Reconnaissance

The adversary is trying to gather information they can use to plan future operations. Reconnaissance consists of techniques that involve adversaries actively or passively gathering information that can be used to support targeting. Such information may include details of the victim organization, infrastructure, or staff/personnel.

- **Active Scanning (T1595)**: Adversaries may execute active reconnaissance scans to gather information that can be used during targeting. Active scans are those where the adversary probes victim infrastructure via network traffic. Adversaries may perform different forms of active scanning depending on what information they seek to gather. These scans can also be performed in various ways, including using native features of network protocols such as ICMP.
    - **Scanning IP Blocks (T1595.001)**: Public IP addresses may be allocated to organizations by block, or a range of sequential addresses and these can be scanned to gather information. Adversaries may scan IP blocks to Gather Victim Network Information, such as which IP addresses are actively in use as well as more detailed information about hosts assigned these addresses. Scans may range from simple pings (ICMP requests and responses) to more nuanced scans that may reveal host software/versions via server banners or other network artifacts.
    - **Vulnerability Scanning (T1595.002)**: Vulnerability scans typically check if the configuration of a target host/application (ex: software and version) potentially aligns with the target of a specific exploit the adversary may seek to use. Vulnerability scans typically harvest running software and version numbers via server banners, listening ports, or other network artifacts.
    - **Wordlist Scanning (T1595.003)**: Involves crawling using wordlists that contains generic, commonly used names and file extensions or terms specific to a particular software. The goal is the identification of content and infrastructure. For example, adversaries may use web content discovery tools and generic or custom wordlists to enumerate a website's pages and directories. This can help them to discover old, vulnerable pages or hidden administrative portals that could become the target of further operations.
- **Gather Victim Host Information (T1592):** Adversaries may gather information about the victim's hosts that can be used during targeting.
    - **Hardware (T1592.001):** Information about hardware infrastructure may include card/biometric readers, dedicated encryption hardware, etc.
    - **Software (T1592.002):** Information about installed software may include antivirus, SIEMs, etc.
    - **Firmware (T1592.003):** Information about host firmware may include configuration, purpose, age/patch level, etc.
    - **Client Configurations (T1592.004):** Information about client configurations may include a variety of details and settings, including operating system/version, virtualization, architecture (ex: 32 or 64 bit), language, and/or time zone.
- **Gather Victim Identity Information (T1589):** Adversaries may gather information about the victim's identity that can be used during targeting. Adversaries may gather this information in various ways, such as direct elicitation via Phishing for Information.
    - **Credentials (T1589.001):** Account credentials gathered by adversaries may be those directly associated with the target victim organization or attempt to take advantage of the tendency for users to use the same passwords across personal and business accounts. Adversaries may also compromise sites then add malicious content designed to collect website authentication cookies from visitors. Credential information may also be exposed to adversaries via leaks to online or other accessible data sets (ex: Search Engines, breach dumps, code repositories, etc.). Adversaries may also purchase credentials from dark web or other black-markets.

    - **Email Addresses (T1589.002):** Even if internal instances exist, organizations may have public-facing email infrastructure and addresses for employees. Adversaries may easily gather email addresses, since they may be readily available and exposed via online or other accessible data sets. Email addresses could also be enumerated via more active means, such as probing and analyzing responses from authentication services that may reveal valid usernames in a system.
    - **Employee Names (T1589.003):** Employee names can be used to derive email addresses as well as to help guide other reconnaissance efforts and/or craft more-believable lures.
- **Gather Victim Network Information (T1590):** Information about networks may include a variety of details, including administrative data (ex: IP ranges, domain names, etc.) as well as specifics regarding its topology and operations.
    - **Domain Properties (T1590.001):** includes what domain(s) the victim owns as well as administrative data (ex: name, registrar, etc.) and more directly actionable information such as contacts (email addresses and phone numbers), business addresses, and name servers.
    - **DNS (T1590.002):** includes registered name servers as well as records that outline addressing for a target's subdomains, mail servers, and other hosts. DNS, MX, TXT, and SPF records may also reveal the use of third-party cloud and SaaS providers.
    - **Network Trust Dependencies (T1590.003):** includes second or third-party organizations/domains (ex: managed service providers, contractors, etc.) that have connected (and potentially elevated) network access.
    - **Network Topology (T1590.004):** includes the physical and/or logical arrangement of both external-facing and internal network environments. This information may also include specifics regarding network devices (gateways, routers, etc.) and other infrastructure.
    - **IP Addresses (T1590.005):** includes which IP addresses are in use. IP addresses may also enable an adversary to derive other details about a victim, such as organizational size, physical location(s), Internet service provider, and or where/how their publicly facing infrastructure is hosted.
    - **Network Security Appliances (T1590.006):** includes the specifics of deployed firewalls, content filters, network-based intrusion detection systems (NIDS) and proxies/bastion hosts.
- **Gather Victim Organization Information (T1591):** Information about an organization may include a variety of details, including the names of divisions/departments, specifics of business operations, as well as the roles and responsibilities of key employees.
    - **Determine Physical Locations (T1591.001):** includes where key resources and infrastructure are housed. Physical locations may also indicate what legal jurisdiction and/or authorities the victim operates within.
    - **Business Relationships (T1591.002):** includes second or third-party organizations/domains (ex: managed service providers, contractors, etc.) that have connected (and potentially elevated) network access. This information may also reveal supply chains and shipment paths for the victim's hardware and software resources.
    - **Identify Business Tempo (T1591.003):** includes operational hours/days of the week. This information may also reveal times/dates of purchases and shipments of the victim's hardware and software resources.
    - **Identify Roles (T1591.004):** Information about business roles may reveal a variety of targetable details, including identifiable information for key personnel as well as what data/resources they have access to.
- **Phishing for Information (T1598):** Phishing for information is an attempt to trick targets into divulging information, frequently credentials or other actionable information. All forms of phishing are electronically delivered social engineering. More generally, adversaries can conduct non-targeted phishing, such as in mass credential harvesting campaigns.
    - **Spearphishing Service (T1598.001):** In this scenario, adversaries send messages through various social media services, personal webmail, and other non-enterprise-controlled services. These services are more likely to have a less-strict security policy than an enterprise.
    - **Spearphishing Attachment (T1598.002):** adversaries attach a file to the spearphishing email and usually rely upon the recipient populating information then returning the file. The text of the spearphishing email

