# Intelligence Gathering

## Contents

# General

This section defines the Intelligence Gathering activities of a penetration test. The purpose of this document is to provide a standard designed specifically for the pentester performing reconnaissance against a target (typically corporate, military, or related). The document details the thought process and goals of pentesting reconnaissance, and when used properly, helps the reader to produce a highly strategic plan for attacking a target.

# Background Concepts

Levels are an important concept for this document and for PTES as a whole. It’s a maturity model of sorts for pentesting. Defining levels allows us to clarify the expected output and activities within certain real-world constraints such as time, effort, access to information, etc.

The Intelligence Gathering levels are currently split into three categories, and a typical example is given for each one. These should guide the adding of techniques in the document below. For example, an intensive activity such as creating a facebook profile and analyzing the target’s social network is appropriate in more advanced cases, and should be labeled with the appropriate level. See the mindmap below for examples.

## Level 1 Information Gathering

(think: Compliance Driven) Mainly a click-button information gathering process. This level of information can be obtained almost entirely by automated tools. Bare minimum to say you did IG for a PT.

Acme Corporation is required to be compliant with PCI / FISMA / HIPAA. A Level 1 information gathering effort should be appropriate to meet the compliance requirement.

## Level 2 Information Gathering

(think: Best Practice) This level can be created using automated tools from level 1 and some manual analysis. A good understanding of the business, including information such as physical location, business relationships, org chart, etc.

Widgets Inc is required to be in compliance with PCI, but is interested in their long term security strategy, and is acquiring several smaller widget manufacturers. A Level 2 information gathering effort should be appropriate to meet their needs.

## Level 3 Information Gathering

(think: State Sponsored) More advanced pentest, Redteam, full-scope. All the info from level 1 and level 2 along with a lot of manual analysis. Think cultivating relationships on SocNet, heavy analysis, deep understanding of business relationships, most likely a large number of hours to accomplish the gathering and correlation.

An Army Red Team is tasked to analyze and attack a segment of the Army’s network in a foreign country to find weaknesses that could be exploited by a foreign national. A level 3 information gathering effort would be appropriate in this case.

# Intelligence Gathering

## What it is

Intelligence Gathering is performing reconnaissance against a target to gather as much information as possible to be utilized when penetrating the target during the vulnerability assessment and exploitation phases. The more information you are able to gather during this phase, the more vectors of attack you may be able to use in the future.

Open source intelligence (OSINT) is a form of intelligence collection management that involves finding, selecting, and acquiring information from publicly available sources and analyzing it to produce actionable intelligence. [1]

## Why do it

We perform Open Source Intelligence gathering to determine various entry points into an organization. These entry points can be physical, electronic, and/or human. Many companies fail to take into account what information about themselves they place in public and how this information can be used by a determined attacker. On top of that many employees fail to take into account what information they place about themselves in public and how that information can be used to to attack them or their employer.

## What is it not

OSINT may not be accurate or timely. The information sources may be deliberately/accidentally manipulated to reflect erroneous data, information may become obsolete as time passes, or simply be incomplete.

It does not encompass dumpster-diving or any methods of retrieving company information off of physical items found on-premises.

# Target Selection

## Identification and Naming of Target

When approaching a target organization it is important to understand that a company may have a number of different Top Level Domains (TDLs) and auxiliary businesses. While this information should have been discovered during the scoping phase it is not all that unusual to identify additional servers domains and companies that may not have been part of the initial scope that was discussed in the pre-engagement phase. For example a company may have a TDL of .com. However, they may also have .net .co and .xxx. These may need to be part of the revised scope, or they may be off limits. Either way it needs to be cleared with the customer before testing begins. It is also not all that uncommon for a company to have a number of sub-companies underneath them. For example General Electric and Proctor and Gamble own a great deal of smaller companies.

## Consider any Rules of Engagement limitations

At this point it is a good idea to review the Rules of Engagement. It is common for these to get forgotten during a test. Sometimes, as testers we get so wrapped up in what we find and the possibilities for attack that we forget which IP addresses, domains and networks we can attack. Always, be referencing the Rulles of Engagement to keep your tests focused. This is not just important from a legel perspective, it is also important from a scope creep perspective. Every time you get sidetracked from the core objectives of the test it costs you time. And in the long run that can cost your company money.

