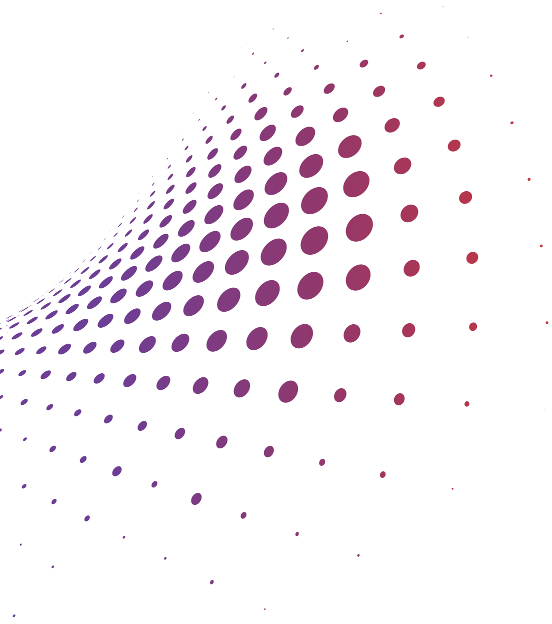
CISO TECH BRIEF

How Anomaly Detection Advances Help Companies, Not Just Detect, But Prevent Data Breaches



In the threat landscape of today, we have various data security solutions ringing alarm bells with varying degrees of severity, yet little to no direction as to where to prioritize. The big question we must answer is this: Are we looking at the users' and machines' behaviors and understanding when they deviate from normal behavior? That's where anomaly detection becomes a critical part of your data security stack.

Anomaly detection is an important technique for identifying fraudulent or suspicious activity, network intrusion, and other abnormal events that may have great significance, but are difficult to detect. Anomaly detection is critical in translating data into critical actionable information and provides useful insights in a variety of application domains.

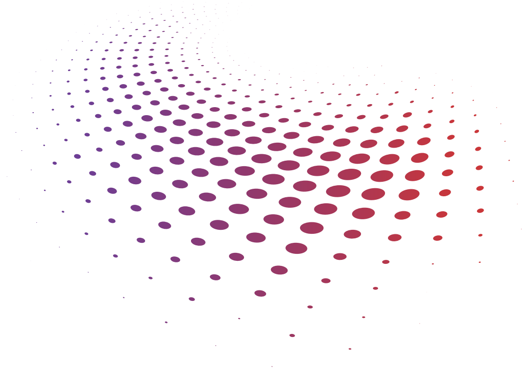
The bottom line is that anomaly detection hasn't lived up to the hype. Broad adoption has been hampered for several reasons, three of which are: 1) the lack of productization of the technology, 2) a clear definition of use cases and actionable insights, and 3) poor performance due to flawed implementations and a lack of knowledgeable resources in house.

Anomaly detection has presented a range of challenges for data security professionals. In this paper, we outline some of the key challenges, along with two significant advances in anomaly detection that are helping organizations overcome those challenges today, to more effectively detect and prevent data breaches.

> **Anomaly detection is the identification of rare events, items, or observations which are suspicious because they differ significantly from standard behaviors or patterns. Anomalies in data are also called standard deviations, outliers, noise, novelties, and exceptions.**

Anomaly Detection Challenges

Implementing an effective anomaly detection solution has not been without its challenges. Here are some of the key challenges that organizations have faced.



1. Building an anomaly detection solution is difficult.

Some organizations contemplate building an anomaly detection solution in house but find it is very difficult because they don’t have the talent in house, nor are they able to hire this type of talent. Therefore few organizations attempt to build an anomaly detection solution from scratch. Additionally, building a solution is not effective due to the following reasons:

- Anomaly detection (AD) systems are either manually built by experts setting thresholds on data, or constructed automatically by learning from the available data through machine learning (ML).
- It is extremely time-consuming to build an anomaly detection system from scratch. This requires an extensive background in ML, AI, neural networks,domain knowledge and—even more difficult to access—foresight.
- Under the lens of chaos engineering, building an anomaly detection solution manually is costly and very difficult to adapt over time.

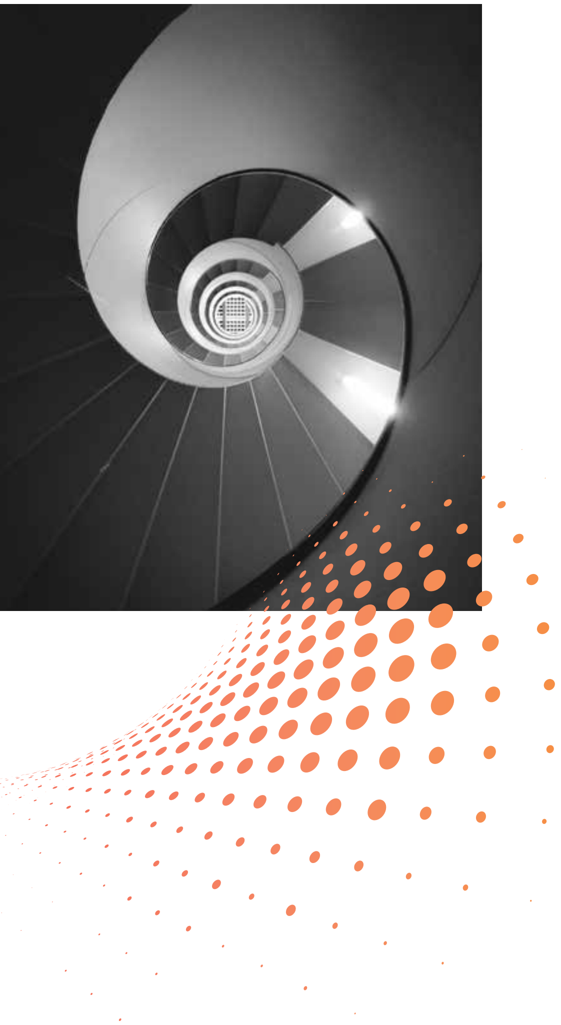
2. Traditional anomaly detection solutions come with two major limitations

