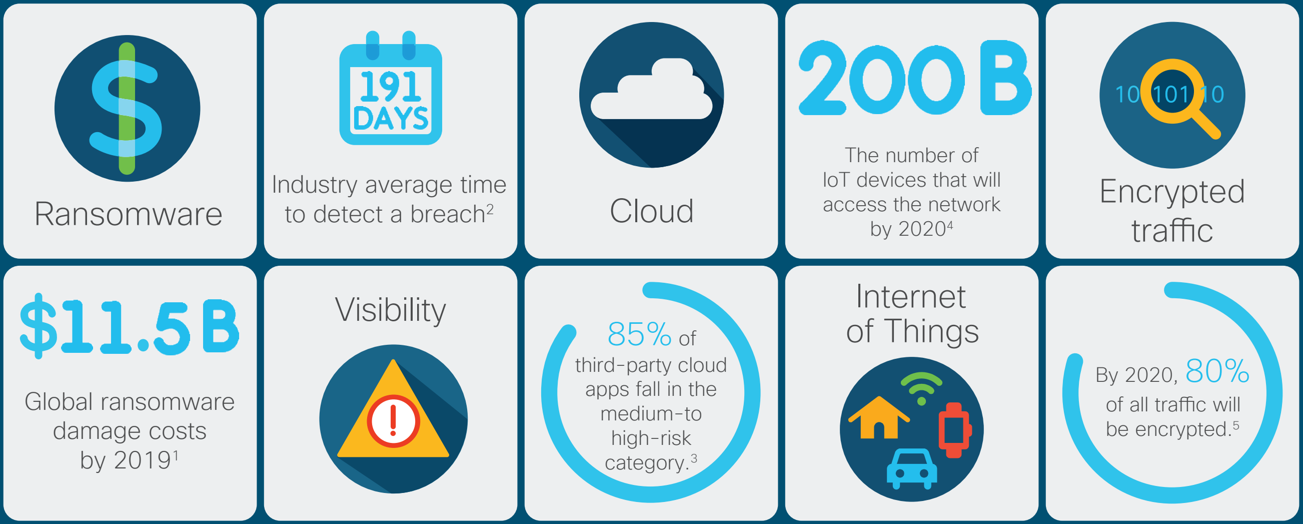
Cisco Security Analytics

A deep-dive into the unique behavioral modeling and machine learning techniques for advanced threat detection.

I. Introduction

The increasing complexity in the enterprise network has created many blind spots. Today, employees are connecting to the network from many places and from multiple devices. The number of smart devices accessing the network and the use of public cloud services continue to grow. And encrypted traffic is on the rise for data protection and privacy. All this has enabled organizations to accelerate business outcomes and move toward digitization. But at the same time, it has also increased the opportunities for threat actors to hide and persist undetected within your digital business. Threats continue to evolve rapidly in terms of scale and sophistication.



Ransomware

191 DAYS
Industry average time to detect a breach[2]

Cloud

200 B
The number of IoT devices that will access the network by 2020[4]

Encrypted traffic

$11.5 B
Global ransomware damage costs by 2019[1]

Visibility

85% of third-party cloud apps fall in the medium-to high-risk category.[3]

Internet of Things

By 2020, 80% of all traffic will be encrypted.[5]

Sources:
1. Cybersecurity Ventures 2. Ponemon Institute 3. 2017 Cisco Annual Cybersecurity Report 4. Intel 5. Gartner

Contents

The main goal of attackers is to get in and remain present or persist. They may make themselves known by demanding ransomware or stealing information, but early detection of their presence is key. As threats continue to evolve, perimeter-based solutions can not be 100 percent effective. Attackers no longer break into your network; they simply log in with stolen credentials. You need a solution that can detect advanced threats early in the attack lifecycle and before they are able to compromise your network and create a significant impact.

