Threat Intelligence

By Dr Ashu Sharma

Understanding Threat Intelligence

- Basic terminology
- Threat: An IT entity, such as a host or website, that is suspected of performing attacks.
- Attack: An instance of malicious activity. An example of an attack is malware directed at a target.
- Attacker: A person or group who attacks others. (Another term for attacker is threat agent.)
- Threat indicator: One or more related data points that imply heightened risk: for example, an IP address located in a specific country.
- Threat intelligence: Insights into threats gained by collecting evidence, such as observing attacks and studying the characteristics of attackers. TI comprises threat indicators and associated threat indicator metadata

Threat indicator types

- IP address
- URL
- malicious content

Threat indicator metadata

- Time stamp One of the simplest, yet most critical, pieces of metadata for a threat indicator is a time stamp, which indicates when the TI related to the threat indicator was collected.
- Score Many TI feeds offer risk scores, which are measures of the relative maliciousness of a given IP address, URL, host, or executable.
- Source Some sources of TI only aggregate existing information — that is, they take free threat feeds from other parties and aggregate them into a single feed of their own. Other sources of TI do their own intelligence gathering, organically collecting threat information from customer networks or out in the wild, by monitoring and analyzing global network traffic, especially traffic on the darknets.
- Geolocation Most sources of TI provide geolocation information and use that information as a major factor in risk scoring. If many attacks are coming from a particular country or region, other activity from the same place is somewhat more likely to be malicious as well.

Threat indicator metadata

- Category TI products and services categorize threats in many ways. The purpose of these categories is to provide insight into the nature of a threat. Here are some sample high-level threat categories:
    - Anonymous proxy: This host has been identified as a node on an anonymous proxy network, potentially indicating attackers' intent to hide their points of origin
    - Bogon: The host's IP address hasn't been allocated by the Internet Assigned Numbers Authority (IANA) or assigned by an authorized registry such as the American Registry for Internet Numbers (ARIN), which means that it shouldn't be used on the Internet.
    - Bot: The host appears to be infected by malware and is under the control of attackers as part of a botnet.
    - Botnet: The host appears to be a command-and-control node for a botnet, giving orders to bots — directing them to attack other hosts, for example.
    - Malware: The host is known to serve malware. An example is a web server with a URL that points to malware hosted on the web server itself.
    - Passive DNS: The host's IP address appears in passive Domain Name System (DNS) records. Passive DNS records often capture information on attacker domains, such as those used for phishing

Why TI Matters

- Improved attack prevention, detection, and response.
    - controls for attack prevention, detection, and response — firewalls, IPSs, and SIEMs — struggle to identify many attacks because these attacks are highly customized. Signature-based detection alone is no longer effective. You need to adopt a whole new way of identifying attacks, and that way is TI.
    - TI doesn’t replace your existing security controls; it enhances them.

- Expedited forensic investigations
    - After an incident occurs, the forensic investigation can take days or weeks, with a good chunk of this time being spent trying to determine who attacked the organization. By using TI services, you can quickly find out the threat history of an IP address or a URL and its threat categorizations. This information expedites investigations, such as by revealing that an attacking host is actually an innocent bystander being controlled by a botnet.

Why TI Matters

- Inputs for risk assessments
    - One of the toughest parts of performing an enterprise risk assessment is accurately portraying the nature of the threats against the organization. Fortunately, you can use TI services in conjunction with your security logs to categorize threats to your organization and their relative frequency. This data lets you see trends over time and helps you ensure that your risk assessment process takes the entire threat landscape into account.

Understanding TI Delivery

- TI reports
  - human-readable TI report, created by human intelligence analysts who research a particular threat, then produce a white paper or booklike report on that threat.
- Machine-readable TI (MRTI)
  - Machine-readable threat intelligence (MRTI) is formatted for automated use. An MRTI feed can become an input into a SIEM, firewall, IPS, or other security control to give it real-time information about threat indicators. This data allows the security control to make better-informed decisions about alerting on, blocking, or otherwise responding to a possible threat
- Console-based TI
  - is essentially MRTI in a human-friendly console interface.
- TI appliances
  - Developed by a TI vendor, a TI appliance typically requires a persistent connection to the vendor’s data feeds. Then the appliance uses this MRTI out of band (to detect likely malicious activity) or inline (to detect and block such activity).

Gathering Threat Intelligence

- Existing data feeds
- Internal customer networks
  - Detecting suspicious activity entering, transiting, or exiting the customer's network
  - Sharing this information (sanitized, of course) with other customers to alert them to threats that they may see in the near future

  Unfortunately, this approach has some significant drawbacks:

  - The vendor can see a customer's sensitive security information, such as which vulnerabilities an attacker has exploited and which of its hosts are compromised.
  - The vendor is responsible for sanitizing all the information it collects.

