Integrate Threat Reputation Services

Overwhelming volume

Today's adversaries use the complex and dynamic nature of cyberspace to conduct a large volume of attacks using varied vectors. Security devices deployed on enterprise networks help protect against malicious activity, but in order to keep systems defended, security pros must relentlessly research the latest threats to keep up with the increasingly large array of malicious activity. Threat Reputation Services (TRS) gather crowdsourced intelligence from sensors worldwide, so that the slightest misstep by an adversary immediately alerts an entire community. Integrating TRS with existing network security devices and applications contributes to a better defended network while freeing valuable resources.

What are Threat Reputation Services?

Fundamentally, TRS are automated interfaces that provide client systems with a machine-readable score that describes the perceived trustworthiness of an object, based on an identifier/indicator. Vendors conduct frequent threat assessments against websites, files, domain names, and other objects to categorize how often they have been associated with malicious activity, based on observed past behavior and shared intelligence. Reputation scores are assigned to the objects to reflect these levels of observed activity. Objects associated with malicious activity are assigned worse scores while known good objects are assigned better scores. The resultant threat level/reputation scores can be used by security devices to detect or block potentially malicious activity. TRS typically use indicators like Uniform Resource Locators (URL), Internet Protocol (IP) addresses, file hashes (like SHA256 or MD5), Domain Name System (DNS) names, and email addresses.

What are the threats?

Attackers often use email as a threat vector, employing phishing to take advantage of users. Information in emails can be crafted to trick users into thinking the emails are legitimate. Malicious emails can contain attachments that download malware, include links to infected websites, and request recipients to respond with sensitive information. Users that fall victim can open the door for attackers to breach not only the computer the user is using, but the rest of the enterprise network too. Since email senders are associated with their email domains and IP addresses, TRS can use email domain authentication with Domain-based Message Authentication, Reporting, and Conformance (DMARC) to prevent spoofing domains, and then lookup the domain and IP address reputations to block malicious emails, protecting networks and users from known bad senders [1].

Websites can contain links to malware downloads, links that redirect to other malicious websites, and login fields that steal usernames and passwords. Drive-by downloads can also occur, where malware is secretly downloaded onto workstations with outdated and vulnerable web browsers the moment a user arrives on malicious websites. TRS addresses website threats by assigning reputation scores to known bad URLs associated with the websites.[1] TRS-enabled network devices and browsers block traffic containing URLs with bad reputation scores.

DNS utilizes a hierarchy of domains for translating website URLs into their associated IP addresses. When a user attempts to access a website, a DNS request for the relevant domain is performed. TRS can block DNS queries to known malicious domains. [2]

Unauthorized scanning can be used to conduct reconnaissance and gain information about networks, such as open ports and operating systems (OS). This information allows adversaries to tailor attacks, mapping them to known vulnerabilities and threats associated with the open ports and OSs. The techniques used to conduct the scanning, whether successful or

[1] Access to malicious websites often requires DNS queries, but not always.

not, could also overwhelm affected network resources, potentially resulting in denial of service. If IP addresses have known bad reputation scores, a TRS can support automated blocking.

Like URLs, IP addresses are also associated with websites. Traffic from malicious websites can also be blocked according to their low-scored IP addresses. Additionally, IP-based TRS can block command and control (C2) and exfiltration traffic.

Malware and malicious files can affect network hosts through a myriad of vectors, including when victims browse infected websites or open malicious emails. Malware spreading from host to host can disrupt the entire network from the inside. In order to keep up with the volume of new malware variants, most anti-virus (AV) solutions support cloud-based lookups to TRS for detection, where they can draw upon a much larger corpus of malware knowledge. Such files can be detected after they are downloaded or dropped onto systems, but before they execute.

The term “zero-day” can also be used to describe malware artifacts that have not yet been identified as malicious and do not have detection indicators developed. While security devices can fail to detect and prevent file artifacts, if C2 infrastructure is reused, then IP- or DNS-based TRS can still be effective. In essence, the real-time updates of TRS force an adversary to simultaneously move away from or “burn” all previously used infrastructure and recompile all malware in order to evade detection. Another way some enterprises with strict policies can use TRS to address zero-day malware is by blocking items with “unknown” reputations, effectively stopping threats that are not recognized as explicitly malicious. However, this approach presents a concern due to its potential to block legitimate items that also have “unknown” reputations and must be carefully applied.

The Pyramid of Pain

Some indicators, when used to detect and identify adversaries, cause the adversary more “pain” by forcing them to create new attack infrastructure or change their tools. This observation inspired the development of David Bianco’s “Pyramid of Pain” concept which is helpful in understanding the value of TRS. [3]

