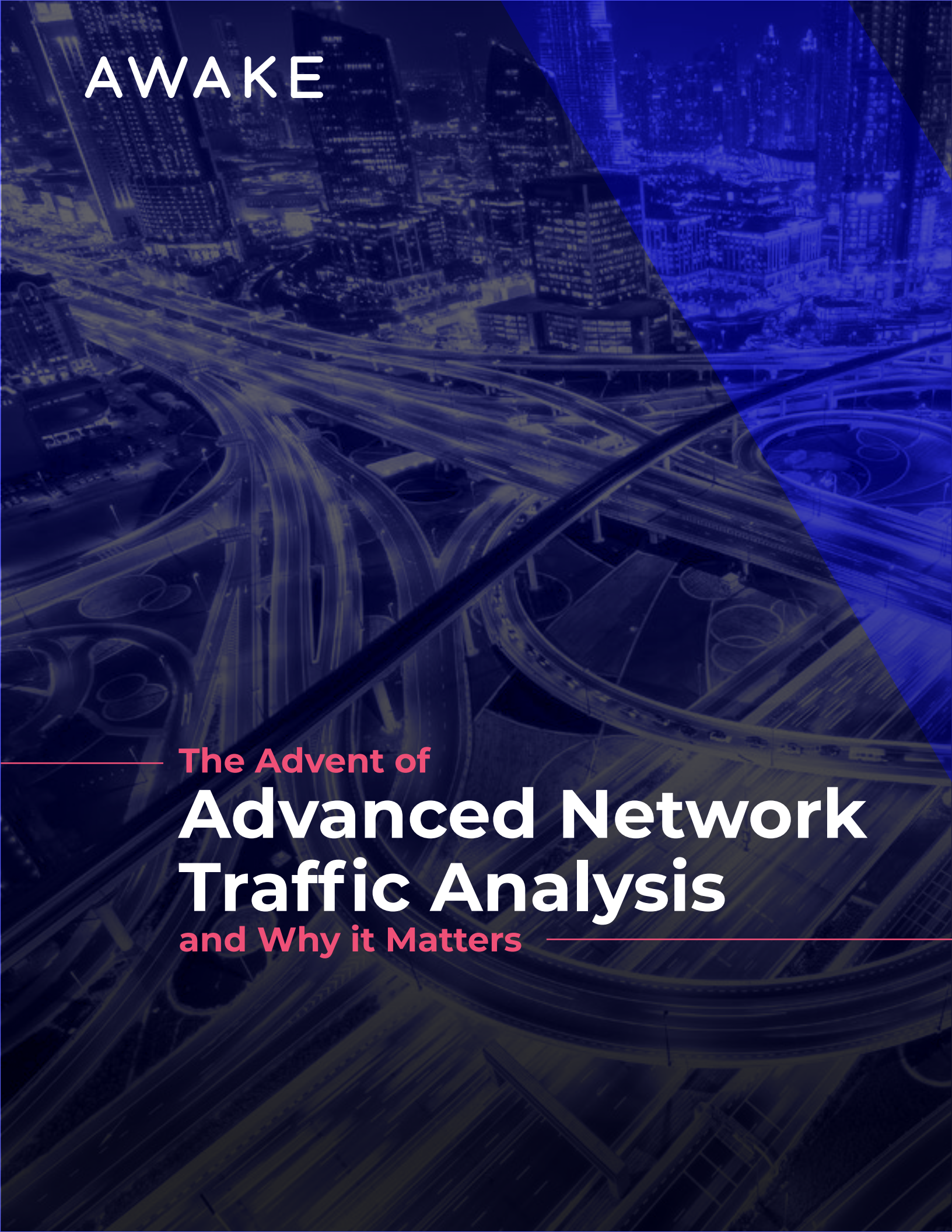


AWAKE

The Advent of

Advanced Network Traffic Analysis

and Why it Matters

AWAKE

The Advent of Advanced Network Traffic Analysis and Why it Matters

Introduction

Over the last few years, so many of the breaches have shown that a prevention-only, perimeter-focused security approach is simply not going to be enough for an organization looking to manage their risk. The attacks have evolved to be long running and the traditional point in time prevention simply hasn't kept up. In fact, industry analysts point out[1][2], that detection and response are now a top priority for organizations. This trend has been accompanied by the embrace of cloud computing, new DevOps processes, the sprawl of devices through IoT, and other such efforts that make the lack of visibility a key impediment in effective detection and response.