Cisco® Stealthwatch is the industry-leading security analytics solution providing comprehensive threat visibility into the extended network. It can detect and respond to advanced threats, and help simplify network segmentation using a combination of **behavioral modeling, multilayered machine learning, and global threat intelligence**. Because attackers aren't employing just one method to breach your network, Stealthwatch employs multiple analytical techniques to detect threats early and helps ensure that the eviction is complete. And it is the first and only solution in the industry that can detect malware in encrypted traffic without any decryption using Encrypted Traffic Analytics

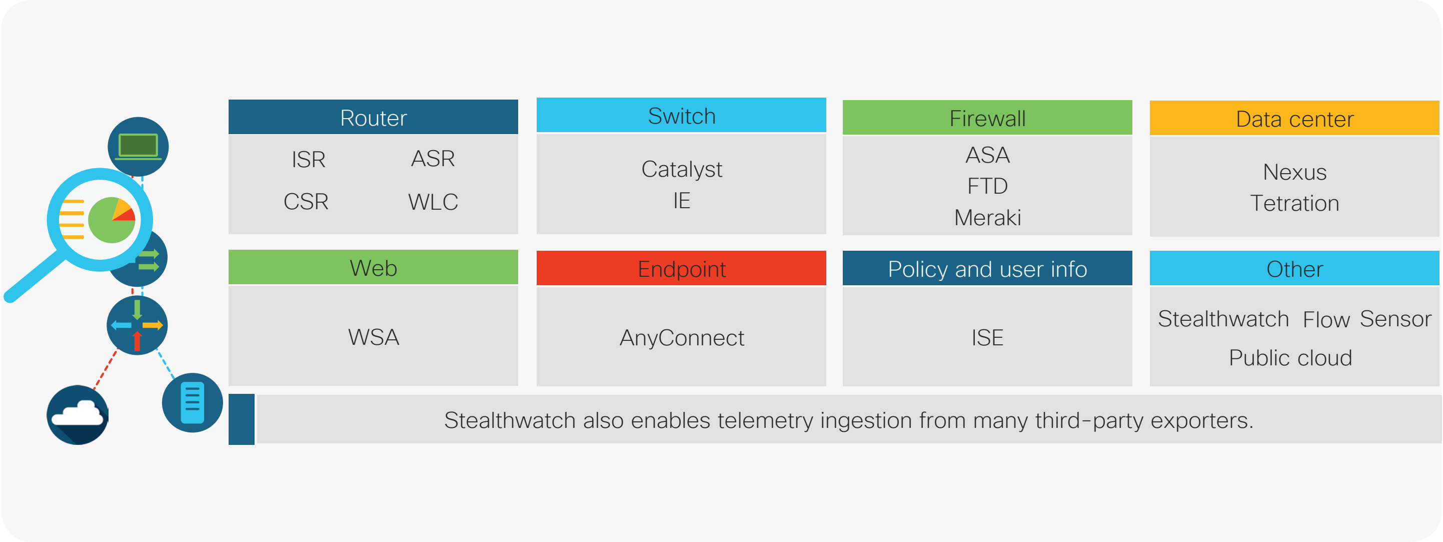
II. Collecting the right data using the existing network

If you want to know your digital business and how it behaves on a daily basis, you must have the necessary and sufficient telemetry. All of your network, including routers, switches, and firewalls, is able to deliver rich telemetry, and now with Encrypted Traffic Analytics, Stealthwatch can also analyze encrypted traffic to detect malware without decryption and govern the quality of network encryption within your digital business.

It also collects metadata from Cisco Identity Services Engine, Cisco AnyConnect® Network Visibility Module, and other supporting systems to make sure you have user and application context to the network behavioral analytics (Figure 1). Stealthwatch acts as a "general ledger" for your digital business, so you have an account of who, what, where, when, and how everything is behaving over time.

With a **single, agentless** solution, you get scalable security that grows with your digital business. Instead of your having to deploy hundreds of sensors throughout your network, the network infrastructure itself becomes the sensor, and this leaves threat actors with nowhere to hide. Also, Stealthwatch is a vendor-agnostic solution that enables telemetry ingestion from many third-party exporters. So it doesn't matter what kind of network infrastructure you have, Stealthwatch can leverage it as a data source to improve security.

