DATASHEET

Managing the Threat Intelligence Lifecycle through Scoring and Expiration

Many companies possess ample threat data and intelligence, yet they often perceive gaps in their protection, even as they add more and more data to their platform. This flood of information often leads to false positives, rather than the desired outcome of closing the gaps. The key deficiency lies in the absence of a method to prioritize this data according to an organization's specific needs. Such prioritization empowers security teams to reduce false alarms and concentrate on critical issues. The ThreatQ Platform's Scoring and Expiration features effectively tackle this challenge.

THREAT INTELLIGENCE - LIFECYCLE MANAGEMENT

When discussing threat intelligence, there is a tendency to focus our attention on using an indicator to block or alert on your platform. Indeed this is a key area. However, to truly maximize our ability to turn data into actionable information we must consider more than the "working" part of the intelligence lifecycle. We need to think about the entire lifecycle– from birth (choosing the right intelligence) to death (retiring data and ongoing review), and every step in between.

BIRTH

CHOOSING THE RIGHT INTELLIGENCE

One of the oldest rules of information management is "garbage in, garbage out." If you can avoid putting noise into your platform you can have high confidence in decisions made based on the information inside it. Choosing the right intelligence can be accomplished in a number of ways, beginning with the careful selection of sources.

One approach to avoid gaps in your intelligence is to collect data from multiple sources. No single entity can be perfect, so there is merit in casting a wider net, especially when your platform will automatically deduplicate. However with respect to intelligence, often "less is more".

Before ingesting all data from a feed, review a sample of the data you receive to see how beneficial it is before making it part of your regular dataset. One important area to evaluate is whether the feed provides useful context to help with prioritization and analysis for your own business purposes. Data without context is not information. At the same time, the source itself may carry enough weight to justify making it part of your regular dataset. For example, if your country's CERT supplies an IP address, it's probably a valid IP address to block.

You should also consider the type of data that the data source provides. If the source is focused on a particular geographical region and you're not in that region, is this feed for you? If the feed provides hashes in various formats, but you only need SHA-512 is it possible to pre-filter that data? If you'll never use it, don't ingest it.

When a new feed is added, automatically marking an indicator as Review will allow a human to check for the quality and context prior to acting upon the intel.

When ingesting from a new source, consider pushing those first few ingestions through a person for review. Once approved, a bulk action can change the indicator status to Active with the final step being to allow all ingested intel from this feed to be Active immediately. Bringing in data as Review before making it Active also allows you to determine and prioritize those indicators most relevant to your organization, further reducing the noise.

LIFE

PRIORITIZATION WITH SCORING

Most organizations have plenty of threat data and threat intelligence, yet they still don't feel they are adequately protected. Each company's unique threat landscape is based on three factors: external threats (to their industry, geography, profile and popularity), their internal infrastructure and the company's risk profile.

What's missing is a way to prioritize the data based on the requirements of a particular organization, both internal and external. This would allow security teams to reduce the noise from feeds, minimize false positives and focus on what matters to their organization and their threat landscape.

To attempt to solve this problem many intelligence providers and "blackbox" TIPs include a threat score. But, those scores aren't specific to you or your industry vertical, Instead, they are generic global risk scores. Over-reliance on these scores can lead to misallocation of resources, and can also limit your choice of threat intel sources and cause you to miss out on key intelligence. How would we prioritize threat intelligence if it did not carry this pre-built risk score? Or what if the same intelligence is scored differently by two sources?

A customer defined scoring methodology allows the team to dictate their own risk posture based on their resources, tools and other team priorities.

An important concept with scoring to remember is "relevance not severity." For example, if there are only Linux servers in an estate, then Windows vulnerabilities and the malware that targets it, and subsequently the indicators related to that malware, are all less relevant no matter their severity.

The same concept can be used to adjust the risk of an attack based on your own business profile. Priority should be given to threat intelligence that has context linking it to your geographic region or industry vertical. An adversary that is focused on high profile U.S. financial institutions is probably not going to pose as high a threat to an Australian local government's infrastructure.

Context from internal sources can also assist with prioritization and scoring. One of the most important types of intelligence is context gained from artifacts gathered through your incident response activities, such as IoCs from the SIEM and any attribution intelligence. If a particular adversary has been probing your network it is likely that they are engaging in a campaign that should be tracked and challenged.

The use of context can allow prioritization of indicators, both in raising and suppressing visibility of that intelligence, while maintaining a full audit trail of reasoning.

