OPTIMIZING THREAT OPERATIONS:

Prioritize Threat Intelligence through Scoring

by Ryan W. Trost, Co-Founder & CTO

THREATQUOTIENT

WHITEPAPER

OPTIMIZING THREAT OPERATIONS:

Prioritize Threat Intelligence through Scoring

WHAT IS SCORING?

Scoring represents the measure of "threat risk" facing an organization based on the calculated intersection derived from supporting internal and external intelligence.

As each day passes, threat intelligence platforms (TIPs) are automatically absorbing hundreds, thousands or potentially millions of indicators, forcing teams to quickly define an 'all in panoramic-view' scoring strategy. Without this capability, any TIP can't fully address the data overload problem that security operations and threat intelligence teams suffer from.

ThreatQ provides the first highly customizable intelligence-scoring platform allowing teams to define scoring parameters that the platform will use to automatically re-score providers' intelligence as it enters their ThreatQ system. The result is a customer-driven score based on their own views of the world and NOT that of the provider!

WHY SCORING MUST BE CUSTOMIZED TO YOUR ENVIRONMENT

To squeeze every ounce of benefit from threat intelligence, it is critical the intelligence conform to the vantage point and mission of the team operationalizing it. The ability to customize the threat intelligence score allows teams to re-align external intelligence to their own risk posture, prioritize threats to their organization (and thus remove noise at the same time) and be more efficient in deploying the right intelligence to the proper tools.

REAL RISK SCORES

Lots of intelligence providers and "blackbox" TIPs include a threat score. But, those scores aren't specific to you or your vertical, but, rather, a generic global risk score. This leads companies to a false sense of risk management and a misallocation of resources because that adversary, attack or indicator might not even be launched against your industry, although it steals cycles from your team because the provider generically scored it high.

PRIORITIZATION

Indicators trigger alerts, which, in turn, initiate analyst investigations. However, alerts are generally not created equal and teams end up wasting a significant amount of time chasing ghost alerts (false positives). A customer-defined scoring methodology allows the team to dictate their own risk posture based on their resources, tools and other team priorities.

THE RIGHT INTELLIGENCE TO THE RIGHT TOOLS AT THE RIGHT TIME

With the panoramic scoring feature, customers quickly become more strategic about WHERE they deploy their intelligence! Teams can export intelligence with more confidence to specific security technologies based on the risk it poses to their organization. For example, threat intelligence with higher threat scores and, thus more reliable, will be deployed to blocking technologies (i.e., firewalls, IPS, DNS, web-proxy, endpoint, etc.), whereas, intelligence that poses less of a threat [read: less reliable] will be distributed to detection technologies (i.e., IDS, netflow, etc.) to minimize any operational impact due to false positives. This is a critical component for companies with limited infrastructure tools already pushed to the sensor's limits.

SCORING BEST PRACTICES

About 12 months ago, threat intelligence providers heard the cries from the industry that, although the provided threat score was appreciated, the industry wanted to know "why?" and "how?" that score was determined. As a result, providers expanded their APIs to include not only the indicator's risk score, but other pertinent information – for instance, target industry, attack vector, role, category, etc. This helped educate and enlighten defenders, but they still DID NOT have a mechanism to modify the third-party score to reflect what it SHOULD be for them.

The underlying premise of our scoring methodology is to:

1) Allow customers to control their own destiny through weights AND prioritization to accurately reflect their risk temperature
2) Allow customers to customize the scoring algorithm – *because nobody knows your environment better than you*
3) Provide a scoring range that is digestible; 1-5 isn't drastic enough and 1-1000 is too difficult to conceptualize and deviates significantly from industry best practices
5) Find middle ground – too few elements does not accurately reflect the risk, whereas, too many elements overwhelms customers and is ignored
6) Offer Score Transparency – AT ALL TIMES displaying how the score was calculated!
7) Ensure the score reflects local environmental variables (i.e., sandbox results, ticket results, observed dates, etc.)
8) Most importantly, updates the indicator score every time a new piece of information is appended to it

SCORING ALGORITHM

The scoring algorithm is an aggregate score of the PARAMETER * SCORE which provides a customizable, multi-dimensional approach.