## Consider time length for test

The amount of time for the total test will directly impact the amount of Intelligence Gathering that can be done. There are some tests where the total time is two to three months. In these engagements a testing company would spend a tremendous amount of time looking into each of the core business units and personal of the company. However, for shorter crystal-box style tests the objectives may be far more tactical. For example, testing a specific web application may not require you to research the financial records of the company CEO.

## Consider end goal of the test

Every test has an end goal in mind - a particular asset or process that the organization considers critical. Having the end result in mind, the intelligence gathering phase should make sure to include all secondary and tertiary elements surrounding the end goal. Be it supporting technologies, 3rd parties, relevant personnel, etc... Making sure the focus is kept on the critical assets assures that lesser relevant intelligence elements are de-prioritized and categorized as such in order to not intervene with the analysis process.

# OSINT

Open Source Intelligence (OSINT) takes three forms; Passive, Semi-passive, and Active.

- **Passive Information Gathering**: Passive Information Gathering is generally only useful if there is a very clear requirement that the information gathering activities never be detected by the target. This type of profiling is technically difficult to perform as we are never sending any traffic to the target organization neither from one of our hosts or "anonymous" hosts or services across the Internet. This means we can only use and gather archived or stored information. As such this information can be out of date or incorrect as we are limited to results gathered from a third party.
- **Semi-passive Information Gathering**: The goal for semi-passive information gathering is to profile the target with methods that would appear like normal Internet traffic and behavior. We query only the published name servers for information, we aren't performing in-depth reverse lookups or brute force DNS requests, we aren't searching for "unpublished" servers or directories. We aren't running network level portscans or crawlers and we are only looking at metadata in published documents and files; not actively seeking hidden content. The key here is not to draw attention to our activities. Post mortem the target may be able to go back and discover the reconnaissance activities but they shouldn't be able to attribute the activity back to anyone.
- **Active Information Gathering**: Active information gathering should be detected by the target and suspicious or malicious behavior. During this stage we are actively mapping network infrastructure (think full port scans nmap –p1-65535), actively enumerating and/or vulnerability scanning the open services, we

are actively searching for unpublished directories, files, and servers. Most of this activity falls into your typically “reconnaissance” or “scanning” activities for your standard pentest.

# Corporate

## Physical

### Locations (L1)

Per location listing of full address, ownership, associated records (city, tax, legal, etc), Full listing of all physical security measures for the location (camera placements, sensors, fences, guard posts, entry control, gates, type of identification, supplier’s entrance, physical locations based on IP blocks/geolocation services, etc… For Hosts/NOC: Full CIDR notation of hosts and networks, full DNS listing of all associated assets, Full mapping of AS, peering paths, CDN provisioning, netblock owners (whois data), email records (MX + mail address structure)

- Owner (L1/L2)
- Land/tax records (L1/L2)
- Shared/individual (L1/L2)
- Timezones (L1/L2)
- Hosts / NOC

### Pervasiveness (L1)

It is not uncommon for a target organization to have multiple separate physical locations. For example, a bank will have central offices, but they will also have numerous remote branches as well. While physical and technical security may be very good at central locations, remote locations often have poor security controls.

**Relationships (L1)**

Business partners, customs, suppliers, analysis via whats openly shared on corporate web pages, rental companies, etc. This information can be used to better understand the business or organizational projects. For example, what products and services are critical to the target organization?

Also, this information can also be used to create successful social engineering scenarios.

- Relationships (L2/L3)

  Manual analysis to vet information from level 1, plus dig deeper into possible relationships.

- Shared office space (L2/L3)
- Shared infrastructure (L2/L3)
- Rented / Leased Equipment (L2/L3)

## Logical

Accumulated information for partners, clients and competitors: For each one, a full listing of the business name, business address, type of relationship, basic financial information, basic hosts/network information.

- Business Partners (L1/L2/L3)

Target’s advertised business partners. Sometimes advertised on main www.

- Business Clients (L1/L2/L3)

  Target’s advertised business clients. Sometimes advertised on main www.

- Competitors (L1/L2/L3)

  Who are the target’s competitors. This may be simple, Ford vs Chevy, or may require much more analysis.

