually well protected by DNS servers
  - Can use **host, dig, nmap** to attempt to gain **AXFR** data about a server
  - Initiating an AXFR zone-transfer request from a secondary server
    - dig +short ns zonetransfer.me
    - dig axfr zonetransfer.me @nsztm1.digi.ninja.
  - May includes information such as server host name, primary contact, serial number,

time between changes, TTL for the domain, MX records, name servers, GPS location, TXT records, IP address mappings
    - Robin Wood Practice Site https://digi.ninja/projects/zonetransferme.php
- **IP Ranges**
    - Some organizations own **blocks of IP addresses**
    - If the target organization owns the IP you can see if they own a block / range
    - https://www.whois.com/whois/
    - Remember to check if the target is the owner of the IP block to ensure authorization
- **Routes**
    - **traceroute –** Unix / Linux command
    - **tracert –** Windows command
        - Returns the path the packet take to the IP
        - Some intermediary steps over public internet may be different each time depending on internet traffic
        - On an internal network traceroute may reveal switches or appliances
        - Switches may allow VLAN hopping
    - **External**
        - **Public BGP – Border Gateway Protocol**
            - Route information
            - The **Border Gateway Protocol (BGP)** is the routing protocol used to exchange routing information across the Internet
            - It makes it possible for ISPs to connect to each other and for end-users to connection more than one ISP
            - https://www.bgp4.as
    - **Internal**
        - Internal routes may identify existing hardware in the IT infrastructure

Security Search Engines

- **Shodan**
    - One of the most popular security search engines
    - Pre-built searches and categories
    - Can search sub-categories such as **ICS, databases, IP cameras, etc**
- **Censys**
    - Similar to Shodan
    - Also provides Geo location

Host Enumeration

- Attempt to build a list of all hosts on a network
- One of the most important first steps to an engagement
- Some methods may miss hosts, so comprehensive approach is required
- Methods include:
    - Ping sweep: **for ip in $(seq 1 254); do ping -c 1 192.168.1.$ip; done;**
    - Nmap: **nmap -sn 192.168.1.0/24**

- Nmap: **nmap -sP 192.168.1.***
- ARP Packet capture and look for broadcast packets (whohas)
- Central management systems like SCCM, Jamf Pro
- **DHCP server logs**
  - Contain the device MAC addresses, associated IPs, and hostnames
  - Can be crucial in rapidly identifying a device that has been indicated as being compromised
  - Stored in the C:\Windows\System32\DHCP folder on Windows
  - Stored in /var/lib/dhcp/dhcpd.leases on Linux DHCP server
- **Router logs / network logs / server logs**
  - Access to server logs can reveal IP addresses of end-points that have accessed the server
- **ARP tables**
  - ARP tables can reveal local end-point MAC and IP addresses without having to scan the network
  - Linux: **$ cat /proc/net/arp**
  - MacOs, Linux, and Windows: **arp -a**
- Passive network packet capture reduces change of detection

Service Enumeration / Port scanning

- Port scan hosts and **associate discovered ports with known services**
- Not always trustworthy as **some services may be changed to non-standard ports**
- **Fingerprinting banners** to get **service version information**
  - **NetCat**
    - Host port enumeration
      - nc -v -z [host IP] [port-range]
    - Port service version scan
      - nc [host IP] [port]
- Most port scanners have:
  - Host discovery
  - Port scanning and service identification
  - Service version identification
  - Operating system identification
  - Port ranges
    - **0-1023 –** Well-known ports / system ports
    - **1024-49151 –** Registered ports, assigned by IANA
    - Cannot assume that open port is running the usual service
    - See list of common port assignments in Security+ notes and add the following:
      - **69 TFTP**
      - **123 NTP –** Network Time Protocol
      - **136-139 –** NetBIOS
      - **445 –** Microsoft AD and SMB
      - **515 –** LPD print services
      - **1434 –** Microsoft SQL Monitor

- **1521 –** Oracle database listener

OS Fingerprinting

- Typically done with **TCP/IP stack fingerprinting** that compares responses to **TCP / UDP packets** send by various OSs
- Some OSs have distinct response such as
  - Packet WIN value (rolling window length)
  - Which TCP options they support
  - The order they send packets
  - The service ports open on them (typical Windows services for example)
- Nmap can attempt OS fingerprint
- Sometimes not 100%

