Cybersecurity

Security analytics use cases for threat hunting

Trusted partner for your **Digital Journey**

Atos

Content

Introduction

Security analytics can be a valuable tool for detecting advanced attacks. However, it must be applied correctly. Too often, the goal of security analytics is reduced to the construction of an AI-driven big data platform, running data science algorithms, machine learning, or statistical packages.

Instead, the starting point should be to identify the risks that cannot be monitored through conventional security products and then to define use cases in security analytics to monitor those risks.

In this paper, we discuss the need for security analytics and how to apply it in a meaningful way within a Managed Detection and Response (MDR) service to achieve results. We then discuss the technology components required to put security analytics in action.

The context for security analytics

Threat actors have evolved. They now employ types of attacks that cannot be detected using signature matching or predefined rules.

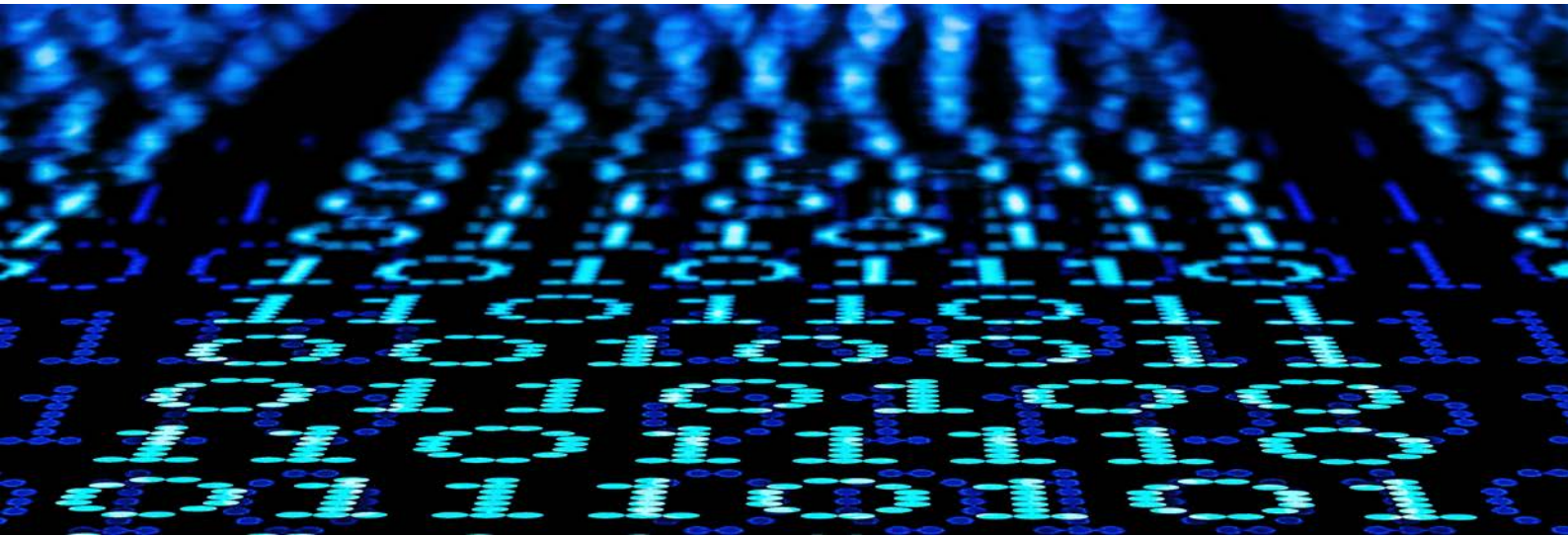
The threat landscape can be mapped using a grid with two dimensions: attacks (whether known or unknown) and attackers (known or unknown). Attacks that are known can be detected using the rule-matching technology of antivirus (AV) software, intrusion prevention systems (IPS), web application firewalls (WAF), data loss prevention (DLP), and security information and event management (SIEM). When attackers are also known, their attributes (IP addresses, URLs, files) can further prioritize the attacks.