- Touchgraph (L1)

  A touchgraph (visual representation of the social connections between people) will assist in mapping out the possible interactions between people in the organization, and how to access them from the outside (when a touchgraph includes external communities and is created with a depth level of above 2).
  The basic touchgraph should reflect the organizational structure derived from the information gathered so far, and further expansion of the graph should be based on it (as it usually represents the focus on the organizational assets better, and make possible approach vectors clear.

- Hoovers profile (L1/L2)

  What: a semi-open source intelligence resource (paid subscriptions usually). Such sources specialize in gathering business related information on companies, and providing a

“normalized” view on the business.
Why: The information includes physical locations, competitive landscape, key personnel, financial information, and other business related data (depending on the source). This can be used to create a more accurate profile of the target, and identify additional personnel and 3rd parties which can be used in the test.
How: Simple search on the site with the business name provide the entire profile of the company and all the information that is available on it. Its recommended to use a couple of sources in order to cross reference them and make sure you get the most up-to-date information. (paid for service).

- Product line (L2/L3)

  Target's product offerings which may require additional analysis if the target does offer services as well this might require further analysis.

- Market Vertical (L1)

  Which industry the target resides in. i.e. financial, defense, agriculture, government, etc

- Marketing accounts (L2/L3)

  Marketing activities can provide a wealth of information on the marketing strategy of the target
  Evaluate all the social media Networks for the target's social personas
  Evaluate the target's past * marketing campaigns

- Meetings (L2/L3)

  Meeting Minutes published?
  Meetings open to public?

- Significant company dates (L1/L2/L3)

  Board meetings
  Holidays
  Anniversaries
  Product/service launch

- Job openings (L1/L2)

  By viewing a list of job openings at an organization (usually found in a ‘careers’ section of their website), you can determine types of technologies used within the organization. One example would be if an organization has a job opening for a Senior Solaris Sysadmin then it is pretty obvious that the organization is using Solaris systems. Other positions may not be as obvious by the job title, but an open Junior Network Administrator position may say something to the effect of ‘CCNA preferred’ or ‘JNCIA preferred’ which tells you that they are either using Cisco or Juniper technologies.

- Charity affiliations (L1/L2/L3)

  It is very common for executive members of a target organization to be associated with charitable organizations. This information can be used to develop solid social engineering scenarios for targeting executives.

- RFP, RFQ and other Public Bid Information (L1/L2)

  RFPs and RFQs often reveal a lot of information about the types of systems used by a company, and potentially even gaps or issues with their infrastructure.
  Finding out who current bid winners are may reveal the types of systems being used or a location where company resources might be hosted off-site.

- Court records (L2/L3)

  Court records are usually available either free or sometimes at a fee.
  Contents of litigation can reveal information about past complainants including but not limited to former employee lawsuits
  Criminal records of current and past employees may provide a list of targets for social engineering efforts

- Political donations (L2/L3)

  Mapping out political donations or other financial interests is important in order to identify pivotal individuals who may not be in obvious power positions but have a vested interest (or there is a vested interes in them).
  Political donation mapping will change between countries based on the freedom of information, but often cases donations from other countries can be traced back using the data available there.

- Professional licenses or registries (L2/L3)

Gathering a list of your targets professional licenses and registries may offer an insight into not only how the company operated, but also the guidelines and regulations that they follow in order to maintain those licenses. A prime example of this is a companies ISO standard certification can show that a company follows set guidelines and processes. It is important for a tester to be aware of these processes and how they could affect tests being performed on the organization.

A company will often list these details on their website as a badge of honor. In other cases it may be necessary to search registries for the given vertical in order to see if an organization is a member. The information that is available is very dependent on the vertical market, as well as the geographical location of the company. It should also be noted that international companies may be licensed differently and be required to register with different standards or legal bodies dependent on the country.