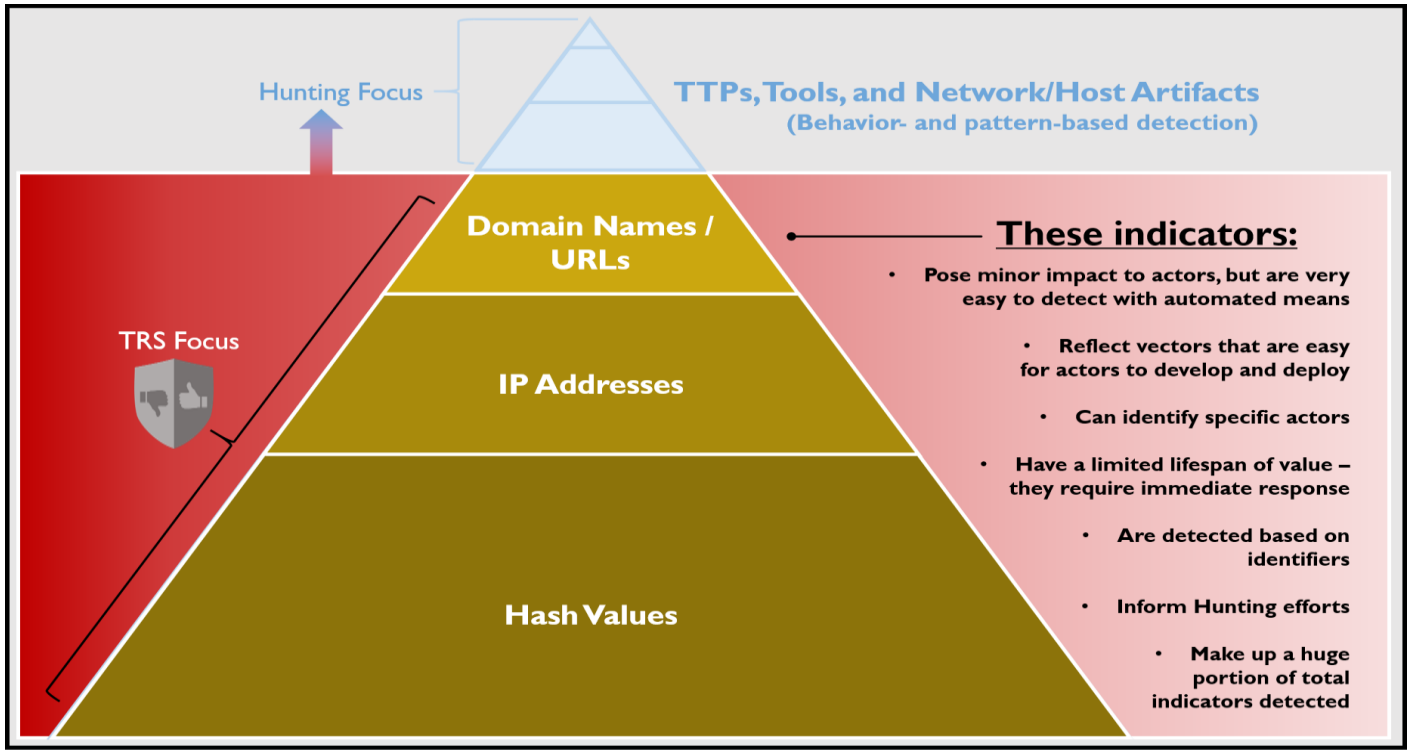

Figure 1: The Pyramid of Pain — Identifier-based Indicators

At the top of the pyramid are **behavior- and pattern-based indicators**, including the adversary's tactics, techniques, and procedures (TTPs). These are the fundamental tools and behaviors integral to an attacker's operation. If defenders can identify these TTPs and use them to detect and block an attack, it forces an attacker to rethink their entire approach, sending them back to the drawing board. But these indicators are not just difficult for attackers to develop: they are also difficult for defenders to detect.

Moving down the pyramid are **identifier-based indicators.** While they are excellent at pinpointing specific actors and activity, these indicators have a narrow lifetime-of-value measured in mere hours. It is easy for adversaries to launch attacks involving changes to IPs, URLs, and file hashes. The only way to keep up with these easy, frequent changes and evasions is complete automation and industrial-scale indicator sharing. TRS focuses on the detection of identifier-based indicators by drawing on shared intelligence, and integrates with network security devices to identify and block threats in real time. This automation frees valuable security resources and informs the threat hunting process.[2]

As defenders mature in their approach to threat hunting, they use increasingly advanced detection like correlation of logs, statistical and visual analysis, machine learning, behavioral analysis, and creation of new and innovative detection capabilities. Defenders require the time and resources needed to develop skills for threat hunting and to bring their threat hunting model to maturity. TRS can be used to detect simple indicators while defenders focus their resources to continuously hunt threats. Combining automated detection with pattern-matching/signature detection and behavioral detection allows network defenders to counter clever adversaries who use new infrastructure and more sophisticated tools.

Implementation

Vendors offer TRS as an additional feature that can be incorporated into endpoint security software and traffic inspection network devices such as firewalls and proxies. This allows a network owner to set a simple policy for the categories and reputation scores allowed, alerted, or blocked. Licenses and subscriptions to the TRS threat databases need to be acquired before TRS filtering can be enabled on a device. Further, licenses and subscriptions for each of the filtering services desired must be acquired, and some vendors split services among item types (e.g., URLs, IP addresses, files, and DNS and email domains).