Parameters: The parameters allow customers to determine how granular and sophisticated they would like scoring to be. Having multiple datapoints provides a more precise scoring algorithm – too few are too broad, whereas, too many is overwhelming and ignored.

Score: The Score is a customer-defined evaluation, which compares elements of equal tier. A Score of -10 to 10 provides a complex but not overwhelming scorecard. The negative score capability caters to more advanced organizations that want to be able to manipulate/man-handle scores beyond 0-10 or 0-100 (example forthcoming).

PARAMETERS

The scoring algorithm takes the following parameters into consideration:

1) Source
The Source of an indicator is a fundamental evaluation of the maliciousness of the information, more specifically, the customer's trustworthiness of the information being published by the Source.

- All indicators must have a Source (internal [ticketing system, sandbox, etc.] or external)
- All attributes have a Source
- Multiple Sources can increase or decrease the risk score based on the historical fidelity of the Source

2) IOC Type
The IOC type maps back to how a customer can/does leverage indicators for detection or blocking. Indicator types have different lifecycles based on team and resources, etc. For example, in a previous life, my USG SOC Team could not operationalize hashes due to customer constraints, but, we still wanted to ingest them for peripheral analysis/investigation purposes. So, those would receive a threat score of 0 or even a negative number.

3) Attributes
An indicator's context is a core ingredient of an indicator's threat score because it helps distinguish the type of threat. Most attributes can be broken down into three categories including: describes the indicator, describes the attack, or describes the adversary. Each of these plays a critical part in the indicator's threat score.

4) Adversary Attribution
Indicators associated with adversary groups are another mechanism to weight the risk levels of an indicator – the accuracy of attribution is a different debate. The scoring of the various individual adversaries provides customers with a granular ability to help score the risk of indicators stemming from a specific adversary.

HOW DOES THE INDICATOR SCORE LOOK ON THE IOC DETAIL'S PAGE?

Because ThreatQ allows the threat score to be completely defined by the customer/team, it is the most important attribute. However, even more crucial is the transparency of how that score was calculated. The key is to display a summary of supporting data and associated scores in a quick efficient manner. Figure 1 provides a quick-at-a-glance breakdown of the calculated score, whereas, Figure 2 highlights the score within the larger indicator detail's page.