To illustrate this changing dynamic, how would you detect an attacker that uses SMB to enumerate file shares, then logs in and executes processes remotely? Or how quickly would you know if someone is tapping into the IP phone system to record and steal voice calls? The answer lies in the network, which sees everything and offers a ground-truth reality that other data sources can struggle with. Rapid network detection and response ultimately lowers the amount of time an attacker operates in the network and thus minimizes the impact. To be effective, however, it has to evolve so customers can protect their organizations even as the very definition of the network itself is changing.

Gartner Says
Detection and Response is Top Security Priority for Organizations

> "The shift to detection and response approaches spans people, process and technology elements and will drive a majority of security market growth over the next five years... While this does not mean that prevention is unimportant or that chief information security officers (CISOs) are giving up on preventing security incidents, it sends a clear message that prevention is futile unless it is tied into a detection and response capability."
>
> **- Sid Deshpande, Principal Research Analyst at Gartner**

The Evolution of Network Security

First Came IDS

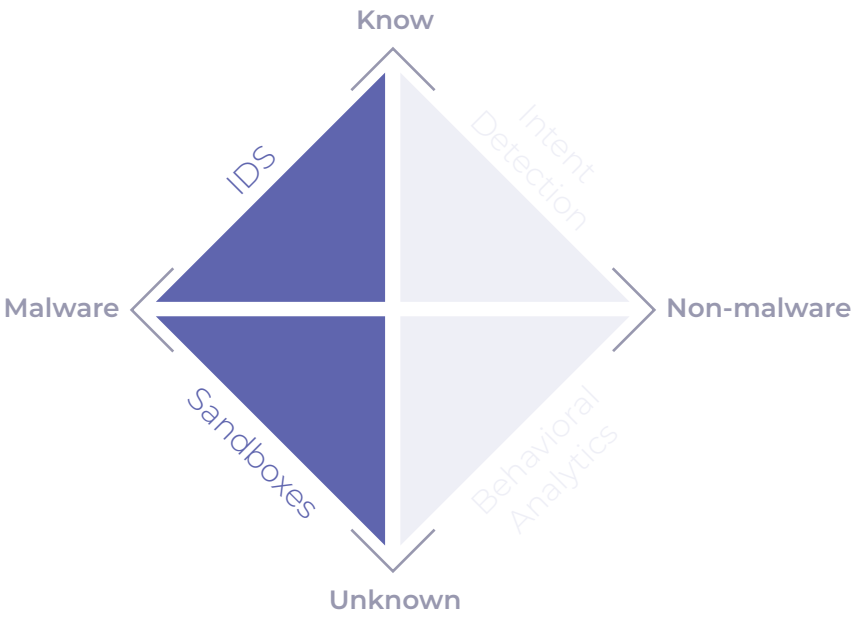
Before we talk about where we are today and where we are headed, it is helpful to establish a framework by which we can understand these evolutions. The premise of network intrusion detection systems (IDS) was detecting known malware by looking at individual packets or sessions for signatures of the malware.

Clearly that model was fraught with challenges including:

[1] https://www.gartner.com/en/newsroom/press-releases/2017-03-14-gartner-says-detection-and-response-is-top-security-priority-for-organizations-in-2017
[2] https://www.gartner.com/en/newsroom/press-releases/2018-08-15-gartner-forecasts-worldwide-information-security-spending-to-exceed-124-billion-in-2019

Then There Were Sandboxes

Attackers were able to bypass IDS detection by making simple, yet subtle changes to the underlying malware or exploiting so-called 0-day vulnerabilities. This brought the era of unknown malware and a fundamental new challenge: without prior knowledge of the vulnerability being exploited, how does one build a signature? Cue: sandboxes. These used a combination of static and dynamic analyses to determine if something is likely malicious. That worked for a while and dealt with some of the IDS challenges.



But attackers never stop innovating. As security teams improved their ability to block malware, attackers changed tactics and stopped using malware as much in their attacks. The attacks also grew to be multi-step and long drawn, often spanning days if not weeks or months and point in time analysis of packets or files just doesn't cut it. Attackers now increasingly focus on the people in the target organization, stealing their legitimate credentials and then using them with the tools that are already deployed in the environment—think scripting languages like Python or system utilities like PowerShell, WMI or psexec and even productivity tools like Microsoft Word or Excel.