Figure 1. End-to-end visibility across any telemetry



Router	Switch	Firewall	Data center
ISR ASR CSR WLC	Catalyst IE	ASA FTD Meraki	Nexus Tetration
Web	**Endpoint**	**Policy and user info**	**Other**
WSA	AnyConnect	ISE	Stealthwatch Flow Sensor Public cloud

Stealthwatch also enables telemetry ingestion from many third-party exporters.

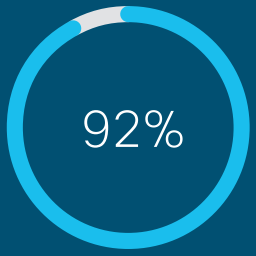
ISR = Cisco Integrated Services Router; ASR = Cisco Aggregation Services Router; CSR = Cisco Cloud Services Router; WLC = Cisco Wireless LAN Controller; IE = Cisco Industrial Ethernet; ASA = Cisco Adaptive Security Appliance; FTD = Cisco Firepower® Threat Defense; WSA = Web Security Appliance; ISE = Identity Services Engine

III. Multiple analytical techniques working together

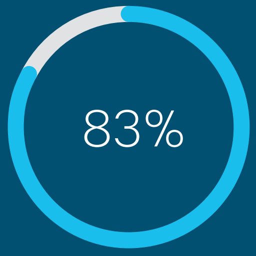
Stealthwatch has three core approaches that work together to leave no stone unturned for catching threats at the earliest point in the attacker's activities.

1. Behavioral modeling

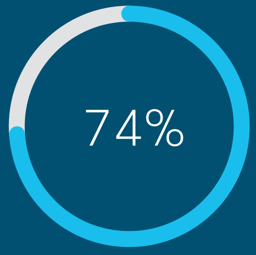
Stealthwatch closely monitors the activity of every device on the network and is able to create a baseline of normal behavior. In addition, it also has a deep understanding of known bad behavior. It applies close to 100 different **security events** or heuristics that look at various types of traffic behavior, such as scanning, beaconing host, brute force login, suspect data hoarding, suspect data loss, etc. These security events feed into high-level logical alarm categories. Some security events can also trigger alarms on their own. So the system is able to correlate multiple, isolated anomalous incidents and piece them all together to determine what kind of attack might be in play, and also tie it to a specific device and user (Figure 2). The incident can be further investigated by time as well as by the associated telemetry. This is context at its best. Physicians examining a patient don't look at a symptom in isolation to figure out what's wrong. They look at the whole picture to provide a diagnosis. Similarly, Stealthwatch records every anomalous activity in the network and looks at it holistically to generate contextual alarms that can help security teams prioritize risks.



The percentage of security professionals who see value in behavioral analytics tools



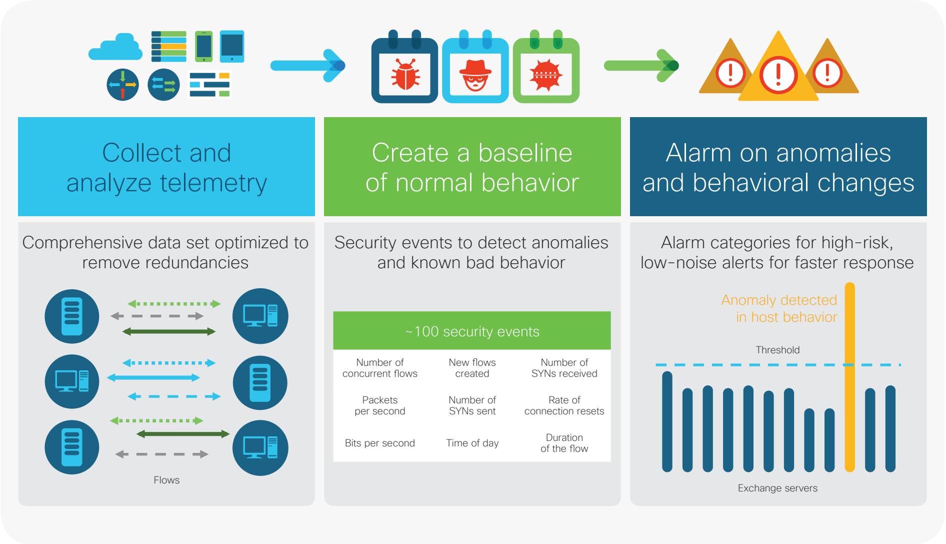
The percentage of CISOs who rely on artificial intelligence to reduce the level of effort for securing the organization



