Information Gathering

First step of an attacker: learn as much about a particular target (human, system, organization) as possible

- Dependencies and third-party interactions are also important
- Example: Target 2013 breach was achieved through the compromise of a third-party HVAC vendor who had access to the internal network
- Peripheral or “forgotten” systems are often less secure than publicized web, application server, and mail endpoints
- Every piece of information counts!

Many complementary approaches

- Information gathering from public sources
- Passive eavesdropping
- Active scanning and probing
- Induce information leakage (e.g., through social engineering)
- Dumpster diving (e.g., recover data from discarded hard disks)
OSINT (Open-source Intelligence Gathering)

Intelligence collected from publicly available sources
As opposed to covert or clandestine sources

Wide variety of information and sources

Search engines: public documents, forgotten web pages, exposed login interfaces, dashboards, historical data, …
Public data: courthouse documents, tax forms, budgets, …
Media: articles, interviews, blog posts, …
Social media: LinkedIn/Facebook/Twitter/etc., mailing lists, …
Professional/academic sources: reports, presentations, …
Metadata: documents, EXIF, executables, email headers, …
…
Search Engines

Google, Bing, Yandex, Baidu, …
  Refined searches for certain kinds of information ("Google-Fu")
  Useful operators: intext, intitle, inurl, filetype, site

Netcraft: uptime and web server info

Internet Archive’s Wayback Machine: old site versions

Google/Yahoo groups: sysadm questions, gossip, …

LinkedIn: persons within an organization, interests, …

Qualys’ SSL report: SSL configuration of public web servers

Many more: phone directories, “people” search, gov/state databases, dark web search, …
Site report for www.cs.stonybrook.edu

Background

<table>
<thead>
<tr>
<th>Site title</th>
<th>SBU - Computer Science Department - HOME</th>
</tr>
</thead>
<tbody>
<tr>
<td>Date first seen</td>
<td>June 2005</td>
</tr>
<tr>
<td>Site rank</td>
<td>428532</td>
</tr>
<tr>
<td>Description</td>
<td>Not Present</td>
</tr>
<tr>
<td>Keywords</td>
<td>Not Present</td>
</tr>
</tbody>
</table>

Network

<table>
<thead>
<tr>
<th>Site</th>
<th><a href="http://www.cs.stonybrook.edu">http://www.cs.stonybrook.edu</a></th>
</tr>
</thead>
<tbody>
<tr>
<td>Domain</td>
<td>stonybrook.edu</td>
</tr>
<tr>
<td>IP address</td>
<td>130.245.27.2</td>
</tr>
<tr>
<td>IPv6 address</td>
<td>Not Present</td>
</tr>
<tr>
<td>Domain registrar</td>
<td>educause.net</td>
</tr>
<tr>
<td>Organisation</td>
<td>State University of New York/ Stony Brook</td>
</tr>
<tr>
<td>Hosting company</td>
<td>State University of New York/ Stony Brook</td>
</tr>
<tr>
<td>Hosting country</td>
<td>US</td>
</tr>
</tbody>
</table>

Hosting History

| Netblock owner         | State University of New York at Stony Brook |
| IP address             | 130.245.27.2                               |
| OS                     | Linux                                      |
| Web server             | Apache/2.2.22 Ubuntu                      |
| Last seen              | 9-Dec-2015                                 |

Security

Netcraft Risk Rating [FAQ] 0/10
You have connected to a router. Administrative access only. If this device is not in your possession, please contact your local network administrator.
"Password=" inurl:web.config -intext:web.config ext:config

web.config
ftp.mvaonline.com/partners.mvacolombia.com/wwwroot/web.config
... connectionString="Data Source=ns1.nightshade.arxive.com;Initial Catalog=dnn_mva;User ID=cballesteros;Password=********" providerName="System.

Copy of web.config - EarSinus.com
carsinus.com/new/Copy%20of%20web.config
... the provider is specified passwordAttemptThreshold= "int" The number of failed password attempts, or failed password answer attempts that are allowed before ...

Web.config
ftp://60.250.85.148/StreamStore/WG/WebService/Web.config
C:\wra10\FC\FcPumps.xml WaterLevel.xml Data Source=127.0.0.1;Initial Catalog=River;User ID=sa;Password=********

web.config - Axis HR
www.axishrpro.co.uk/wwwroot/web.config
SQLExpress;Database=hrpro;User ID=hrpro;Password=********" />
<connectionStrings>
<appSettings>
<add key="SQLServerConn" value="Server=

D:\IMG_Catalogazione\server=192.168.0.157 ...
D:\IMG_Catalogazione\server=192.168.0.157,Trusted_Connection=false;User ID=sa;Password=********" Initial Catalog=marubi_web_cp; server=192.168.0.157 ...

web.config - PASA
www.pasaweb.com/forum/web.config
... during which failed password attempts and failed password answer attempts are tracked enablePasswordRetrieval="[true/false]" Should the provider support ...
Google Hacking Database (GHDB)