Gathering Threat Intelligence

- External networks (not paid customers)
  - External network monitoring allows a TI vendor to find attack traffic in many places on the Internet, not just on customer networks. This technique is more proactive than waiting until customers get attacked, and it provides a much more comprehensive picture of current threats, identifying thousands or even millions of attacking hosts around the world.

Automated Sources of TI

- Anonymous proxies
  - an anonymous proxy is a node on an anonymous proxy network. Such a network is designed to anonymize users' computing activity so that their actions can't be traced to their origin. For this reason, many criminals use anonymous proxy networks to conceal their identities when they use the Internet to commit crimes
- Crawlers
  - A crawler is a program that methodically scans web servers and other hosts to identify the content that they serve to others. Crawling every web server in the world would be much too resource-intensive for TI vendors — comparable in scope to what major Internet search companies such as Google might do.
- Free services
  - Some TI vendors provide free services to the Internet community as a whole, such as free email and free Domain Name System (DNS) hosting. The latter service is particularly popular with attackers, who often need to acquire and stage a domain name in a hurry so that they can launch a phishing attack and then tear down the site in a matter of hours.

Automated Sources of TI

- Geolocation
  - Geolocation information serves a few important purposes:
  - It can identify hosts in countries that don’t need to interact with the organization. Some organizations that have a limited clientele and particularly high security needs choose to block all activity involving certain geographic locations, but this posture tends to be rather extreme.
  - TI customers commonly factor geolocation information into risk scores. Suppose that most of the attacks against an organization come from a particular city. When new suspicious activity involving that city is observed, it may make sense to give those threat indicators higher scores than usual because of their association with the city.
  - Geolocation information can be used to correlate threats. Initially, threats may appear to come from the same geographic region, but a more detailed analysis may indicate that they’re coming from the same building, suggesting a single attacker.
- Honeypots
  - security researchers and security-minded organizations have used honeypots — hosts whose only purpose is to lure attackers and record their activities for analysis — to detect malicious activity. An attacker (or malware acting on behalf of the attacker) scans a honeypot, thinks that it’s a legitimate host, and attacks it. Although the honeypot may be configured to appear to contain valuable resources, it doesn’t; it’s simply a trap designed to record malicious activity.

Automated Sources of TI

- IANA and Internet registries
  - The Internet Assigned Numbers Authority (IANA) oversees IP addresses globally. IANA controls the high-level allocation of IPv4 and IPv6 addresses by delegating the allocation of chunks of address space to authorized Internet registries, each of which then allocates IP addresses in a different region of the world. Only addresses that have been allocated by IANA to a registry and then allocated by that registry to an organization (such as an Internet service provider, company, or educational institution) may be used on the Internet. The use of other addresses on the Internet, such as reserved private addresses (such as the 10 net), indicates spoofed traffic, misconfigured network devices, a darknet, or other suspicious activity.
- Internet Relay Chat (IRC)
  - Another common source of TI is the Internet Relay Chat (IRC) protocol. IRC, which has been around for many years, provides a text-based chat mechanism that works over the Internet in a standardized way. People around the world use thousands of IRC servers to communicate with one another. Unfortunately, attackers take advantage of the communications infrastructure provided by IRC servers. Many botnets leverage IRC channels to communicate with their bots, for example. By monitoring these channels, TI vendors can identify hosts that are bots or botnet command-and-control nodes, which give orders to and collect information from bots.

Automated Sources of TI

- P2P
  - Peer-to-peer (P2P) networks are used for many reasons, both benign and malicious. The basic purpose of P2P is to allow people to share files. Unfortunately, many people use P2P to share illegal content, such as pirated music and software. P2P networks are also used to share data stolen during attacks on organizations. TI vendors can monitor P2P networks to determine which hosts are involved in providing or acquiring illegal content.

Calculating Scores

- No magic formula exists for calculating risk scores. Each TI vendor has its own way of analyzing TI and determining the relative priority of each factor. Think of the factors as variables in a giant formula. Dozens or hundreds of variables could be used in a single formula, with high-level categories such as these:
  - Geography and geolocation
  - Routing changes and IP registration validity
  - Domain Name Service (DNS) reverse lookup for IP address
  - Frequency of searches for a threat indicator
  - Threat category (bot, anonymous proxy, and so on)

Aging Score

- Score aging is the process of reassessing a score after a given period of time.
  - Suppose that you see clearly malicious activity from a particular host, so you assign that host a score of 100 on a 0-to-100 scale.
  - Now imagine that it’s a week later, and you’ve seen no subsequent malicious activity from this host.
  - Are you still completely confident that the host is malicious? Maybe the host is a victim that has been cleaned up, or maybe the IP address has been reassigned.
  - To reflect this gradual loss of confidence in the data, it’s absolutely critical that risk scores be aged. Recent data, including lack of observations of new malicious activity, should be given greater weight than older data.