The percentage of CISOs who rely on automation to reduce the level of effort for securing the organization

Source: Cisco 2018 Security Capabilities Benchmark Study

Figure 2. Anomaly detection using behavioral modeling


Collect and analyze telemetry
Comprehensive data set optimized to remove redundancies
Flows
Create a baseline of normal behavior
Security events to detect anomalies and known bad behavior
~100 security events
Number of concurrent flows
New flows created
Number of SYNs received
Packets per second
Number of SYNs sent
Rate of connection resets
Bits per second
Time of day
Duration of the flow
Alarm on anomalies and behavioral changes
Alarm categories for high-risk, low-noise alerts for faster response
Anomaly detected in host behavior
Threshold
Exchange servers

2. Multilayered machine learning

Stealthwatch also applies **machine learning**, both supervised and unsupervised, to discover advanced threats and malicious communications. It integrates with a cloud-based multistage machine learning analytics pipeline which correlates threat behaviors seen in the enterprise with those seen globally.

The system analyzes user and device behavior to discover malware infections, command-and-control communications, data exfiltration, and potentially unwanted applications operating in an organization's infrastructure. There are multiple layers of processing, where a combination of techniques from artificial intelligence, machine learning and mathematical statistics helps the network to self-learn its normal activity so it can identify malicious activity.

This network security analytics pipeline, which collects telemetry from every part of the extended network, including encrypted traffic, is unique to Stealthwatch. It gradually builds a notion of "what is anomalous," then classifies actual individual pieces of "threat activity," and finally arrives at a final conviction of whether or not a device or user is in fact compromised. It is through a very careful analysis and correlation that we are able to bring in small pieces of evidence that all together will allow us to finally convict a compromised entity.

This capability is important because a typical enterprise may receive tons of alerts daily, and it's not possible for resource-strapped security teams to investigate all those alerts. The machine learning engine processes massive amounts of data in near real time to discover critical incidents with high confidence and is also able to suggest clear courses of actions to remediate quickly.

3. Global threat intelligence

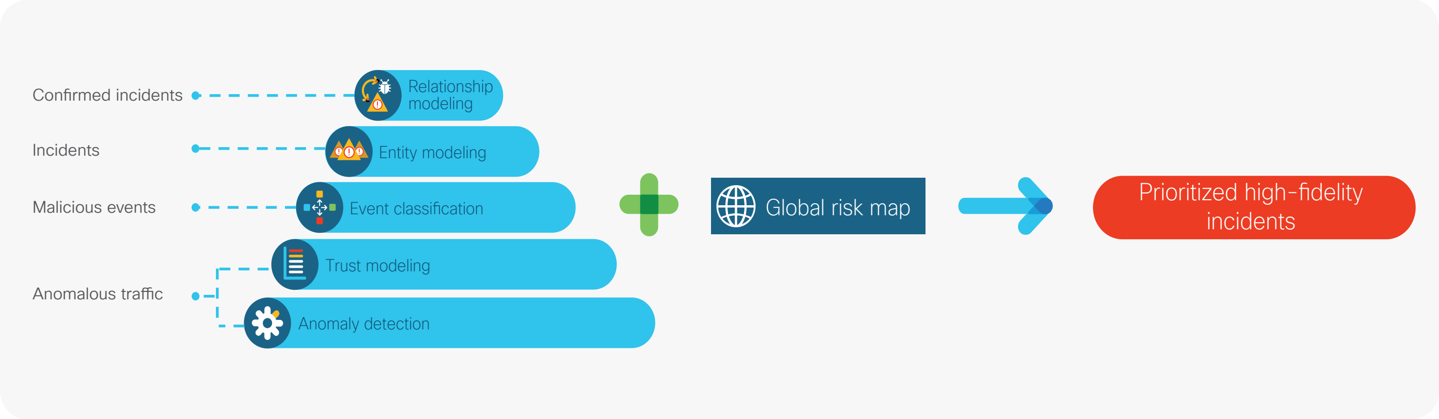
One of the advantages that attackers have is that they can apply the same attack on multiple targets, and the odds are that they'll be successful across them all because these victims are all constrained to their local view of the threat activity. But what if you had information about malicious IPs and domains, or a new strain of malware used in a campaign, and could map the alerts to this global threat intelligence? You would greatly reduce the time to detection as well as increase the fidelity of detection.