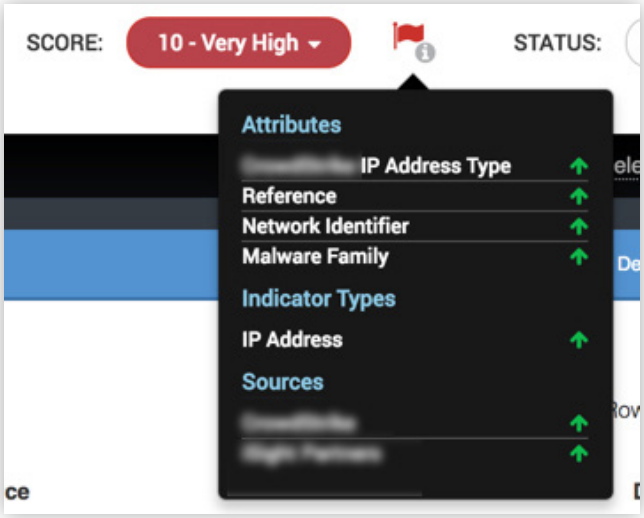

Figure 1. Summary Breakdown of an Indicator's Score

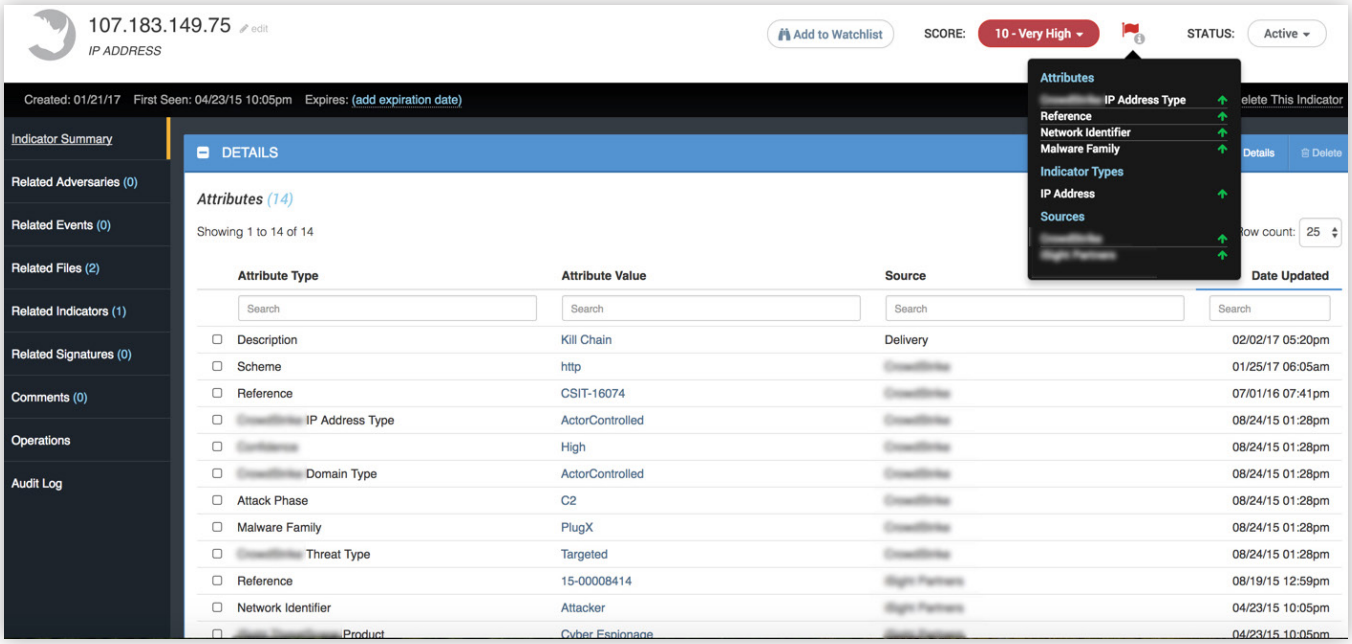

Figure 2. Risk Score Placement on an Indicator Detail Page

The design of the intelligence-scoring algorithm is based on real-world use cases from managing a Defense Industrial Base (DIB) SOC to ensure the algorithm works for less mature teams and more mature teams. In order to really maximize your threat intelligence, teams need to automatically re-calculate the third-party score during the ingestion process, whether from a commercial feed, third-party services (i.e., VirusTotal, DomainTools, etc.) or even from a "community fight club."

THREATQ SCORING IN ACTION

Let's take a closer look at how ThreatQ accomplishes this through two different use cases. In these examples, let's say we are a SOC Team within the DIB. Our primary purpose is to identify and remediate cyber intrusions as quickly as possible. However, due to the high volume of intelligence we are consuming, we cannot validate everything and deploy it in a timely manner without overwhelming our own defenses. We need a way to "customize" the intelligence during the ingestion workflow to more accurately align with the threats targeting the organization, team and partners. The following use cases will help describe the power and flexibility allotted your team to discern which threats are real – and which are forcing you to chase ghosts.

USE CASE — DETERMINING "MY REAL" RISK SCORE

Let's focus on an attribute that a feed vendor provides called "Target Industry." If an indicator's target industry is Universities or Retail with a risk score of HIGH, does that translate evenly if my company is part of the DIB?! Absolutely not! ***Does the indicator still pose a threat? Absolutely, BUT NOT at the same threat level as the industry being targeted.*** *As a result, intelligence being consumed is now more precise to my needs as seen in Figure 3 (red on the right representing the bigger threat to your organization).*