Network Topography

- Collect **IP addresses / MAC addresses** and **services scans**
- Visualize with Zenmap or other network topography visualization tool
- Items to map include:
  - Internal LAN hosts / servers with service scans
  - Include subnets
  - Hubs / switches / VLANs
  - Router / gateway / bridges
  - Appliances such as IDS / IPS / mail-server / web-servers / HTTPS proxies
  - Cloud services
  - WAN external network
  - VPN concentrators
  - Peripherals / printers / fax machines / scanners / hybrid business machines
  - ICS – Industrial control systems
  - SCADA – Supervisory Control and Data Acquisition
  - IoT – Internet of things
  - Security cameras / alarm systems

Eavesdropping / Packet Capture

- Capture **ARP WhoHas? management frames** to find IP endpoints on the network
- Capture **addresses of external network resources**
- Capture **passwords and login credentials** if in cleartext
- Capture **session cookies** and authentication processes

SNMP Sweeps

- **SNMP –** Simple Network Management Protocol
- Requires internal access to a network
- Requires the **community string** used by the network devices
- Four possible reasons for lack of response:

    - Wrong community string
    - Unreachable host (offline or fire-walled)
    - SNMP service not running
    - UDP dropped the packet
- **snmpwalk / snmpget**
    - Retrieving information from SNMP-enabled devices
    - **snmpget** uses simple **GET** requests
    - **snmpwalk** uses multiple **GETNEXT** requests

Packet crafting / Inspection

- Manual or tool based packet creation allows **sending malformed or modified / custom packets**
- Packet review and decoding
- Assembling packets from scratch
- Editing existing packets to modify content
- Replaying packets
- **Hping –** Allows creation of custom packets
- **MITM – Man In the Middle**
    - Packet modification / injection for unencrypted traffic
    - Sniff DNS requests

User Enumeration

- Gather usernames for attempting to access unauthorized resources
- Sources:
    - On a host such as Samba (SMB) server
    - Social media / website accounts
    - CMS framework such as WordPress
    - Gather usernames to attempt unauthorized logins

Email Address Enumeration

- Can be used for **phishing campaign**
- To find email addresses associated with organization
    - Check for **mail server response** to service email addresses (**info@, sales@**)
    - Look for organization's email pattern policy (**first.last@, f.last@**)
    - **theHarvester**
        - **-d domain.com**
    - **Metasploit**
        - > use /auxiliary/gather/search_email_collector
        - > set domain domain.com
        - > set outfile filename.txt
        - > exploit
- Use Internet search to check for account breach data
    - **h8mail**

        - https://github.com/khast3x/h8mail
    - https://whatismyipaddress.com/breach-check
    - https://haveibeenpwned.com/
    - https://www.avast.com/hackcheck/friends-check

API and Interface Enumeration

- Discover patterns in API traffic and extract tokens, passwords, etc.
- **Replay** extracted authentication tokens
- Discover **unauthorized data** accessible in an API

Certificate Enumeration and Inspection

- May provide a list of other subdomains or domains located on the server

Code Analysis

- Static code analysis
- Dynamic code analysis
- Script analysis
- Software decompilation
- Software debugging

Reconnaissance Data Storage

- Some scanner applications have **built in databases** to store collected information
    - theHarvester
    - Maltego
    - Recon-ng
    - Metasploit
- Otherwise data can be stored in r**elational database** or **file formats**
- Data collected from **packet capture, port scans,** or other scans should be kept for documentation of pen-test
- Data should be encrypted in storage whenever possible
- Fully encrypted hard-drives help prevent data disclosure of sensitive data

Defences Against Active Reconnaissance

- Limit exposure by removing unneeded services
- Using an IPS or similar appliance to limit or stop probes and scans
- Honeypot / monitoring and alerting systems
- Port knocking

Preventing Passive Information Gathering

- Blacklisting systems / networks / users / IP addresses
- Use **CAPTCHAs** to prevent bots
- Use privacy services to prevent organizational information appearing in domain registration
- Google-dork the domain for **filetype:PDF** or **filetype:DOC** and ensure all documents are classified for public disclosure
- Implement **rate limiting** on server software / WAF to prevent scanning
- Do not publish zone files if possible (gTLDs are required to publish)