A global threat intelligence feed powered by the Cisco Talos™ intelligence platform provides an additional layer of protection against botnets and other sophisticated attacks. It correlates suspicious activity in the local network environment with data on thousands of known command-and-control servers and campaigns to provide high-fidelity detection and faster threat response. We have assumed the viewpoint of these threat actors, and from this position, we gain the advantage.

IV. What is multilayered machine learning and how effective is it?

Let's take a closer look at the multiple machine learning techniques used by Stealthwatch. When an incident is passed on to Stealthwatch's machine learning engine, it goes through a funnel of security analytics that uses a combination of supervised and unsupervised machine learning (Figure 3).

Figure 3. Power of multilayered machine learning



Layer 1: Anomaly detection and trust modeling

This layer is able to discard 99 percent of the traffic with the help of statistical **anomaly detectors**. These detectors together maintain complex models of what is normal and, in contrast, what is anomalous. But what is anomalous might not necessarily be malicious. There are too many things going on in your network that have nothing to do with threat—they may be just **weird**. But it's important to classify them apart from threat behavior. For this reason, the output of these detectors is analyzed further to remember those behaviors that are odd but can be explained and trusted. The end result is that only a small percentage of the most relevant flows and requests pass on to layers 2 and 3. Without the application of such machine learning techniques, the operational costs to separate the signal from the noise are too expensive.

Benefits of security analytics powered by multilayered machine learning and the Global Risk Map include:

- Detection of "known-unknown" threats (previously unseen variations of known threats, malware subfamilies, or related new threats) and "unknown-unknown" threats (net-new malware), in addition to known threats
- High-fidelity and automated alerts allowing resource-strapped security teams to focus on high-risk incidents
- Correlation of local threats to global campaigns for faster mitigation
- Scalable security with the ability to process vast amounts of telemetry

Anomaly detection: In the first step, anomaly detection employs statistical machine learning methods in order to separate the statistically normal traffic from anomalous traffic. More than 70 individual detectors process conversational telemetry records collected by Stealthwatch for traffic traversing your network perimeter, select internal Domain Name System (DNS) traffic, and proxy data if available. Each request is processed by the more than 70 detectors, and each detector applies a different statistical algorithm, generating a score for detected anomalies. These scores are combined and produce a single score per individual request by again applying multiple statistical methods. The aggregate score is then used to separate normal and anomalous traffic.

Trust modeling: Next, similar requests are grouped together, and the anomaly score for those groups is aggregated as a long-term average. Over time, more requests are analyzed to produce a long-term average anomaly score, thereby reducing false positives and false negatives. The results of trust modeling are used to select a subset of traffic with anomaly scores above a certain dynamically determined threshold to move on to the next layer of processing.

Layer 2: Event classification and entity modeling

This layer is focused on classifying the findings of the previous stages into specific malicious events. Event classification is performed by a magnitude of machine learning classifiers aimed at a consistent precision rate of more than 90 percent. These include:

- Neyman-Pearson-based linear models
- Support vector machines using multiple-instance learning
- Neural networks and random forests

Next, these isolated security events are associated with a single endpoint over time. This is where the threat narrative is being assembled, resulting in a complete picture of how this threat actor escalated to a particular outcome.

Event classification: The statistically anomalous subset from the previous layer is classified into 100 or more categories using classifiers. Most classifiers are based on individual behavior, group relationships, or behavior on a global or local scale, while others can be very specific. For example, a classifier may indicate command-and-control traffic, a suspicious extension, or an unauthorized software update. The output of this phase is a set of categorized anomalous events with security relevance.