Today’s anomaly detection solutions are typically deployed at the firewall or network level, rather than at the data access level. This prevents them from detecting data requests that are benign at the access level but still malicious at the data level.

Secondly, log file and user behavior analysis tools, such as Splunk, do not operate in real-time. They can help organizations discover hacking/intrusion and unauthorized access as part of a forensic investigation, but they do not enable a company to interrupt and prevent unauthorized access in real-time. This means that organizations must still have some type of anomaly detection technology in place to ensure that intruders are recognized and blocked before unauthorized access is granted and a breach occurs. This is where machine learning (ML) and quarantining suspicious behavior is required.

Anomaly Detection Advances Address These Challenges

Two significant advances in anomaly detection are now helping organizations detect and prevent data breaches more effectively.



1. Anomaly Detection at the Data Level

Anomaly detection typically identifies bad actors at the network level. These bad actors have access to the data whether it is a set of files or a database, which makes it nearly impossible to prevent data breaches at the network level. The best analogy is to imagine a house where the front door has been locked, but the back window is still ajar. If a neighbor pretends to be the owner of the house, he will still be able to get past the front door. This is equivalent to gaining privileged user access by pretending to be someone “known’’ or “privileged.” A network attacker is able to get through the cracked window, thereby infiltrating the house.

Today, we can employ anomaly detection at the data level to protect the actual assets that are the targets of malicious actors. By protecting data at the data level rather than the network level, this means that even if an attacker is able to infiltrate the network, he or she is most likely unable to get past the data level. This is equivalent to the key protecting the most sensitive data in the vault where it is being kept safe, therefore, this means it becomes irrelevant whether someone entered through a window or door. No matter where the attacker is located in the house, he or she is unable to get past the vault, where all the sensitive data is stored - at the data level.

2. Real-time Anomaly Detection with Machine Learning

Machine learning (ML) has proven highly effective for advancing anomaly detection accuracy and helping companies and organizations manage big data. Algorithms enable ML systems to learn by their own experience, thus refining their analytical and predictive capacity on their own, is a valuable feature for accurate anomaly detection.

So, what is the advantage of an anomaly detection solution method enriched with ML technology? The first undeniable benefit is the ML system’s ability to handle unlabeled data proactively, determining what is normal and what may be regarded as a data anomaly. Second, ML systems are much more sensitive to distinguishing data anomalies from noise, allowing them to differentiate data units based on the degree of their deviation from the norm.

Real-time anomaly detection combined with machine learning enables organizations to proactively detect malicious attempts to access, use, and steal information. Algorithms that enable pattern detection while the model constantly learns with every signal or event enable organizations with the ability to process large amounts of data while eliminating mistakes. For example, some anomaly detection solutions devise a combined probability score that is frequently referred to as a threat score between 0 and 100. If the anomaly or threat score is high enough, the system has the ability to quarantine or stop requests in real time so there’s no data loss. The threat score can use multiple thresholds across a range of scores to trigger actions. It can allow anomalies to pass through while logging every anomaly, but if an anomaly is deemed above a certain threshold, it can trigger a quarantine action and thereby refuses the execution of the anomalous query.

Some ML algorithms are able to enable the analysis of each data access request and are able to review and categorize them based on their threat potential. Anomaly detection solutions can achieve this by real-time, self-learning ML models that enable threats to not only be detected, but also stopped for execution and quarantining. Typically, anomaly detection solutions create a baseline of what is normal versus anomalous over a period of time by either: 1) having a baseline window where the system trains itself or 2) processing historical logs on which the system can train itself.

Conclusion

As organizations try to protect their sensitive data at all times, regardless of where it is located, they must monitor and understand the user behavior and receive alerts when that user is behaving outside of the norm. Threat actors today watch and learn, but will almost always perform an action that is abnormal to a typical user. With anomaly detection, organizations can be alerted to unusual behavior so security operations (SOC) teams can take immediate remediation actions to block the adversary.

→ **Contact Sotero to learn how an anomaly detection solution can enable your organization to successfully deploy anomaly protection to not only detect, but also prevent data breaches.**

Contact a data security expert today.

ABOUT SOTERO

Sotero is the global innovator and leader in next generation data security. Sotero's data security platform enables our customers with a way to protect data anytime, anywhere, regardless of data store, integration mechanisms, and user tools. The platform is able to control, access, operate, and use data to extract information that drives organizations' business outcomes and innovation. With Sotero, organizations have the ability to proactively detect and prevent malicious attempts to access, use and steal information. It's anomaly detection technology allows organizations complete control over their data's privacy, compliance, auditability, and governance.

Sotero

99 S. Bedford Street, Suite 106
Burlington, MA 01803
www.soterosoft.com