Network Traffic Analysis: Behavioral Analytics

The security industry responded to these evolved attacks with network traffic analysis (NTA)—shifting the approach from identifying the 'known bad' to establishing a baseline of what is 'normal or good' and then detecting anomalies from that baseline as 'potentially bad.' This approach has been heavily marketed with words like "data science," "machine learning," and "artificial intelligence"—with these techniques being applied, essentially, to network activity information at layers 3 and 4, for instance, IP address, ports and protocols. While these are great techniques, they are ultimately a means to an end, and in this case the end is the ability to spot anomalies such as, "This IP normally does not see connections from China. Alert if any such connection occurs."

Behavioral analytics and anomaly detection run into some significant challenges:

Training
The algorithms need to understand what is "normal". That takes time—often, 30 to 90 days, which can be frustrating when you are trying to evaluate the technology or deploy it in your environment. But even worse: what if the environment is already compromised? Then the attacker's behavior becomes part of the baseline! *(Oh yes, we always exfiltrate data to random servers...)*

Weak Attribution
The baselining in most cases also attributes behaviors to layer 3 IP addresses. However, IP addresses change often in any modern network: in the course of a single day, a given device may have multiple IP addresses, and a unique IP address may be assigned to multiple devices. If the system alerts based on IP addresses, it will mix together behaviors from multiple devices, and fail to track and characterize the behaviors of actual devices and users that move across IPs. If, on the other hand, it relies on MAC addresses, it loses visibility beyond the first network hop. So how can you truly know what is anomalous for a given real-world entity like a device or user?

Behaviors Change All The Time
Modern networks are dynamic with new behaviors legitimately appearing and others going away. For instance, consider what happens when the baseline is established and then IT rolls out a new backup solution. Every system in the network is now exhibiting an anomaly from the baseline since it is uploading data to a previously unknown location. This leads to the false positives and the need to retrain.

Ultimately while NTA solutions have shown promise, security teams have sometimes struggled to both see ongoing value, and to keep operational costs of tuning the solution to a minimum.

Advanced Network Traffic Analysis

The last decade has seen tremendous innovation in the areas of network processing, data analytics and security research. A new category of solutions is now emerging that take advantage of those advances to specifically solve the shortcomings we described above. Importantly, they also add capabilities that help you combat the threats we are faced with today. Advanced network traffic analysis solutions go beyond the first-generation behavioral analytics solutions in several ways.

Data Sources

Firstly, what is the network?

In a borderless world where workloads are moving to the cloud and remote workers are often the rule rather than the exception, the definition of the network is changing. Advanced NTA solutions are focused on analyzing communications whether those are traditional TCP/IP style packets, "virtual network traffic" crossing a vswitch, traffic from and within cloud workloads and API calls to SaaS applications or serverless computing instances. These solutions also focus on operational technology networks that are often otherwise completely invisible to the security team. This broad visibility ensures security is not a hindrance to broader automated and connected workplace initiatives.

Full packet analysis vs. just meta-data

The first generation of NTA solutions primarily processed layer 3 and 4 meta data like protocol headers or netflow information. Why? Because full packet data through layer 7 is significantly more voluminous and harder to process in real-time. But that volume also means much more signal—signal that is useful to improve detection fidelity, track entities and, perhaps ironically, actually help the solution scale to large and complex networks. For instance, signals like the user information available in Kerberos packets wouldn't show up in protocol headers or netflow. Full packets also give you the ability to understand and store the activity record and then go back in time and retrospectively detect behaviors that may have not been recognized as being malicious when they first occurred. This activity record is also significantly smaller in storage footprint than the full packet data.

Activities Vs. Meta Data An Analogy

If you were in law enforcement and monitoring criminal communications, looking for registrations of new phone numbers is one way to do that. Needless to say, this is likely to be highly ineffective—noisy and easy to miss a real threat. An improvement is monitoring phone logs. They tell you who is talking to who, for how long, and so on. So, if you have a target for your investigation, you can sort through the "meta data" phone records and attempt to differentiate between calls to grandma vs. those to a criminal accomplice. What if instead you were automatically alerted when an actual criminal conversation occurred based on what was said; or if you knew instantly when a new burner phone number showed up on the network?

Breaking Visibility Barriers

Consequential Artifacts

Consider this scenario: say you have lateral movement between two hosts in the Kansas City office. One way to see that activity is to tap the network between those two devices. But in any network except the smallest, that is impractical. However, many communications result in network artifacts that are produced as a side effect. In the lateral movement example, this could be the Kerberos ticket that is issued from the domain controller for one device to access the other. Observing and deeply parsing these has the advantage that NTA sensors only need to be deployed in a relatively limited number of locations while still providing broad visibility. In this case, the Kerberos ticket "consequential artifact" provides evidence of the lateral movement without needing to witness the communication first hand. It is also important to note that in most cases the attacker has no control over these consequential artifacts.