Entity modeling: If the amount of evidence supporting the malicious hypothesis about a specific entity exceeds the significance threshold, a threat is created. The events that contributed to the threat creation are linked to that threat, and become part of a discrete long-term model of the entity. As evidence accumulates over time, the system creates new threats when the significance threshold is reached. This threshold is dynamic and intelligently adjusts based on threat risk level and other factors. The threat is then visible in the web interface dashboard and continues into the next layer.

Layer 3: Relationship modeling

Relationship modeling aims to synthesize the previous layers from a global perspective, offering not only the local context, but also the global context of this incident. It is here where you can determine how many organizations have seen this attack to know if you are being targeted specifically or are affected by a global campaign.

Incidents are either **confirmed** or **detected**. A confirmed incident carries with it a **99 to 100 percent confidence** because these techniques and tools were previously observed in the larger global scope. Detected incidents are unique to you and part of a very targeted campaign. The former findings are delivered with known courses of actions, saving you time and resources in your responses. The latter are delivered with the investigation tools you will need to understand your attacker and the extent of the targeted campaign on your digital business. As you can imagine, confirmed incidents far outnumber the detected for the simple reason that confirmed incidents are inexpensive for threat actors to deliver, while detected incidents are expensive for the actors because the incidents need to be new and customized. By enabling the detection of confirmed incidents, the economics of the game finally shift in the favor of the defenders, giving them some advantage.

Global Risk Map

The Global Risk Map is the result of the analysis applied by the machine learning algorithms on one of the largest data sets of its kind in the industry. It provides broad behavioral statistics about the servers on the Internet, even if they are unknown. These servers are related to attacks, may be exploited, or may be used as part of an attack in the future. This is not a blacklist, but a holistic picture of the server in question from a security perspective. This context-driven information about the activity of these servers allows Stealthwatch's machine learning detectors and classifiers to accurately predict the risk level associated with the communication to these servers.

On a given day, attackers might be able to evade one of these detection techniques, but it's impossible to hide when all are applied together in a pipeline and feed relevant data to each other.

It is very difficult to verify everyday claims of various security products related to the application of machine learning and artificial intelligence. That's why we have **published research** to explain the advanced analytical techniques (including machine learning) that are applied to the telemetry. The product team is constantly working to improve the detectors and classifiers to stay ahead of evolving threats. We also collect feedback from users on every alert we generate to determine whether it was useful or not. Our numbers for alerts being useful remain **well above 90 percent**, and we continue to strive for this performance indicator. In the end, we are not here to be fancy data scientists; we are here to be useful to our customers.

V. Great, so what kind of things can this analytics pipeline find?

More importantly, let's discuss some of the threats that Stealthwatch is able to detect by the application of various analytical techniques mentioned earlier.

Malware detection in encrypted traffic without decryption

Encryption technology has enabled much greater privacy and security for enterprises that use the Internet to communicate and transact business online. However, businesses are not the only ones to benefit from encryption. Threat actors have leveraged these same benefits to evade detection and to secure their malicious activities. Traditional threat inspection with bulk decryption, analysis, and reencryption is not always practical or feasible, for performance and resource reasons.

Encrypted Traffic Analytics is a new technology that is enabled by the new Cisco network and Stealthwatch, one of the major outcomes of which is the ability to detect malware in encrypted traffic without any decryption–an industry first. Encrypted Traffic Analytics produces two new data elements: the sequence of packet lengths and times and the initial data packet. The initial data packet is a treasure trove of metadata, because, remember, all encrypted sessions start out unencrypted initially. Cisco's unique Application-Specific Integrated Circuit (ASIC) architecture provides the ability to extract these data elements without slowing down the data network.

This enhanced telemetry is passed on to Stealthwatch, which applies the multiple analytics techniques described in the previous section to detect malware in encrypted traffic with high fidelity.

Advanced Persistent Threats (APTs)

APTs are highly targeted attacks against an organization with the primary purpose of stealing valuable information without being detected, so they can continue to persist over a long period of time. Using advanced behavioral modeling, Stealthwatch can gain a deep understanding of the normal behavior within the organization combined with the knowledge of known bad behavior using the Global Risk Map. Thus, it's able to detect activities associated with APTs such as reconnaissance, scanning, command-and-control communications, suspicious lateral network behavior, etc.