## Org Chart (L1)

- Position identification
  - Important people in the organization
  - Individuals to specifically target
- Transactions
  - Mapping on changes within the organization (promotions, lateral movements)
- Affiliates
  - Mapping of affiliate organizations that are tied to the business

## Electronic

### Document Metadata (L1/L2)

- What it is? Metadata or meta-content provides information about the data/document in scope. It can have information such as author/creator name, time and date, standards used/referred, location in a computer network (printer/folder/directory path/etc. info), geo-tag etc. For an image its' metadata can contain color, depth, resolution, camera make/type and even the co-ordinates and location information.
- Why you would do it? Metadata is important because it contains information about the internal network, user-names, email addresses, printer locations etc. and will help to create a blueprint of the location. It also contains information about software used in creating the respective documents. This can enable an attacker to create a profile and/or perform targeted attacks with internal knowledge on the networks and users.
- How you would do it? There are tools available to extract the metadata from the file (pdf/word/image) like FOCA (GUI-based), metagoofil (python-based), meta-extractor, exiftool (perl-based). These tools are capable of extracting and displaying the results in different formats as HTML, XML, GUI, JSON etc. The input to these tools is mostly a document downloaded from the public presence of the 'client' and then analyzed to know more about it. Whereas FOCA helps you search documents, download and analyzes all through its GUI interface.

### Marketing Communications (L1/L2)

- Past marketing campaigns provide information for projects which

might of been retired that might still be accessible.
- Current marketing communications contain design components (Colors, Fonts, Graphics etc..) which are for the most part used internally as well.
- Additional contact information including external marketing organizations.

## Infrastructure Assets

### Network blocks owned (L1)

- Network Blocks owned by the organization can be passively obtained from performing whois searches. DNSStuff.com is a one stop shop for obtaining this type of information.
- Open Source searches for IP Addresses could yield information about the types of infrastructure at the target. Administrators often post ip address information in the context of help requests on various support sites.

### Email addresses (L1)

- E-mail addresses provide a potential list of valid usernames and domain structure
- E-mail addresses can be gathered from multiple sources including the organizations website.

### External infrastructure profile (L1)

- The target's external infrastructure profile can provide immense information about the technologies used internally.
- This information can be gathered from multiple sources both passively and actively.

- The profile should be utilized in assembling an attack scenario against the external infrastructure.

### Technologies used (L1/L2)

- OSINT searches through support forums, mailing lists and other resources can gather information of technologies used at the target
- Use of Social engineering against the identified information technology organization
- Use of social engineering against product vendors

### Purchase agreements (L1/L2/L3)

- Purchase agreements contain information about hardware, software, licenses and additional tangible asset in place at the target.

### Remote access (L1/L2)

- Obtaining information on how employees and/or clients connect into the target for remote access provides a potential point of ingress.
- Often times link to remote access portal are available off of the target's home page
- How To documents reveal applications/procedures to connect for remote users

### Application usage (L1/L2)

Gather a list of known application used by the target organization. This can often be achieved by extracting metadata from publicly accessible files (as discussed previously)

### Defense technologies (L1/L2/L3)

Fingerprinting defensive technologies in use can be achieved in a number of ways depending on the defenses in use.

### Passive fingerprinting

- Search forums and publicly accessible information where technicians of the target organisation may be discussing issues or asking for assistance on the technology in use
- Search marketing information for the target organisation as well as popular technology vendors
- Using Tin-eye (or another image matching tool) search for the target organisations logo to see if it is listed on vendor reference pages or marketing material

### Active fingerprinting

- Send appropriate probe packets to the public facing systems to test patterns in blocking. Several tools exist for fingerprinting of specific WAF types.
- Header information both in responses from the target website and within emails often show information not only on the systems in use, but also the specific protection mechanisms enabled (e.g. Email gateway Anti-virus scanners)

### Human capability (L1/L2/L3)

Discovering the defensive human capability of a target organization can be difficult. There are several key pieces of information that could assist in judging the security of the target organization.

- Check for the presence of a company-wide CERT/CSIRT/PSRT team
- Check for advertised jobs to see how often a security position is listed
- Check for advertised jobs to see if security is listed as a requirement for non-security jobs (e.g. developers)
- Check for out-sourcing agreements to see if the security of the target has been outsourced partially or in it's entirety
- Check for specific individuals working for the company that may be active in the security community

## Financial

### Reporting (L1/L2)

The targets financial reporting will depend heavily on the location of the organization. Reporting may also be made through the organizations head office and not for each branch office. In 2008 the SEC issued a proposed roadmap for adoption of the International Financial Reporting Standards (IFRS) in the US.