Filtering the massive volume of new indicators based on relevance is also essential. Many leading threat intel providers publish lists of indicators which provide new content at a magnitude of hundreds of thousands of indicators every 24 hours. Even with an aggressive expiration policy, downstream devices from the TIP would become overwhelmed with data after only a short time. Assigning a score to each indicator, using some sort of logic, enables the application of a simple filter that allows the user to efficiently prioritize while still making use of multiple layers of logic to define the score. When this reasoning is based on logic, it can be applied and modified en-masse without excessive effort.

This score should not be simply a copy of the CVSS severity. For example,a Remote Code Execution vulnerability affecting all Windows 11 devices is not an issue for a company that only uses Linux servers. The reality is never this simple, but this should demonstrate that relevance is as important as severity in managing the load on your downstream systems.

RETIREMENT: WHY EXPIRE DATA?

While it may be tempting to keep all data, even the best financed companies have limited resources. In addition, pressures on adversaries to avoid detection and capture mean that elements of their infrastructure regularly change. So the usefulness of data fades over time.

Malware detection in the past was based on binary hash matching, so virus writers added polymorphic routines to constantly change the hash. This results in a high volume of hashes created on a daily basis, with a reduced chance of a valid detection for every day after first ingestion, leading to a low return for the resource burden imposed on your detection systems.

The same can be said for infrastructure - for example, when considering IP addresses allocated to cloud providers. In this high churn environment, something that might have been malicious in the past may be perfectly legitimate at some point in the future.

Threat actors evolve their TTPs quickly, so holding onto old data for too long can strain resources and provide little security value.

Consider the following scenario. An IP address is used by an adversary as a C2 server in January. Authorities issue a takedown in February and the server is taken offline. The ASN owner re-issues the IP address to a new owner who builds a web application in July. How should that IP address be treated if detected in the SIEM, or in a connection attempt from the proxy?

It makes sense to focus our attention and resources on the most likely candidates, to avoid over burdening our systems and avoid false positives triggered by legitimate use of a previously malicious item. This means we need to start talking about a subset of all the indicators we have seen that are still actively being used, or in ThreatQ terms are still “Active”. This subset of indicators may be reviewed on a regular basis to remove them from our Active dataset, ensuring the tools leveraging the indicators for detection and protection are at maximum effectiveness.

WHY KEEP DATA AFTER IT HAS EXPIRED

In the example above, can the IP address be removed from the Active dataset, or should it be removed entirely? While it may not be desirable to block activity against the indicator, or have it alert in the SIEM, there can be benefits to retaining the full context related to an indicator. It may be related to an ongoing incident investigation, or you may want to retain the full context of a prior investigation. This makes it sensible to be able to keep the indicator longer than the duration that it is Active.

In the future it’s possible to discover potentially malicious activity that was not detected at the time of infection, and requires investigation in historical logs to understand the context that was relevant at the time. This helps to ensure that no other devices were infected and to identify the original adversary and malware.

For investigation and threat hunting purposes, it is worthwhile to create a policy to retain data for a certain period of time after it has expired.

Finally, with Advanced Persistent Threat (APT) re-use of infrastructure, a previously used and dormant C2 server or similar may be enlivened and we would want to retain full context of the original use to minimize duplicating research.

ALTERNATIVES TO ACTIVE AND EXPIRED STATUS

Sometimes, adding content to the TIP that flows directly to your downstream security stack can be important, as you need to protect your business as soon as possible. But what if you’re not sure of the data? Flowing that straight to your firewall when you have little confidence in it could cause a great deal of problems for the TI department. For this reason, it can be a good idea to ‘birth’ data in your system with a “Review” status. Once you’ve looked over it, and possibly enriched it with context, you can make an informed decision about the future status for this intelligence.

Quite often artifacts collected from analysis of a malware process can include infrastructure that was not itself malicious, but is an important part of identifying a particular malware signature. As an example, although non-malicious themselves, the Google DNS (8.8.8.8) and Google drive FQDN (drive.google.com) can be used as part of malware infection. A more nuanced example is a malicious FQDN related to an IP address that might further support several

FQDNs, most of which are simply sharing infrastructure. While it is probably unwise to block access to this IP, it could be valuable to record the use of this IP to aid an incident response investigation. For this purpose we can label an indicator with a status of “Indirect”.

If something is added inadvertently to the TIP, like a Google DNS IP, or even a company’s own IP or Domain name. we may end up blocking essential business processes and cause an incident ourselves. We can immunize against this outcome by storing known-good artifacts that may appear in the TIP in the future, such as your own domain name and Google, and maybe some important domains like google.com, youtube.com, and then setting them as known good with a status of “Whitelisted.” By making this status protected, even if someone were to accidentally ingest these domains, they would not become Active and not end up downstream.

Indicators aren’t black and white – there’s a lot of gray – so expand Active and Expired statuses to include Review, Whitelisted and maybe even your own custom status.