Insider threats

Among the most valuable assets that an organization has are its intellectual property, confidential information, and information stored in the company networks. Data breaches cost organizations millions of dollars. The average cost per lost or stolen record is $158, which amounts to about $4 million per year on data breaches. Insider threats, such as compromised user credentials or disgruntled employees, hoard data in order to exfiltrate it to the outside world for financial gain or just to cause harm.

Using behavioral modeling to detect anomalous behavior, a "data hoarding" or "exfiltration" alarm is generated. The host that triggers the alarm can be investigated with one click. Stealthwatch uses the network to provide additional contextual information about the host such as username, MAC address, location, etc. And if needed, Stealthwatch can **quarantine** the suspected host off the network. This is enabled by the integration with Cisco Identity Services Engine.

Malware propagation

Stealthwatch can not only detect if a host is infected with malware, but also track how that malware has spread within the network and what other hosts have been compromised. This knowledge is important for the remediation of a threat. Lastly, when infected hosts are identified, incident responders can ask Stealthwatch what other hosts or machines were in a relationship with the infected party to help ensure that no backdoor was created when the threat actor was operating within the window of compromise.

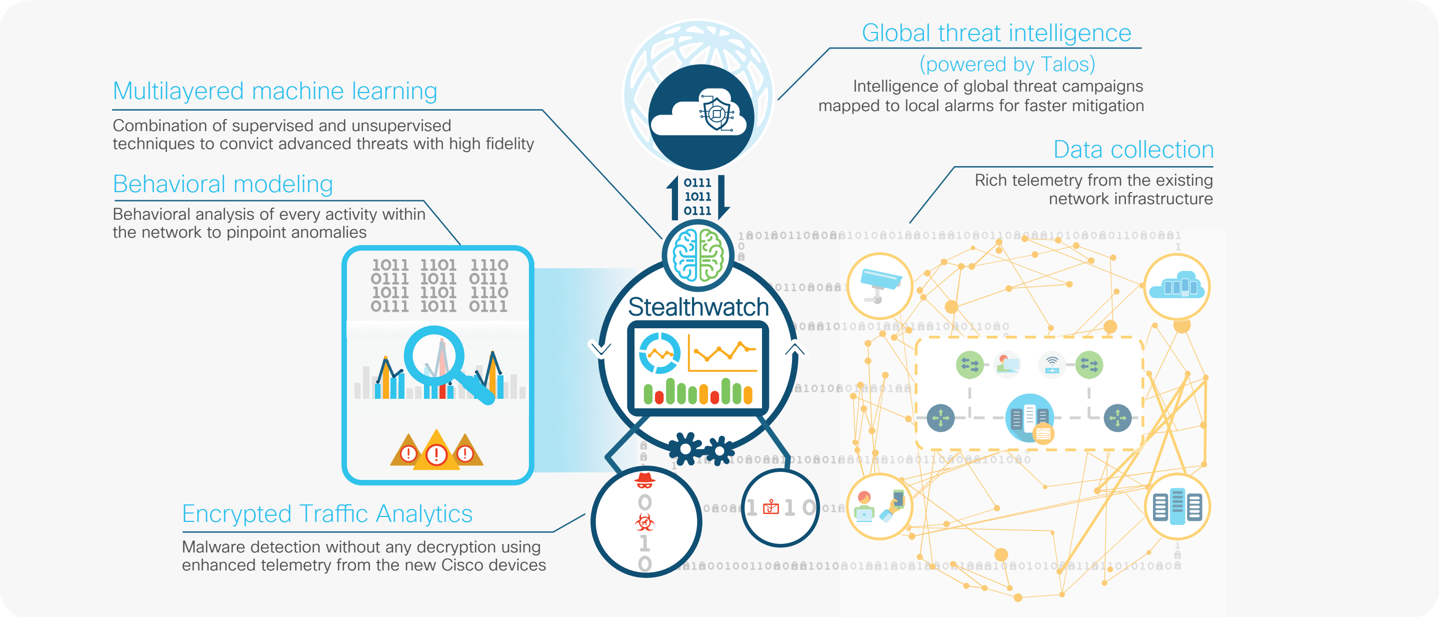
Setting and monitoring segmentation policies