IFRS Adoption per country --> http://www.iasplus.com/en/resources/use-of-ifrs

### Market analysis (L1/L2/L3)

- Obtain market analysis reports from analyst organizations (such as Gartner, IDC, Forrester, 541, etc...). This should include what the market definition is, market cap, competitors, and any major changes to the valuation, product, or company in general.

### Trade capital

- Identify is the organization is allocating any trade capital, and in what percentage of the overall valuation and free capital it has. This will indicate how sensitive the organization is to market fluctuations, and whether it depends on external investment as part of it's valuation and cash flow.

### Value history

- Charting of the valuation of the organization over time, in order to establish correlation between external and internal events, and their effect on the valuation.

### EDGAR (SEC)

- What is it: EDGAR (the Electronic Data Gathering, Analysis, and Retrieval system) is a database of the U.S. Security and Exchanges Commission (SEC) that contains registration statements, periodic reports, and other information of all companies (both foreign and domestic) who are required by law to file.
- Why do it: EDGAR data is important because, in additional to financial information, it identifies key personnel within a company that may not be otherwise notable from a company's website or other public presence. It also includes statements of executive compensation, names and addresses of major common stock owners, a summary of legal proceedings against the company, economic risk factors, and other potentially interesting data.
- How to obtain: The information is available on the SEC's EDGAR website (http://www.sec.gov/edgar.shtml). Reports of particular

interest include the 10-K (annual report) and 10-Q (quarterly report).

# Individual

## Employee

### History

- Court Records (L2/L3)
  - What is it: Court records are all the public records related to criminal and/or civil complaints, lawsuits, or other legal actions for or against a person or organization of interest.
  - Why you would do it: Court records could potentially reveal sensitive information related to an individual employee or the company as a whole. This information could be useful by itself or may be the driver for gaining additional information. It could also be used for social engineering or other purposes later on in the penetration test.
  - How you would do it: Much of this information is now available on the Internet via publicly available court websites and records databases. Some additional information may be available via pay services such as LEXIS/NEXIS. Some information may be available via records request or in person requests.
- Political Donations (L2/L3)
  - What is it: Political donations are an individual's personal funds directed to specific political candidates, political parties, or special interest organizations.
  - Why you would do it: Information about political donations could potentially reveal useful information related to an

individual. This information could be used as a part of social network analysis to help draw connections between individuals and politicians, political candidates, or other political organizations. It could also be used for social engineering or other purposes later on in the penetration test.

  - How you would do it: Much of this information is now available on the Internet via publicly available websites (i.e., http://www.opensecrets.org/) that track political donations by individual. Depending upon the laws of a given state, donations over a certain amount are usually required to be recorded.
- Professional licenses or registries (L2/L3)
  - What is it: Professional licenses or registries are repositories of information that contain lists of members and other related information for individuals who have attained a particular license or some measure of specific affiliation within a community.
  - Why you would do it: Information about professional licenses could potentially reveal useful information related to an individual. This information could be used to validate an individual's trustworthiness (do they really have a particular certification as they claim) or as a part of social network analysisto help draw connections between individuals and other organizations. It could also be used for social engineering or other purposes later on in the penetration test.
  - How you would do it: Much of this information is now available on the Internet via publicly available websites. Typically, each organization maintains their own registry of

information that may be available online or may require additional steps to gather.

### Social Network (SocNet) Profile

- Metadata Leakage (L2/L3)
    - Location awareness via Photo Metadata
- Tone (L2/L3)
    - Expected deliverable: subjective identification of the tone used in communications – aggressive, passive, appealing, sales, praising, dissing, condescending, arrogance, elitist, underdog, leader, follower, mimicking, etc…
- Frequency (L2/L3)
    - Expected deliverable: Identification of the frequency of publications (once an hour/day/week, etc…). Additionally - time of day/week in which communications are prone to happen.
- Location awareness (L2/L3)

    Map location history for the person profiled from various sources, whether through direct interaction with applications and social networks, or through passive participation through photo metadata.