When attackers are known through threat intelligence feeds, and the attack methods are not known, a big data platform can be used to collect Netflow, proxy and user access data, and match this against known malicious IP address or other known indicators of compromise (IOC) related to the attackers. Threat intelligence is also used for carrying out threat hunting activities within the network. Both of these approaches need specialized tools and may be grouped under the broad umbrella of security analytics. Signature, rules, and threat intelligence fail when detecting unknown attacks from unknown threat actors. Undetected attackers can stay within the network of an organization for longer, navigating towards critical or valuable assets. Their attacks usually fall into the category of advanced persistent threats (APT) or advanced targeted attacks (ATA). FireEye Mandiant report (M-Trends 2020: A View from the Front Lines) states that it takes a median time of 60 days to detect a threat.The damage is often high by the time the attack is detected. Reducing the time to detect such attacks can considerably reduce the impact of breaches.

The critical need for security analytics is for detecting such unknown attacks. By comparison, although incremental value might be gained in applying security analytics in the other three quadrants of the grid, the benefit to effort ratio is small. As a result, whenever security analytics is considered for attack detection, the acid test is to see if such attacks can be detected through rules, because the attack, the attacker, or both are known. If a rule-based approach is insufficient, next step is defining the appropriate use case(s) to apply security analytics.
We will illustrate this approach with a few examples of use cases in security analytics.

"Reducing the time to detect such attacks can considerably reduce the impact from breaches."

"The acid test is to see if attacks can be detected through rules. If not, the next step is to define use cases to apply security analytics."



Use cases

1 Detection of malware beaconing activity

Security analytics can detect beaconing activity from unknown malware and from that detect the unknown malware itself. This applies to malware used in APTs and ATAs. A rule-based system such as SIEM or a signature-based system including IPS, anti-malware, and WAF cannot detect such malware attacks.
Even a sandbox technology approach for detecting malware will fail, because today's malware stops executing when a virtual environment is detected.

Before deciding to use security analytics, we should recognize two other ways to identify unknown malware. One of them is to use external threat feeds. Most malware sends out regular heart beat information to its command-and-control (C&C) server. Using external threat intelligence, these communications can be identified and flagged to security investigators. The other technique is to look for a fixed pattern of beaconing: for instance, data packets of a specific size are sent at a certain frequency. If the size and frequency are known, such rules can be modeled in a SIEM. Advanced attackers use command-and-control (C&C) servers that are not yet on any threat feed. They also use beaconing tactics that do not follow a known pattern of size and frequency. This is the use case we need the AI-driven security analytics to solve - advanced malware with no known C&C server and beaconing pattern.

There are nonetheless traces of malware activity in different sources in the IT environment. One such source is a proxy. It will contain the heart beat traffic, even though that traffic cannot be detected by rules. Using analytics, this heart beat can be detected by applying entropy techniques. Entropy in data science terms refers to "uncertainty of data". When we look at proxy data in general, the data related to user interaction with URLs is expected to be randomly distributed in terms of the interaction size. After all, people visit a variety of websites and upload and download a variety of data. Therefore, data sizes are expected to be highly variable. In this case, when we examine the entropy of the byte size of the communication between a user and a URL, the entropy and "uncertainty of data" should be high.

The heart beat information being beaconed out by many malware is characterized by similarity and regularity, even though they are not known for us to build rules. If we apply an entropy function to this data, the entropy will be very low, since the byte size and frequency are relatively uniform when interacting with the C&C URLs. This enables us to detect the unknown attack, even though the attack signature or attacker is unknown

"Advanced malware with no known C&C server and beaconing pattern must be solved by AI-driven security analytics."

"AI-driven security analytics enables us to detect cyber attacks for which the attack signature or attacker is unknown."