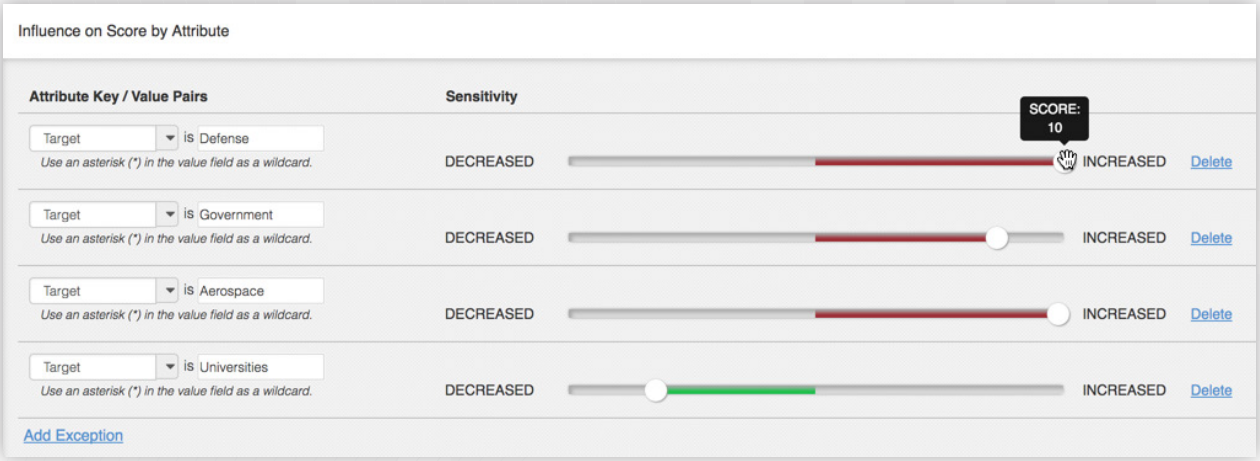

Figure 3. Customizing Scoring within a Single Attribute [Target Industry, in this case] is SUPER Simple

USE CASE — QUALITY VS. QUANTITY

My main concern is heeding the volume of indicators being deployed to my sensor grid.

Another feed vendor provides critical indicator context to help recipients better determine what to do with the indicator. ThreatQ allows customers to take advantage of that information and prioritize it against all the other intelligence coming into the system. In this example, the indicators posing the bigger risk are bubbling to the top, whereas, the less threatening indicators are only being brought in for peripheral investigations, as seen below in Figure 4. You will notice that indicators associated with DomainRegistrants, DomainResellers, or DomainAdministrators have negative scores associated to them; whereas, SpearphishSender and C2 have the maximum threat scores associated to them. Diligent companies will maintain visibility to email addresses linked to broader operations, but are only operationalizing the email-related indicators that pose a threat.

IMPORTANT NOTE: The great aspect about ThreatQ is how versatile it is. For example, DomainRegistrants, DomainResellers, and DomainAdministrators are NOT indicators most teams will put into defenses because they are secondary indicators and are not used in the actual attacks. However, teams use those indicators as proactive "alert triggers" through WHOIS, DomainTools, etc. So, although they received negative indicator scores due to the lack of "DIRECT" operational risk posed to the organization, the team can still create an IOC Export with those specific indicator attributes and build that into a proactive workflow to be alerted when the associated email address does register a new domain.