Using Scores

- A threshold is the lowest score that matters to your organization. Every organization has a different threshold that makes sense based on its risk tolerances and the combination of threats it faces. TI vendors often suggest a threshold, but organizations should be prepared to alter it
- Setting a threshold involves determining what level of false positives and false negatives is acceptable — how many benign transactions can be blocked and how many malicious transactions can be permitted.

TI to Support Incident Response

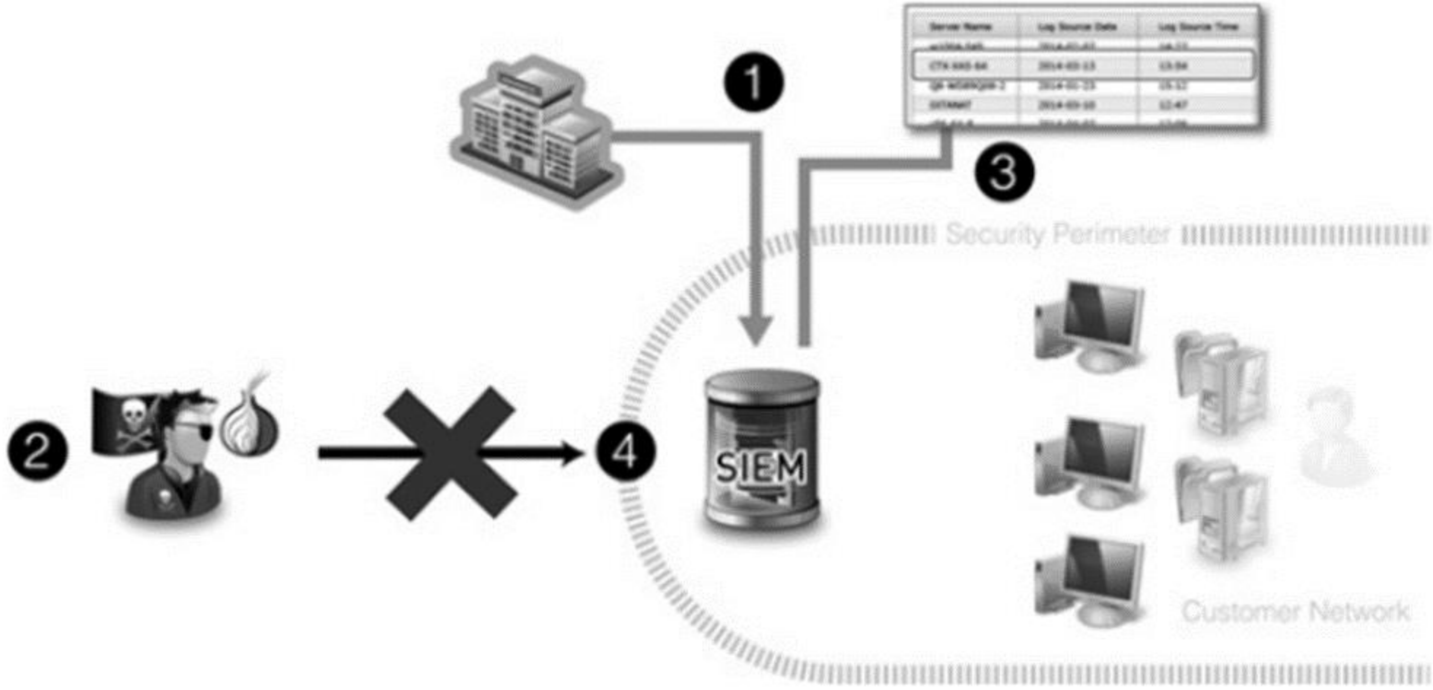
- Incident response is a critical capability for organizations, because security incidents can cause untold damage in a short time.
- TI is invaluable for incident response. The forms of TI most often used to support incident response efforts are machine- readable TI (MRTI), console-based TI, and TI reports.

Improving incident detection

- TI can improve the detection of incidents in multiple ways:
- TI can speed incident detection by indicating which suspicious activity is linked to known malicious hosts. An example is having MRTI ingested by your security information and event management (SIEM) system so that when analysts review suspicious log entries, they have context provided by the MRTI.
- TI can find compromised systems within your own organization, such as malware-infected hosts that are acting as bots. You can identify these hosts by regularly checking MRTI for your organization’s external IP addresses.
- MRTI can provide real-time information about attacks on other organizations. This information allows you to block the hosts at the source of these attacks from accessing your networks, thereby preventing breaches before they occur.

Using MRTI

1. MRTI is transferred regularly from a TI vendor to your SIEM solution.
2. An attacker launches attacks against your organization, and the SIEM records this activity.
3. The SIEM does its standard analysis of events and correlates suspicious activity with the attacker’s IP address, which was included in the latest MRTI update.
4. The SIEM acts appropriately to stop the activity, such as alerting a human to intervene or directing the firewall to block the connection.


1
Server Name
Log Source Date
Log Source Time
3
Security Perimeter
2
4
SIEM
Customer Network