2 Detection of a watering hole attack

A 'watering hole' attack is used to infect hosts by luring users to a location (URL/IP) where the malicious code is hosted. In a similar way to the previous use case, if the URL hosting malicious code is represented in blacklists of threat feeds, this is not a use case for security analytics. It can be detected using SIEM with external threat feed integration.
Also, if the file size or file pattern is known, it can be spotted by writing specific rules based on IPS signatures, URL filters, and sandboxing for any one of a number of security products.

On the other hand, we need security analytics to detect infection through water holing where the file size, URL, and file pattern are all unknowns.
Detection of a waterhole attack requires three threat hunting steps. First, outliers in URL access data must be identified, based on the number and frequency of accesses over the past few days. Second, an entropy function is applied to the file size/download content to identify low entropy URLs. This generates a repository of all potentially compromised hosts and likely water holing URLs they have visited. As the third and final step, this data is converted into a tree-map to help security analysts quickly visualize the hosts and URLs, and investigate the tree-map's larger nodes.

"Detection of a waterhole attack requires three threat hunting steps - isolate access data outliers, identify low entropy URLs, and visualize the results to take action."

Use cases

3 Data exfiltration

The application of security analytics to data exfiltration makes it possible to detect new data leakage scenarios beyond those identified by standard data loss prevention (DLP) solutions. The standard solutions can detect leakage if we know the data or the pattern of the data we want to protect: for instance, card information, specific data fields, or file signatures. However, when we want to detect data leakage instances beyond known data or when the communication is encrypted, we need security analytics.

As an example, HTTP POST is a common method used to upload files. Such uploads could be valid business requirements. However, they could also be malicious data exfiltration by malware or a rogue insider. Detecting such illicit exfiltration is possible by base lining HTTP POST traffic from each source system in an organization and then detecting outliers for this traffic. This enables us to detect any abnormal traffic movement from any system in the enterprise, and mitigate it. In an AI driven machine learning system that builds on itself, baselines are created and updated over a period of time. There are many other channels (e.g. email) to which similar analytical techniques can be applied to detect data exfiltration.

The above three use cases are samples. There are many more use cases that can be addressed using security analytics. **Speak to an Atos Security Professional for more.**

"An AI-driven machine learning system, which builds on itself, detects data exfiltration."

Use case best fit

Overall, we can classify use cases in three broad categories: real-time rule-based use cases, real-time security analytics use cases, and batch security analytics use cases.

Real-time rule-based use cases – Use cases for attacks or attackers that are known and that do not need to be compared with past attack history. They can be defined and detected using rule-based approaches such as SIEM, IPS, WAF, and DLP. As an example, an attack originating from a blacklisted IP address corresponds to a simple rule for matching the event's source IP address with the available blacklist or global threat database. Similarly, rules for known attacks are signature rules in SIEM, IPS, WAF, and other rule-based systems. Known indicators, such as a high number of login failure attempts within a short time, can also be defined in rule-based systems. Compliance related measures including logins after office hours, logins from suspicious geographies, and access using non-standard remote clients can all be configured as use cases in a SIEM system.

Real time security analytics use cases – Use cases for triaging incoming alerts from other real-time systems (SIEM, IPS, WAF, etc.) or using machine learning to match a pattern that needs extended periods for detection. As an example, a SIEM alert can be better prioritized by quickly assessing the attacker's IP address in real-time for the past volume of attacks from that IP address, the severity of those attacks, other destinations targeted by the same address, and user parameters including vacation information. Other similar parameters that can also be used to ascertain if the event is worth the effort of remediation. All of this is achievable using AI-driven real-time security analytics to leverage statistical models and rapid searches of large datasets.