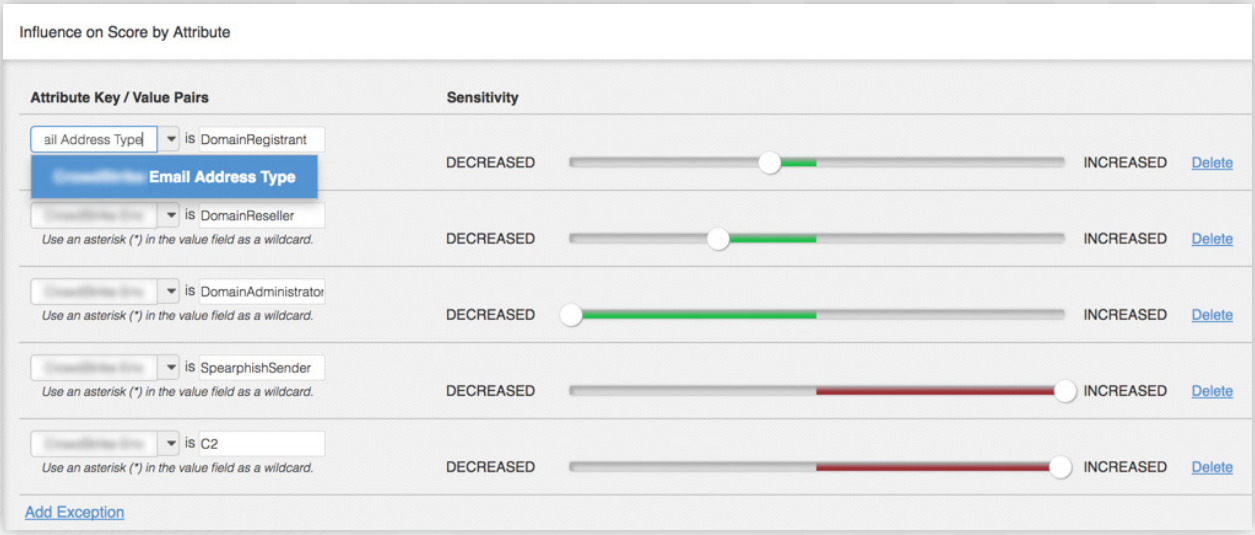

Figure 4. Customizing Scoring across Multiple Attributes is SUPER Simple

TQ SCORING APPLIED TO ITS THREAT LIBRARY

One of the primary purposes of scoring intelligence is to have the utmost control and accuracy over the whole dataset – not just a single indicator or its attributes. By applying the ThreatQ scoring customization, you really get a sense of its power in a common operational maturity model.

1) **Entry Approach:** Newer teams will leverage all the indicators equally. A quick view of this shows a total number of ingested indicators in the dataset are **910,321**.

2) **Maturing Approach:** The next stage of evolving is focusing on the commercial's alert severity. The breakdown of indicators by the provider's confidence score (low/medium/high) is below, which still remains a hefty volume of indicators.

   Malicious Confidence = HIGH = **319,754**
   Malicious Confidence = MEDIUM = 497,969
   Malicious Confidence = LOW = 80

3) **Seasoned Approach:** The most powerful capability is applying a custom score that aligns with our resources, threats and capabilities. In this use case, let's maintain the vantage point of a SOC Manager within the Defense industry and apply higher custom scores to the attributes that are more meaningful to "my" organization. In Figure 5, we focus on "increasing" the pertinent risk scores versus strategically allocating negative values for attributes on the opposite end of the spectrum which pose little to no risk to our organization.

By applying the custom scoring algorithm, the ~1M is broken down into risk categories as followed:

Risk Category	Number of Indicators*[1]	Percentage
Very High	27,358	~3%
High	45,651	~5%
Medium	312,623	~34%
Low	248,211	~27%
Very Low	276,478	~30%

These results help the teams focus on what's important – not only from a "what to deploy" mindset, but, more importantly, what information should the intelligence teams scrutinize and pivot from! This is a great example of the value ThreatQ's Scoring capability offers – the ability to empower teams to re-prioritize intelligence based on their own team, resources, true risk, tools, etc. We take the ~1M indicators, remove noise and significantly reduce ~97% to a more digestible amount based on the company's and team's risk posture.

*Note: there are several complex strategies that can be applied given our scoring flexibility. However, in this approach, as highlighted in Figure 5, the idea is the more "relevant" intelligence that supports the indicator, the greater risk that indicator poses to the organization. This is a more mature model over blindly categorizing everything coming from commercial feeds as HIGH fidelity as seen above in #2.

1This analysis is a snapshot in time. With ThreatQ's scoring capability, the Threat Library will self-tune as more data and context is received. In addition, aging and expiration will affect results differently over time.