    - Bing Map Apps
    - Foursquare
    - Google Latitude
    - Yelp
    - Gowalla
- Social Media Presence (L1/L2/L3)

Verify target’s social media account/presence (L1). And provide detailed analysis (L2/L3)

### Internet Presence

- Email Address (L1)
    - What it is? Email addresses are the public mail box ids of the users.
    - Why you would do it? Email address harvesting or searching is important because it serves multiple purposes - provides a probable user-id format which can later be brute-forced for access but more importantly it helps sending targeted spams and even to automated bots. These spam emails can contain exploits, malware etc. and can be addressed with specific content particularly to a user.
    - How you would do it? Email addresses can be searched and extracted from various websites, groups, blogs, forums, social networking portals etc. These email addresses are also available from various tech support websites. There are harvesting and spider tools to perform search for email addresses mapped to a certain domain (if needed).
- Personal Handles/Nicknames (L1)
- Personal Domain Names registered (L1/L2)
- Assigned Static IPs/Netblocks (L1/L2)

### Physical Location

- Physical Location
    - Can you derive the target's physical location

### Mobile Footprint

- Phone number (L1/L2/L3)
- Device type (L1/L2/L3)
- Use (L1/L2/L3)
- Installed applications (L1/L2/L3)
- Owner/administrator (L1/L2/L3)

#### "For Pay" Information

- Background Checks
- For Pay Linked-In
- LEXIS/NEXIS

# Covert Gathering

# Corporate

## On-Location Gathering

Selecting specific locations for onsite gathering, and then performing reconnaissance over time (usually at least 2-3 days in order to assure patterns). The following elements are sought after when performing onsite intelligence gathering:

- Physical security inspections
- Wireless scanning / RF frequency scanning
- Employee behavior training inspection
- Accessible/adjacent facilities (shared spaces)
- Dumpster diving
- Types of equipment in use

## Offsite Gathering

Identifying offsite locations and their importance/relation to the organization. These are both logical as well as physical locations as per the below:

- Data center locations
- Network provisioning/provider

# HUMINT

Human intelligence complements the more passive gathering on the asset as it provides information that could not have been obtained otherwise, as well as add more “personal” perspectives to the intelligence picture (feelings, history, relationships between key individuals, “atmosphere”, etc…)

The methodology of obtaining human intelligence always involves direct interaction - whether physical, or verbal. Gathering should be done under an assumed identity, that would be created specifically to achieve optimal information exposure and cooperation from the asset in question.

Additionally, intelligence gathering on more sensitive targets can be performed by utilizing observation only - again, either physically on location, or through electronic/remote means (CCTV, webcams, etc…). This is usually done in order to establish behavioral patterns (such as frequency of visitations, dress code, access paths, key locations that may provide additional access such as coffee shops).

## Results

- Key Employees

- Partners/Suppliers
- Social Engineering

# Footprinting

WHAT IT IS: External information gathering, also known as footprinting, is a phase of information gathering that consists of interaction with the target in order to gain information from a perspective external to the organization.

WHY: Much information can be gathered by interacting with targets. By probing a service or device, you can often create scenarios in which it can be fingerprinted, or even more simply, a banner can be procured which will identify the device. This step is necessary to gather more information about your targets. Your goal, after this section, is a prioritized list of targets.

# External Footprinting

## Identify Customer External Ranges

One of the major goals of intelligence gathering during a penetration test is to determine hosts which will be in scope. There are a number of techniques which can be used to identify systems, including using reverse DNS lookups, DNS bruting, WHOIS searches on the domains and the ranges. These techniques and others are documented below.

## Passive Reconnaissance

### WHOIS Lookups

For external footprinting, we first need to determine which one of the WHOIS servers contains the information we're after. Given that we should know the TLD for the target domain, we simply have to locate the Registrar that the target domain is registered with.

WHOIS information is based upon a tree hierarchy. ICANN (IANA) is the authoritative registry for all of the TLDs and is a great starting point for all manual WHOIS queries.

- ICANN - http://www.icann.org
- IANA - http://www.iana.com
- NRO - http://www.nro.net
- AFRINIC - http://www.afrinic.net
- APNIC - http://www.apnic.net
- ARIN - http://ws.arin.net
- LACNIC - http://www.lacnic.net
- RIPE - http://www.ripe.net

Once the appropriate Registrar was queried we can obtain the Registrant information. There are numerous sites that offer WHOIS information; however for accuracy in documentation, you need to use only the appropriate Registrar.