usually tries to give a plausible reason why the file should be filled-in, such as a request for information from a business associate.
    - **Spearphishing Link (T1598.003):** the malicious emails contain links generally accompanied by social engineering text to coax the user to actively click or copy and paste a URL into a browser. The given website may be a clone of a legitimate site (such as an online or corporate login portal) or may closely resemble a legitimate site in appearance and have a URL containing elements from the real site. URLs may also be obfuscated by taking advantage of quirks in the URL schema, such as the acceptance of integer- or hexadecimal-based hostname formats.
    - **Spearphishing Voice (T1598.004):** adversaries use phone calls to elicit sensitive information from victims. Known as voice phishing (or "vishing"), these communications can be manually executed by adversaries, hired call centers, or even automated via robocalls. Voice phishers may spoof their phone number while also posing as a trusted entity, such as a business partner or technical support staff.
- **Search Closed Sources (T1597):** Adversaries may search in different closed databases depending on what information they seek to gather. Information from these sources may reveal opportunities for other forms of reconnaissance.
    - **Threat Intel Vendors (T1597.001):** Threat intelligence vendors may offer paid feeds or portals that offer more data than what is publicly reported. Although sensitive details (such as customer names and other identifiers) may be redacted, this information may contain trends regarding breaches such as target industries, attribution claims, and successful TTPs/countermeasures. Threat actors may seek information/indicators gathered about their own campaigns, as well as those conducted by other adversaries that may align with their target industries, capabilities/objectives, or other operational concerns.
    - **Purchase Technical Data (T1597.002):** Information about victims may be available for purchase within reputable private sources and databases, such as paid subscriptions to feeds of scan databases or other data aggregation services. Adversaries may also purchase information from less-reputable sources such as dark web or cybercrime black markets. Details from purchased data include employee contact information, credentials, or specifics regarding a victim's infrastructure.
- **Search Open Technical Databases (T1596):** Information about victims may be available in online databases and repositories, such as registrations of domains/certificates as well as public collections of network data/artifacts gathered from traffic and/or scans.
    - **DNS/Passive DNS (T1596.001):** DNS information may include a variety of details, including registered name servers as well as records that outline addressing for a target's subdomains, mail servers, and other hosts. Threat actors can query nameservers for a target organization directly, or search through centralized repositories of logged DNS query responses (known as passive DNS). Adversaries may also seek and target DNS misconfigurations/leaks that reveal information about internal networks.
    - **WHOIS (T1596.002):** WHOIS data is stored by regional Internet registries (RIR) responsible for allocating and assigning Internet resources such as domain names. Anyone can query WHOIS servers for information about a registered domain, such as assigned IP blocks, contact information, and DNS nameservers.
    - **Digital Certificates (T1596.003):** Digital certificates are issued by a certificate authority (CA) in order to cryptographically verify the origin of signed content. These certificates, such as those used for encrypted web traffic (HTTPS SSL/TLS communications), contain information about the registered organization such as name and location. Digital certificate data may also be available from artifacts signed by the organization (ex: certificates used from encrypted web traffic are served with content).
    - **CDNs (T1596.004):** CDNs allow an organization to host content from a distributed, load balanced array of servers. Adversaries may seek and target CDN misconfigurations that leak sensitive information not intended to be hosted and/or do not have the same protection mechanisms (ex: login portals) as the content hosted on the organization's website.
    - **Scan Databases (T1596.005):** Various online services continuously publish the results of Internet scans/surveys, often harvesting information such as active IP addresses, hostnames, open ports,

certificates, and even server banners. Threat actors can use online resources and lookup tools to harvest information from these services. Adversaries may seek information about their already identified targets or use these datasets to discover opportunities for successful breaches.

- **Search Open Websites/Domains (T1593):** Information about victims may be available in various online sites, such as social media, new sites, or those hosting information about business operations such as hiring or requested/rewarded contracts.
    - **Social Media (T1593.001):** Social media sites may contain various information about a victim organization, such as business announcements as well as information about the roles, locations, and interests of staff. Threat actors may passively harvest data from these sites, as well as use information gathered to create fake profiles/groups to elicit victim's into revealing specific information.
    - **Search Engines (T1593.002):** Search engine services typical crawl online sites to index context and may provide users with specialized syntax to search for specific keywords or specific types of content (i.e. filetypes). Threat actors may use search engines to harvest general information about victims, as well as use specialized queries to look for spillages/leaks of sensitive information such as network details or credentials.
    - **Code Repositories (T1593.003):** Victims may store code in repositories on various third-party websites such as GitHub, GitLab, SourceForge, and BitBucket. Users typically interact with code repositories through a web application or command-line utilities such as git. Public code repositories can often be a source of various general information about victims, such as commonly used programming languages and libraries as well as the names of employees. Adversaries may also identify more sensitive data, including accidentally leaked credentials or API keys.
- **Search Victim-Owned Websites (T1594):** Victim-owned websites may contain a variety of details, including names of departments/divisions, physical locations, and data about key employees such as names, roles, and contact information. These sites may also have details highlighting business operations and relationships.