Batch security analytics use cases – Unknown attacks and attackers are best handled in batch or near real-time analytics use cases. In these cases, detection involves using deeper statistical models and profiling large data sets. These are periodic threat hunting jobs that run on data to produce output that is then visualized. The time to process this type of output is usually a function of the type and quantity of hardware deployed. Accordingly, the periodicity of the analytical models used may range from a day to less than an hour. The use cases discussed above for malware beaconing, watering hole attacks, and data exfiltration fall into this category.

"Unknown attacks and attackers are best detected in batch or near real-time analytics use cases, with deeper statistical models and large data sets."

Security analytics in threat hunting

No matter the specific use case, there is one activity where security analytics are required today—Threat Hunting. Atos Threat Hunting service deploys a streamlined AI platform that combines network, end-point, application, and user behavior threat analytics to uncover threats and attack campaigns that traditional security monitoring mechanisms often miss. Threats within each of these analytics channels are complex, hard to detect, and rarely conform to known attack patterns or signatures. Furthermore, modern threats often operate across more than one of these channels, and potentially anomalous behavior within one must not only be detected, but also correlated against behavior in others. Due to the complexity of this task, the massive volume of data involved, and the unknown nature of modern threats, only AI combined with Threat Hunters can effectively hunt for and uncover threats within your data pools.

Security analytics architecture

The architecture must support the functionality that we have discussed above. After data input using technology such as Flume or Logstash, real-time threat hunting analytics on large data sets can be performed with Spark and Spark streaming. Since datasets are large, they will need to be stored in suitable data stores such as Hive and HBase. Solr can support the fast search feature on datasets to match patterns. Batch analytics that require deeper statistical models can be performed using statistical packages such as R and MLib. Interactive visual querying can be achieved using D3 (Data-Driven Documents) functions. The components described here are to illustrate the possibilities. Other similar products can also be used to build the big data analytics platform that drives the MDR program.

Summary

Security analytics needs a clear definition of use cases before designing the platform and data science packages.

A suitable method for defining use cases is to identify the risks that an organization wants to monitor and then find the gaps in existing rule-based systems. These gaps can be addressed using data science techniques. Some of the techniques must be executed on a real-time platform. Others are better served by a batch processing platform.

Atos AI-driven MDR service

Atos MDR service combines Artificial Intelligence (AI) techniques & machine learning with skilled security experts to provide high-speed cyberdefense. You get end-to-end threat management. Our left-to-right-of-hack services include:

Threat Anticipation
We collect, analyze, and proactively respond to global threats.

Threat Hunting
We combine machine learning models with expert hunters to detect hidden threats.

Security Monitoring
We combine security rules with 24x7 SOC analysts to alert on known threats.

Incident Analysis
We investigate threats for attack chain, impact, and threat actors using forensic automation and skilled analysts.

Auto Containment
Contain and recover swiftly with agile response from machine learning.

Incident Response
We execute playbooks via response automation and expert responders.

In total, our AI-driven MDR service quickly contains breaches and reduces their damage. Learn more about our AI-driven MDR service by **clicking here.**

About Atos

Atos is a global leader in digital transformation with 110,000 employees and annual revenue of € 12 billion. European number one in cybersecurity, cloud and high performance computing, the group provides tailored end-to-end solutions for all industries in 73 countries. A pioneer in decarbonization services and products, Atos is committed to a secure and decarbonized digital for its clients. Atos operates under the brands Atos and Atos|Syntel. Atos is a SE (Societas Europaea), listed on the CAC40 Paris stock index.

The purpose of Atos is to help design the future of the information space. Its expertise and services support the development of knowledge, education and research in a multicultural approach and contribute to the development of scientific and technological excellence. Across the world, the group enables its customers and employees, and members of societies at large to live, work and develop sustainably, in a safe and secure information space.

Find out more about us
atos.net
atos.net/careers

Atos Global Head Office
River Ouest, 80 quai Voltaire
95877 Bezons cedex – France
+33 1 73 26 00 00

Let's start a discussion together

For more information: **contact us here.**

CT-210105-AR-MDR-WP-Threat-Hunting-EN-V2