- InterNIC - http://www.internic.net/ http://www.internic.net]

Typically, a simple whois against ARIN will refer you to the correct registrar.

### BGP looking glasses

It is possible to identify the Autonomous System Number (ASN) for networks that participate in Border Gateway Protocol (BGP). Since

BGP route paths are advertised throughout the world we can find these by using a BGP4 and BGP6 looking glass.

- BGP4 - http://www.bgp4.as/looking-glasses
- BPG6 - http://lg.he.net/

## Active Footprinting

### Port Scanning

Port scanning techniques will vary based on the amount of time available for the test, and the need to be stealthy. If there is zero knowledge of the systems, a fast ping scan can be used to identify systems. In addition, a quick scan without ping verification (-PN in nmap) should be run to detect the most common ports avialable. Once this is complete, a more comprehensive scan can be run. Some testers check for only open TCP ports, make sure to check UDP as well. The http://nmap.org/nmap_doc.html document details port scan types. Nmap ("Network Mapper") is the de facto standard for network auditing/scanning. Nmap runs on both Linux and Windows.

You can find more information on the use of Nmap for this purpose in the PTES Technical Guideline

Nmap has dozens of options available. Since this section is dealing with port scanning, we will focus on the commands required to perform this task. It is important to note that the commands utilized depend mainly on the time and number of hosts being scanned. The more hosts or less time that you have to perform this tasks, the less that we will interrogate the host. This will become evident as we continue to discuss the options.

IPv6 should also be tested.

### Banner Grabbing

Banner Grabbing is an enumeration technique used to glean information about computer systems on a network and the services running its open ports. Banner grabbing is used to identify network the version of applications and operating system that the target host are running.

Banner grabbing is usually performed on Hyper Text Transfer Protocol (HTTP), File Transfer Protocol (FTP), and Simple Mail Transfer Protocol (SMTP); ports 80, 21, and 25 respectively. Tools commonly used to perform banner grabbing are Telnet, nmap, and Netcat.

### SNMP Sweeps

SNMP sweeps are performed too as they offer tons of information about a specific system. The SNMP protocol is a stateless, datagram oriented protocol. Unfortunately SNMP servers don't respond to requests with invalid community strings and the underlying UDP protocol does not reliably report closed UDP ports. This means that "no response" from a probed IP address can mean either of the following:

- machine unreachable
- SNMP server not running
- invalid community string
- the response datagram has not yet arrived

### Zone Transfers

DNS zone transfer, also known as AXFR, is a type of DNS transaction. It is a mechanism designed to replicate the databases containing the DNS data across a set of DNS servers. Zone transfer comes in two flavors, full (AXFR) and incremental (IXFR). There are numerous tools available to test the ability to perform a DNS zone transfer. Tools commonly used to perform zone transfers are host, dig and nmap.

### SMTP Bounce Back

SMTP bounce back, also called a Non-Delivery Report/Receipt (NDR), a (failed) Delivery Status Notification (DSN) message, a Non-Delivery Notification (NDN) or simply a bounce, is an automated electronic mail message from a mail system informing the sender of another message about a delivery problem. This can be used to assist an attacker in fingerprint the SMTP server as SMTP server information, including software and versions, may be included in a bounce message.

This can be done by simply creating a bogus address within the target's domain. For instance, asDFADSF_garbage_address@target.com could be used to test target.com. Gmail provides full access to the headers, making it an easy choice for testers.

### DNS Discovery

DNS discovery can be performed by looking at the WHOIS records for the domain's authoritative nameserver. Additionally, variations of the main domain name should be checked, and the website should be checked for references to other domains which could be under the target's control.

### Forward/Reverse DNS

Reverse DNS can be used to obtain valid server names in use within an organizational. There is a caveat that it must have a PTR (reverse) DNS record for it to resolve a name from a provided IP address. If it does resolve then the results are returned. This is usually performed by testing the server with various IP addresses to see if it returns any results.