Encrypted traffic analysis

The increasing use of encryption is often discussed as the primary reason network traffic is not a viable long-term source for visibility and detections. Advanced NTA offerings are embracing this challenge. Encrypted traffic analysis allows analysts to uncover threats by analyzing the full payload without actually peeking into it[3]. This avoids the policy and privacy implications of decryption as well as the technical challenges of doing so in a TLS 1.3 world[4].

[3] https://awakesecurity.com/webinars/ja3-reasons-to-rethink-your-encrypted-traffic-analysis-strategies-webinar/
[4] https://www.zdnet.com/article/snooping-on-https-is-about-to-get-harder-tls-1-3-internet-encryption-wins-approval/

Smarter Data Science

Entity Tracking

One of the challenges with the first generation of behavioral analytics solutions was that behaviors were attributed to IP addresses that are ephemeral. With the deeper signal available in full packet data, advanced NTA solutions can not only track, but also profile the entities on the network—the devices, the users, the applications and the destinations among others. The machine learning and analytics can then attribute the behaviors and importantly the relationships to these named entities. Clearly this entity view is a lot more meaningful to human analysts than a list of IP addresses.

A Better Baseline

Most IT environments are changing constantly, for fully legitimate reasons. Without constant retraining, first-generation NTA will produce false positives when changes happen, but retraining is cumbersome and leaves you blind until it is completed. A better approach advanced NTA solutions are adopting is to track behaviors that are unique to an entity or a small number of entities when compared to the bulk of the entities in the environment. Compared to the traditional temporal statistical models, which baseline to past behavior, these models are easier to learn, as the underlying data is available immediately instead of waiting for weeks or months to build the model. In addition, the baseline evolves in real-time as behaviors change across the enterprise. And, given entity tracking, the baseline also benefits greatly from an understanding of the source and the destination entities, in addition to the traffic patterns. This better baseline thus avoids identifying anomalies when IT rolls out a new backup system, as in our example above.

Broader Use Cases

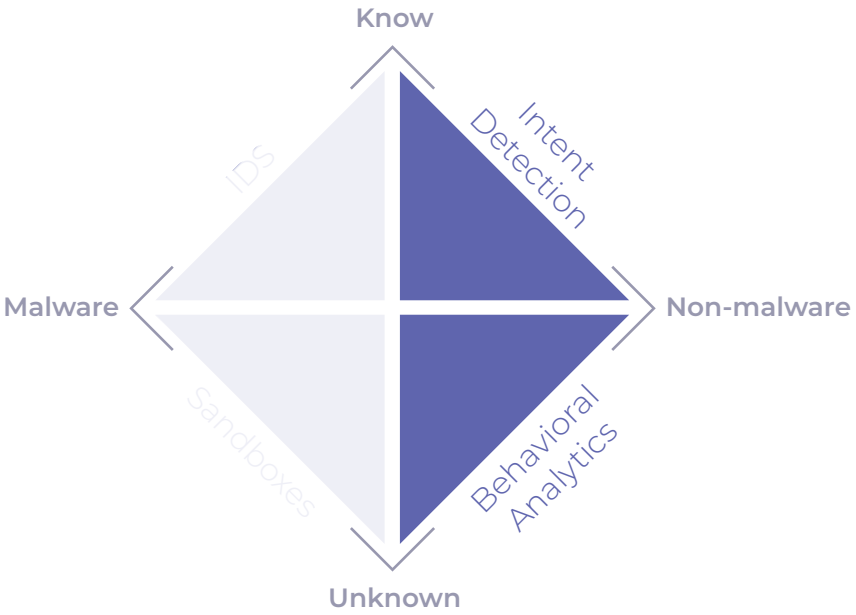
Detection and Response vs. Just Detection

Because advanced NTA solutions attribute behaviors to entities, a wealth of rich context is available not just for detection but also to enable investigation and response workflows. Analysts no longer must look through multiple data sources such as DHCP and DNS logs, configuration management databases and directory service infrastructure to piece the dots together. This is especially powerful for unmanaged devices which most often would not even show up in those databases to begin with. In the age of IoT, cloud and shadow IT, where two out of three organizations struggle with comprehensive visibility[5], security efficacy is heavily dependent not only on how quickly you detect something but also how quickly and decisively you track it down, determine root cause and react.

Intent Detection

What all these capabilities combine to deliver is something we call malicious intent detection. For instance, if you see PowerShell being used to connect to Twitter and that behavior appears on a regular basis, it is highly likely to be some kind of command and control channel.