You shouldn’t consider yourself limited to the built-in statuses of Active, Expired, Review and Whitelisted. There are some use-cases that suggest alternative statuses. As an example you may want to differentiate between Active indicators that are to be blocked, and those you merely detect. This detection-only status can be useful for collecting threat intelligence of a campaign without tipping your hand that you are aware and protected. To implement this, we can create a custom status that represents the passive detection and use “Active” to represent the detect and block. It’s now a simple process to filter out the “Passive” indicators wherever required.

DEATH

WHEN TO SAY GOODBYE TO DATA

So, you are doing what’s required to manage the bulk of your threat intelligence lifecycle thoughtfully. You are careful in your selection of feed suppliers. You only ingest the actionable indicator types you need. And you prioritize so that the most relevant indicators are being used for correlation in the SIEM and blocking on your perimeter and XDR, before they finally expired and are sitting in your TIP.

Now, what do you do with the subset of data stored in your TIP that has no sightings against it, has a low score, and expired three months ago. Is it still providing value?

The fundamental truth of threat intelligence is that time is of the essence.

Just as you need to leverage automation to minimize the time between discovery of intelligence and leveraging it in your estate, you also need processes in place to handle retirement efficiently because threat intel gets old quickly.

Knowing that an APT has a campaign targeting your industry with a particular malware against a specific vulnerability is extremely valuable before the campaign reaches you, and is useful only for attribution after the campaign completes. Similarly, blocking a URL that is being used in phishing emails is a great way to keep your system safe. But if your business was never targeted and the campaign has finished, does that URL have value on your system? What about the related FQDN? Or the IP address that the FQDN resolves to? Does it still have that value six months after it was last reported?

If you are ingesting 50,000 new indicators per day, that is 18 million in a year. After five years of running your TIP you’d have nearly 100 million indicators, which is a lot of data to make sense of, store and process.

Data policies should be created with a “business first” approach, compiled into a data retention plan and then applied to your TIP. Regular review of the performance of this plan can help understand if your data is growing or shrinking, and at what rate.

ONGOING REVIEW

Consistent with all things related to cyber security, “don’t set and forget”. As technology continues to evolve, so do cyber threats and the methods for protecting against them. Your scoring needs should change along with your business and changes to the threat landscape. New adversaries arise, new software is deployed, and geo-political changes affect priorities too.

What worked on day one of your Threat Intelligence program may not be ideal by day 90 or day 360.

It's important to regularly review the quality of data being received from a vendor. The number of false positives and the number of true positives should be part of this review, as well as your scoring filter thresholds and the balance between relevance and severity.

This doesn't mean scoring is a full-time job, as changes can be quickly applied to your system in bulk through automation. However, regular review of your scoring should be scheduled along with a defined process for reporting and managing false positives. And if shifts in scoring reveal that the data overall from a particular source is no longer relevant, it may be time to circle back to the beginning of the lifecycle and review a new source of intelligence to bring into your program.

CONCLUSION

Threat intelligence lifecycle management is a pivotal element for any team and can serve as a force multiplier by setting the rhythm of daily operations, aligning efforts with a shared mission, and maximizing resource efficiency. To truly excel, scoring and expiration methods must be transparent and customizable, reflecting the team's unique parameters. ThreatQ's scoring and status capabilities empowers teams to regain control over their intelligence endeavors, allowing them to tailor intelligence strategies according to their risk thresholds.

The automation and deployment of threat intelligence have brought the industry to a crossroads—those who blindly operationalize it face the frustration of chasing illusory threats, while teams that integrate intelligence with their insights bolster their defenses. Where do you stand?

To explore a strategic approach to operationalizing threat intelligence through customizable scoring and expiration aligned with your environment, reach out to us at info@threatq.com for a demonstration.

> ***"ThreatQ's customer-defined Scoring is huge. We currently have one false positive per month, whereas eight months back we had ten per day."***
>
> *– Threat Intelligence Manager, Fortune 500 Technology Company*

ABOUT THREATQUOTIENT™

ThreatQuotient improves security operations by fusing together disparate data sources, tools and teams to accelerate threat detection, investigation and response (TDIR). ThreatQ is the first purpose-built, data-driven threat intelligence platform that helps teams prioritize, automate and collaborate on security incidents; enables more focused decision making; and maximizes limited resources by integrating existing processes and technologies into a unified workspace. The result is reduced noise, clear priority threats, and the ability to automate processes with high fidelity data. ThreatQuotient's industry leading integration marketplace, data management, orchestration and automation capabilities support multiple use cases including threat intelligence management and sharing, incident response, threat hunting, spear phishing, alert triage and vulnerability management. For more information, visit www.threatquotient.com.