Search the Google Hacking Database or browse GHDB categories

<table>
<thead>
<tr>
<th>Date</th>
<th>Title</th>
<th>Category</th>
</tr>
</thead>
<tbody>
<tr>
<td>2016-03-24</td>
<td>intitle:vood act-index Gateway &gt; Login</td>
<td>Pages containing login portals</td>
</tr>
<tr>
<td>2016-03-24</td>
<td>intext:&quot;powered by webcamXP 5&quot;</td>
<td>Various Online Devices</td>
</tr>
<tr>
<td>2016-03-23</td>
<td>intitle:&quot;VOOD - Welcome to Vood Residential Gateway &gt; Login&quot;</td>
<td>Pages containing login portals</td>
</tr>
<tr>
<td>2016-03-23</td>
<td>intitle:&quot;Residential Gateway Configuration:&quot; intext:&quot;Cable Modem Information.&quot;</td>
<td>Various Online Devices</td>
</tr>
<tr>
<td>2016-03-23</td>
<td>intitle:&quot;Login Page&quot; intext:&quot;Phone Adapter Configuration Utility&quot;</td>
<td>Pages containing login portals</td>
</tr>
<tr>
<td>2016-03-22</td>
<td>(intext:&quot;index of /git&quot;) (&quot;parent directory&quot;)</td>
<td>Sensitive Directories</td>
</tr>
<tr>
<td>2016-03-16</td>
<td>inurl:/sap/bc/webdynpro/sap/ &quot;sap-system-login-oninputprocessing&quot;</td>
<td>Pages containing login portals</td>
</tr>
<tr>
<td>2016-03-14</td>
<td>intext:&quot;sap-system-login&quot;</td>
<td>Pages containing login portals</td>
</tr>
<tr>
<td>2016-03-14</td>
<td>inurl:&quot;sap/hrpcf_a_startpage_ext_cand&quot;</td>
<td>inurl:&quot;sap/hrpcf_a_pw_via_emailExtern&quot;</td>
</tr>
<tr>
<td>2016-03-14</td>
<td>intitle:&quot;Logon - SAP Web Application Server&quot;</td>
<td>Pages containing login portals</td>
</tr>
</tbody>
</table>
Non-technical Information

Information about persons, operations, behaviors, is useful for targeted attacks

- Spear phishing: messages that appear to come from trusted sources
- Watering hole attacks: target the members of a group by infecting websites they are known to visit

Social networks, corporate websites, partners/third-parties, mailing lists, impersonation, social engineering, …

Public actions may also reveal actionable information

- Example: a system administrator of a particular company asks on ServerFault how to secure Nginx
What is this?
--------

theHarvester is a tool for gathering e-mail accounts, subdomain names, virtual hosts, open ports/banners, and employee names from different public sources (search engines, pgp key servers).

Is a really simple tool, but very effective for the early stages of a penetration test or just to know the visibility of your company in the Internet.
WHOIS

Protocol for querying databases with registration information about assignees of internet resources

- IP address blocks, domain names, and autonomous systems
- Top registries: AFRINIC, APNIC, ARIN, IANA, ICANN, LACNIC, NRO, RIPE, InterNic

whois command-line utility

# whois stonybrook.edu
# whois 130.245.27.2

Also web interfaces of registrars and third-party services

Useful information

- Registrar information, domain creation/expiration dates, primary DNS name servers associated with the domain
- Registrant information such as First Name, Last Name, Organization, physical address, phone number, and e-mail address
- Assigned domain administrator, billing contact, technical contact
DNS

Valuable information about individual hosts
  IP addresses (A, AAAA) of certain domains
  Name (NS) and mail (MX) servers of a domain
  Name aliases (CNAME) and reverse mappings (PTR)

Other useful records
  SRV: generic service locator (protocol, host, port) for domain services
       (e.g., Kerberos, LDAP, SIP, XMPP)
  TXT: SPF, DKIM, DMARC, and other custom information
  HINFO: CPU, OS, and other host-related information

Various utilities: nslookup, dig, host

Zone transfers (AXFR) provide all entries for a domain
  Used mostly for replication across secondary DNS servers
  Wealth of information, often very sensitive: subdomains, internal IP addresses and hosts, services used, …
DNS Brute Forcing

Zone transfers are usually restricted only among authorized servers

   Although misconfigurations are common…
   Alternative: guess valid DNS records

Dictionary attack using A/AAAA record requests

   Query based on list of commonly used subdomains, hostnames, words, etc. (e.g., www, mail, vpn, webaccess, msexchange)

DNSSEC NSEC and NSEC3 zone walking

   The NSEC record is used to give negative answers to queries, but has the side effect of allowing enumeration of all names
   NSEC3 mitigates this, but still allows for dictionary attacks
Network Scanning

Identify accessible hosts, running services, service and OS versions, ...