When you have visibility across how the digital business operates, you can create smart segmentation policies to control access to critical resources. This ability is very important to prevent threats from spreading and creating a significant impact. However, it is also important to make sure that the policies being set do not disrupt critical business workflows. Therefore, Stealthwatch provides the ability to model the proposed policies before setting them using firewall or software-defined networking overlays, for example. Contextual alarms are triggered in case the set policies are violated. These alarms can then be investigated, and after the business impact has been evaluated, the policies can be set using the firewall or software-defined networking overlays. You can also verify that the set policies are being enforced effectively through the same alarms.

VI. Summary

Effective network security analytics is not a function of applying just one technique. To stay ahead of evolving threats, a network visibility and analytics solution needs to be able to use a combination of methods. This begins by collecting the right data for comprehensive visibility and using analytical techniques such as behavioral modeling and machine learning. All this is supplemented by global threat intelligence that is aware of the malicious campaigns and maps the suspicious behavior to an identified threat for increased fidelity of detection. And Stealthwatch employs each one of these methods to help organizations identify and thwart attacks that might have crossed the perimeter, or even threats originating within or hiding in encrypted traffic.

Figure 4. Multiple analytical approaches within Cisco Stealthwatch working together to improve security



Global threat intelligence
(powered by Talos)
Intelligence of global threat campaigns mapped to local alarms for faster mitigation

Multilayered machine learning
Combination of supervised and unsupervised techniques to convict advanced threats with high fidelity

Behavioral modeling
Behavioral analysis of every activity within the network to pinpoint anomalies

Data collection
Rich telemetry from the existing network infrastructure

Encrypted Traffic Analytics
Malware detection without any decryption using enhanced telemetry from the new Cisco devices

VII. Learn more

- Product and solution pages:
  - Cisco Stealthwatch Enterprise
  - Cisco Stealthwatch Cloud
  - Encrypted Traffic Analytics
  - Free visibility assessment

VII. Resources

Basic

- Blog: Detecting Encrypted Malware Traffic (Without Decryption)
- Blog: Cognitive Research: Learning Detectors of Malicious Network Traffic
- Blog: Cognitive Threat Analytics: Transparency in Advanced Threat Research
- Blog: Cognitive Threat Analytics: Turn Your Proxy Into Security Device
- Blog: Securing Encrypted Traffic on a Global Scale
- Blog: Closing One Learning Loop: Using Decision Forests to Detect Advanced Threats

Advanced

- Anderson, B., and McGrew, D. (2016). Identifying encrypted malware traffic with contextual flow data. AISec '16
- Grill, M., Pevny, T., and Rehak, M. (2017). Reducing false positives of network anomaly detection by local adaptive multivariate smoothing. Journal of Computer and System Sciences. 83(1):43–57
- Komarek, T., and Somol, P. (2017). End-node fingerprinting for malware detection on HTTPS data. In Proceedings of the 12th International Conference on Availability, Reliability, and Security (p. 77). ACM
- Jusko, J., Rehak, M., Stiborek, J., Kohout, J., and Pevny, T. (2016). Using behavioral similarity for botnet command-and-control discovery. IEEE Intelligent Systems. 31(5):16–22
- Bartos, K., and Rehak, M. (2015). IFS: intelligent flow sampling for network security–an adaptive approach. International Journal of Network Management. 25(5):263–282
- Letal, V., Pevny, T., Smidl, V., and Somol, P. (2015). Finding new malicious domains using variational Bayes on large-scale computer network data. In NIPS 2015 Workshop: Advances in Approximate Bayesian Inference (pp. 1–10)
- Rehak, M., Pechoucek, M., Grill, M., Stiborek, J., Bartos, K., and Celeda, P. (2009). Adaptive multiagent system for network traffic monitoring. IEEE Intelligent Systems. 24(3)