### DNS Bruteforce

After identifying all the information that is associated with the client domain(s), it is now time to begin to query DNS. Since DNS is used to map IP addresses to hostnames, and vice versa we will want to see if it is insecurely configure. We will seek to use DNS to reveal additional information about the client. One of the most serious misconfigurations involving DNS is allowing Internet users to perform a DNS zone transfer. There are several tools that we can use to enumerate DNS to not only check for the ability to perform zone transfers, but to potentially discover additional host names that are not commonly known.

### Web Application Discovery

Identifying weak web applications can be a particularly fruitful activity during a penetration test. Things to look for include OTS applications that have been misconfigured, OTS application which have plugin functionality (plugins often contain more vulnerable code than the base application), and custom applications. Web application fingerprinters such as WAFP can be used here to great effect.

### Virtual Host Detection & Enumeration

Web servers often host multiple "virtual" hosts to consolidate functionality on a single server. If multiple servers point to the same DNS address, they may be hosted on the same server. Tools such as MSN search can be used to map an ip address to a set of virtual hosts.

## Establish External Target List

Once the activities above have been completed, a list of users, emails, domains, applications, hosts and services should be compiled.

### Mapping versions

Version checking is a quick way to identify application information. To some extent, versions of services can be fingerprinted using nmap, and versions of web applications can often be gathered by looking at the source of an arbitrary page.

### Identifying patch levels

To identify the patch level of services internally, consider using software which will interrogate the system for differences between versions. Credentials may be used for this phase of the penetration test, provided the client has acquiesced. Vulnerability scanners are particularly effective at identifying patch levels remotely, without credentials.

### Looking for weak web applications

Identifying weak web applications can be a particularly fruitful activity during a penetration test. Things to look for include OTS applications that have been misconfigured, OTS application which have plugin functionality (plugins often contain more vulnerable code than the base application), and custom applications. Web application fingerprinters such as WAFP can be used here to great effect.

### Identify lockout threshold

Identifying the lockout threshold of an authentication service will allow you to ensure that your bruteforce attacks do not intentionally lock out valid users during your testing. Identify all disparate authentication services in the environment, and test a single, innocuous account for lockout. Often 5 - 10 tries of a valid account is enough to determine if the service will lock users out.

# Internal Footprinting

## Passive Reconnaissance

If the tester has access to the internal network, packet sniffing can provide a great deal of information. Use techniques like those implemented in p0f to identify systems.

## Identify Customer Internal Ranges

When performing internal testing, first enumerate your local subnet, and you can often extrapolate from there to other subnets by modifying the address slightly. Also, a look a the routing table of an internal host can be particularly telling. Below are a number of techniques which can be used.

DHCP servers can be a potential source of not just local information, but also remote IP range and details of important hosts. Most DHCP servers will provide a local IP gateway address as well as the address of DNS and WINS servers. In Windows based networks, DNS servers tend to be Active Directory domain controllers, and thus targets of interest.

### Active Reconnaissance

Internal active reconnaissance should contain all the elements of an external one, and in addition should focus on intranet functionality such as:

- Directory services (Active Directory, Novell, Sun, etc...)
- Intranet sites providing business functionality
- Enterprise applications (ERP, CRM, Accounting, etc...)
- Identification of sensitive network segments (accounting, R&D, marketing, etc...)
- Access mapping to production networks (datacenters)
- VoIP infrastructure
- Authentication provisioning (kerberos, cookie tokens, etc...)
- Proxying and internet access management

## Identify Protection Mechanisms

The following elements should be identified and mapped according to the relevant location/group/persons in scope. This will enable correct application of the vulnerability research and exploitation to be used when performing the actual attack - thus maximizing the efficiency of the attack, and minimizing the detection ratio.

## Network Based Protections

- "Simple" Packet Filters
- Traffic Shaping Devices

- DLP Systems
- Encryption/Tunneling

## Host Based Protections

- Stack/Heap Protections
- Application Whitelisting
- AV/Filtering/Behavioral Analysis
- DLP Systems

## Application Level Protections

- Identify Application Protections
- Encoding Options
- Potential Bypass Avenues
- Whitelisted Pages

## Storage Protections

- HBA - Host Level
- LUN Masking
- Storage Controller
- iSCSI CHAP Secret

## User Protections

- AV/Spam Filtering Software

  > SW Configuration which limit exploitability can be considered antispam / antiAV