Active: target network can observe probe requests
As opposed to passive reconnaissance or querying of public sources
Stealthiness matters! IDSes can easily detect noisy scans

Two main dimensions

**Horizontal scanning:** scan a subnet (or the whole internet) on a particular port number
E.g., find all hosts running a vulnerable service (internet worms)

**Vertical scanning:** scan all (or a subset of) ports on a given host
Scan common ports first

Manual scanning using ping and netcat can be used for quick assessments
Affected System: Printing
Reported On: Sunday, March 27, 2016 - 11:30pm
Incident Started: Saturday, March 26, 2016 - 11:00am
Estimated Restoration: Unknown

- 03/27/2016 - 11:30pm - [Investigating Incident]

The Division of Information Technology (DoIT) is investigating the anonymous delivery of an anti-Semitic message to a network printer at Stony Brook University. Other universities are reporting similar incidents.

Upon being made aware of the situation, DoIT applied necessary filters to incoming Internet traffic in order to block these types of messages. Initial reports show this to be successful, however, it is important to note that the filters applied may interfere with printing to West Campus printers from off-campus, and from Computer Science, CEWIT, the Hospital, and the Residence Halls. Problems like this often arise when printers are not secured. While DoIT’s Client Support team has been working with departments over the past several months to better secure printers on our network, there is more work to be done. Information about printer security will be posted on the DoIT website this week. Should you have any trouble printing, please open a case in the Client Support Help Portal (http://service.stonybrook.edu) or contact Client Support for assistance.
Nmap

De facto tool for network scanning

Support for many port scan types

- `-sS` TCP SYN scan: just wait for the ACK
- `-sT` TCP connect scan: full connection (useful for non-root)
- `-sU` UDP scan: protocol-specific payload for known ports
- `-sA` ACK scan: determine if a firewall is stateful
- `-sO` IP protocol scan: determine IP protocols (TCP, ICMP, IGMP) used
- `-p` Specify port range (default: 1000 most common ports)

Beyond simple port scanning: extensible framework with support for third-party scripts

auth, broadcast, brute, default, discovery, dos, exploit, external, fuzzer, intrusive, malware, safe, version, vuln
Service Fingerprinting

After identifying that a port is open, try to gather more information about the service

```
# nmap -sV 192.168.0.1 -p 22
```

Complete the connection and attempt to determine the software type and version

Version detection “interrogates” those ports to determine more about what is actually running

Server-initiated dialog: banner grabbing

Upon connection, the server transmits a banner string that often includes version information (e.g., SSH)

Client-initiated dialog: send probe application requests

Nmap has about 6,500 dialogue patterns for more than 650 protocols such as SMTP, FTP, HTTP, etc.
Idle Scan

Hide scan attempts by blaming another “zombie” host
Zombie must be mostly idle (e.g., network printer)
Zombie should have sequential/predictable IPID behavior

Probing the zombie’s IPID

Spoofing a SYN from the zombie

Probing the zombie’s IPID again

(1) SYN/ACK
SYN_ack=0
IPID=X

(2) SYN/ACK
SYN_ack=0
IPID=X+1

(3) SYN/ACK
SYN_ack=0
IPID=X+2
**ARP Scan**

Useful technique for host enumeration in a LAN

Find every active IPv4 device in the same subnet

Send a “who has” broadcast packet for each IP address of interest

Example: try all 254 host IP addresses for a /24 subnet

Retry a couple of times if no response is received

**Linux command-line tool: arp-scan**

```bash
# arp-scan 192.168.0.0/24
```
Fast Internet-wide Scanning

http://zmap.io

Scan the entire IPv4 address space for a given port in ~45 minutes using a single machine and a gigabit link

Speed gains

- Eliminate per-connection state by overloading packet header fields (src port, initial Seq No.) – similar concept to SYN cookies
- Bypass TCP stack: raw socket for packet transmission, libpcap to receive responses
- Send as many probes as NIC can support
- Don’t wait for timeouts – just send a fixed number of probes (usually one is enough to achieve decent coverage)
Shodan: let others do the scanning for you
Opportunistic Discovery

Use case: IPv6 address harvesting by joining pool.ntp.org

Non-published (but publicly accessible) random IPv6 addresses suddenly started getting scanned

How were they discovered?
Random guessing is ruled out: 128-bit wide addresses...

Hosts were Linux devices running an NTP daemon for time synchronization

Periodic queries to pool.ntp.org (default configuration)

Observation: IPv6 clients using brand new addresses to connect to pool.ntp.org are subsequently scanned

Probes originated from *.scan6.shodan.io hosts
The NTP servers involved were later removed from the pool
Vulnerability Scanning

Identify vulnerabilities in exposed services
  Typical next step after network scanning
  Exploitable bugs, misconfigurations, default passwords, …

OpenVAS (open-source), Nessus (free/commercial, proprietary), Qualys (commercial), Nexpose (commercial), …

New “vulnerability tests” released every day
  45,000 in total for OpenVAS as of Feb. 2016

Usually come with user-friendly GUI for configuration, policy management, and report generation