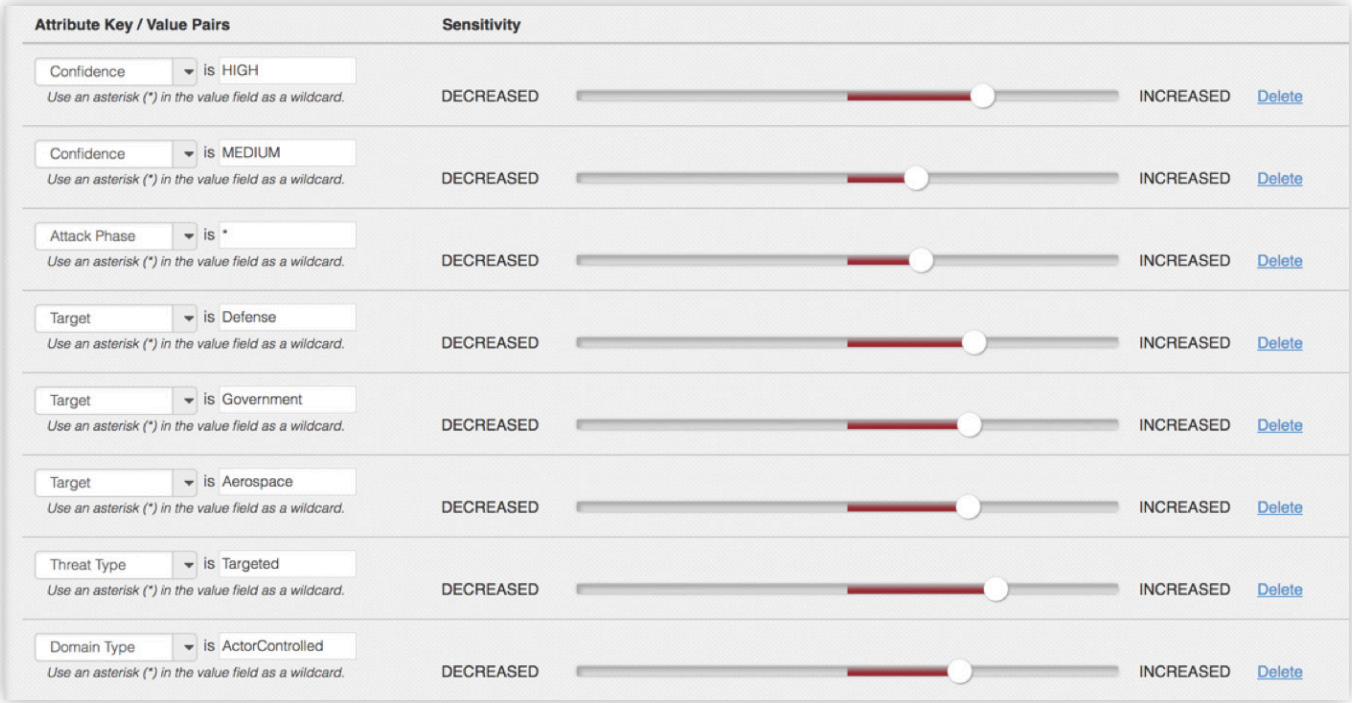

Figure 5. Applying Customizing Scoring to Improve Accuracy Based on MY Needs

CONCLUSION

Scoring is a critical component for any team because it sets the day-to-day pace, aligns teams to a mission and supports efficiency across resources allowing teams to appear bigger than they are. But, to be truly effective, the scoring methodology should be transparent and customizable using parameters the team sets. ThreatQ's Scoring capability offers a chance for teams to take back control of their intelligence efforts and redefine intelligence based on their own risk levels. The ability to automatically consume and deploy threat intelligence has led the industry to a crossroads – those who blindly operationalize it feeling the pain of chasing ghosts, and teams who overlay the intelligence with their own insights to build stronger defenses. Which one are you?

To learn more about how to take a strategic approach to operationalizing threat intelligence through scoring that aligns to your environment, contact us at info@threatq.com for a demo.

ABOUT THREATQUOTIENT

ThreatQuotient™ understands that the foundation of intelligence-driven security is people. The company's open and extensible threat intelligence platform, ThreatQ, empowers defenders to ensure the right threat intelligence is utilized within the right tools, at the right time. Leading global companies are using ThreatQ as the cornerstone of their threat intelligence operations and management system, increasing security effectiveness and efficiency.

For additional information, please visit threatq.com.