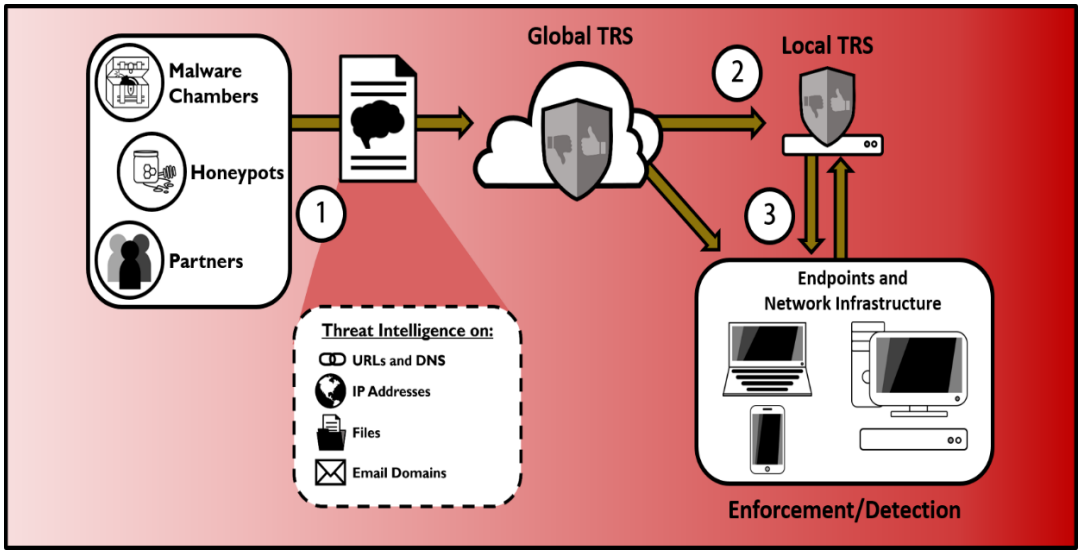

Figure 2: How TRS works

1. Intelligence data on reputation items is submitted to the Global TRS.
2. Local/customer owned TRS can pull reputation data from the Global/cloud TRS to supplement their own reputation data.
3. As clients reach out to external networks/enclaves, or the internet, network security devices like firewalls, now integrated with TRS and drawing on reputation data in real time, detect and block known threats.

[2] For more information on threat hunting, please refer to "Continuously Hunt for Network Intrusions," part of the NSA Cybersecurity Top 10 Mitigations.

Components

Reputation Categories. A category is a group of reputation items with similar characteristics. For example, some types of websites are known for having high amounts of malicious content and are categorized accordingly. Some examples of categories applied to websites include, but are not limited to, malware, social media, news, pornography, gambling, and drugs. The categories to be scanned are typically selected when adding TRS rules to an access policy.

Reputation Scores. A reputation score reflects the threat level assigned to URLs, IP addresses, files, and DNS and email domains. The scoring uses a threat behavior scale from known bad behavior, to unknown, to known good behavior, and it varies based on the TRS vendor. Some scales are numeric, for example, ranging from zero to ten or one to five. Items at the "known bad" end of the scale are assumed to be infected with malicious content. The "known good" end of the scale signifies reputation items with known legitimate sources which are believed to be good. Mid-scale reputation items typically have not yet been identified as good or bad. Most enforcement points allow the network security administrator to set a threshold for which reputation scores to block.

Reputation Databases. The last component is the reputation database. TRS vendors store and manage their reputation details in proprietary databases. The databases are constantly updated as new reputations are assigned and old reputations are updated. The database is offered as a cloud service where products automatically either download the latest reputation data or make real-time queries to the TRS. Some products support a local database which can augment global community reputation scores with local scores from file detonation chamber results or other sources specific to the customer.

Dynamic adaptation

Network security devices like firewalls and proxies must be constantly updated to prevent the latest threats. Security administrators can get overwhelmed trying to keep up with persistent adversaries, but TRS assesses threat levels for millions of websites, files, and domains for enterprise network customers, helping to prevent threats by blocking known bad sources. Resources like time and effort required to collect and apply threat intelligence can be allocated to other critical security tasks. TRS solutions enhance a layered, defense-in-depth strategy. Integrating TRS crowdsources the collection of threat intelligence, achieves resource- and cost-savings, and helps defenses to dynamically adapt with the rise and diversification of malicious activity.

Works Cited

[1] "Measuring the Impact of DMARC's Part in Preventing Business Email Compromise." The Global Cyber Alliance, 2019. [Online] Available: https://www.globalcyberalliance.org/dmarc-economics-benefits-report/
[2] "The Economic Value of DNS Security." The Global Cyber Alliance, 2019. [Online] Available: https://www.globalcyberalliance.org/dns-economic-value-report/
[3] D. Bianco, "The Pyramid of Pain." 2014 January 17. [Online blog] Available: https://detect-respond.blogspot.com/2013/03/the-pryamind-of-pain.html

Disclaimer of Warranties and Endorsement

Contact

Client Requirements / General Cybersecurity Inquiries: Cybersecurity Requirements Center, 410-854-4200, Cybersecurity_Requests@nsa.gov
Media inquiries / Press Desk: 443-634-0721, MediaRelations@nsa.gov