In fact, this is typical of modern attacks, where the adversary uses not malware but some preexisting tool like Power-Shell with malicious intent. Today, only the most sophisticated security teams are able to find this kind of activity and they do so primarily through manual hunting. As the name suggests, this involves pulling a thread until you hit a dead-end or find something and then repeating the process, again and again.



This is time consuming, requires a high degree of skill, and is not easily replicable. The fundamental challenge well intentioned security teams run into, is that most analysts can express what they want to look for in spoken word, but struggle to express it in

[5] https://www.helpnetsecurity.com/2018/09/11/internal-dysfunction-security-risk/

a way that can be automated through their security technology stack. For instance, it is one thing to say, "alert me if I see a connection from a country I haven't encountered before." It's a lot harder to express something like, "alert me if anyone connects to this database server and then transfers data 2x or more the historical average volume."

Advanced NTA solutions focus on automating this process and reducing the skills and effort barrier that prevents most organizations from hunting. To do this, they support not just machine learning and behavioral analytics, but also rule-based detection that can look for very specific attacker tactics, techniques and procedures (TTPs). The rules themselves are easy to define and can automatically correlate across entities, time, protocols and other relevant parameters while mapping to a known attacker kill chain or a framework like Mitre's ATT&CK[6]. This allows a researcher or analyst to look for sequences of events over weeks or months.

For instance, you can define a rule looking for "SMB file transfers that start within minutes of a HTTP file download that also match the size of the download." As the examples illustrate, the rules allow the search for high level behaviors (e.g. the use of non-browser clients to access the Internet) and not just low-level primitive data (e.g. port and protocol parameters). This behavioral rule mechanism is thus more accessible to a broader population of analysts while also being more robust and longer lasting than traditional signatures.

With the addition of rule-based detection, advanced NTA solutions can also offer visibility and control to the organization and the security teams within. As Gartner analyst, Avivah Litan stated, "[Machine learning] Vendors can't sell black boxes."[7] Instead, organizations that adopt advanced NTA solutions are able to adapt these platforms to nuances of a particular network without the need for a data science team to modify training sets or algorithms. In addition, this approach allows for custom detection of threats that are organization-specific (e.g. typo squatting of your supplier domains).

Of course, no detection paradigm is perfect. Combining machine learning and mathematical analysis with high-level rule-based approaches allows for higher fidelity detection with low false positives and negatives. It allows you to detect behaviors that are traditionally seen in the north-south direction (threats that manifest themselves in the interactions between your organization and the outside world) as well as the east-west direction (threats that manifest themselves in the interactions between entities within your organization). Said differently, intent detection applies just the same whether the attacker is an insider that is either accidentally or maliciously compromising your security, or if it's an external attacker that has broken into and is now making their way through your environment.

From the security team's perspective, the rich context that is used to detect, is also available to speed up response, eliminating the need for pivots to other systems for investigating the threat. And when a new pattern of maliciousness is discovered, detecting it is simply a matter of creating a new rule and having the system autonomously then hunt for the behavior.

Conclusion

Network data offers ground truth about the behavior of entities that is impossible to replicate through logs or end-point agents. Attackers can disable agents, delete their traces from logs or file systems, but they cannot "unsend" a packet and they cannot avoid the consequential artifacts that result. Properly instrumented, the network also has a memory, allowing for both real-time and retrospective detection. Recent advances in network processing, analytics and security research have enabled a new era of detection and response capabilities that eliminate many of the challenges of traditional network security. Advanced network traffic analysis solutions are tapping into the evolving network, from on-premise to cloud, virtual and SaaS to deliver value quickly without long drawn deployments and training / re-training. Every organization would do well to consider these as part of their security architecture.

[6] https://attack.mitre.org/wiki/Main_Page
[7] https://blogs.gartner.com/avivah-litan/2017/07/27/can-we-trust-black-box-machine-learning-when-it-comes-to-security-or-is-there-a-better-way/

For additional information about Awake please visit awakesecurity.com

About Awake Security

Awake Security is the only advanced network traffic analysis company that delivers a privacy-aware solution capable of detecting and visualizing behavioral, mal-intent and compliance incidents with full forensics context. Powered by Ava, Awake's security expert system, the Awake Security Platform combines federated machine learning, threat intelligence and human expertise. The platform analyzes billions of communications to autonomously discover, profile and classify every device, user and application on any network. Through automated hunting and investigation, Awake uncovers malicious intent from insiders and external attackers alike. The company is ranked #1 for time to value because of its frictionless approach that delivers answers rather than alerts.