opentext™ | Cybersecurity

Position Paper

What Makes ArcSight Intelligence Different Part 1

As one of many vendors in the cybersecurity space competing for budget based on buzzwords like AI, machine learning, threat hunting, accelerated threat detection, advanced analytics, big data, deep learning, neural networks, and more, we understand your skepticism when you see and hear those exact words from us.

This pair of ArcSight Intelligence position papers demonstrates the key differentiators that customers tell us are unique to us after evaluating dozens of security analytics or user and entity behavioral analytics (UEBA) products.

Table of Contents

Our AI Is Different. Our Approach Is Different

In this paper, you will learn why ArcSight Intelligence's approach to using AI to detect advanced threats is unique and how it improves security operations center (SOC) team efficiency.

Knowing the difference between a mallet, a hammer, and a sledgehammer changes the results, and the same is true for different machine learning and AI methods.

Our AI Is Different: Unsupervised Machine Learning, Unique Normal, and a Useful Risk Score

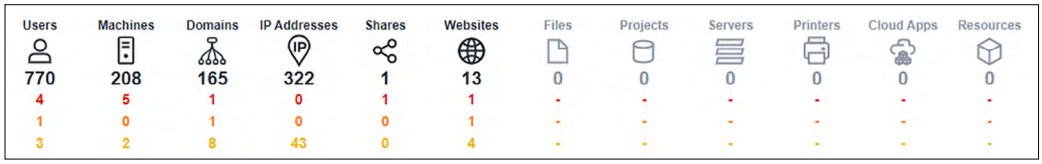
Much of what you heard previously about machine learning in the cybersecurity industry was actually one particular subset of machine learning—supervised learning, or learning by labels. More recently, we've witnessed the introduction of generative AI along with the many hopes, fears and misconceptions that go with it.

Knowing the difference between a mallet, a hammer, and a sledgehammer changes the results, and the same is true for different machine learning and AI methods. Advanced data science teams who have evaluated our machine learning typically respond with, "Wow, that's hard!" This first paper explains those data scientists' reactions by looking at how ArcSight Intelligence uses unsupervised machine learning for measurement of "unique normal" for a meaningful risk score and markedly more effective threat detection.

ArcSight Intelligence Security Analytics Threat Detection Platform

ArcSight Intelligence is unique because it is the only solution that makes extensive use of unsupervised machine learning algorithms to discover new patterns and subsequently find new threats, enabling threat hunters, security practitioners, and security operations center (SOC) teams to effectively measure risk and choose a resource-efficient response.

ArcSight Intelligence's analytics platform is built from the ground up to execute unsupervised machine learning algorithms at enormous scale. These algorithms extract the available entities (individual users, machines, IP addresses, web servers, printers, etc.) from log files and observe events that relate to these entities to determine what is normal or expected behavior. As new information comes through the analytics process, it is evaluated against previously observed behavior, as well as dynamically measured statistical peer groups, to assess potential risk.



Users	Machines	Domains	IP Addresses	Shares	Websites	Files	Projects	Servers	Printers	Cloud Apps	Resources
770	208	165	322	1	13	0	0	0	0	0	0
4	5	1	0	1	1	-	-	-	-	-	-
1	0	1	0	0	1	-	-	-	-	-	-
3	2	8	43	0	4	-	-	-	-	-	-

Figure 1. Entity risk breakdown report

The Importance of Unsupervised Machine Learning

ArcSight Intelligence's most significant differentiator is its extensive use of unsupervised machine learning, which is critical for detecting attacks such as advanced persistent threats and insider incidents. This is a different category of machine learning from supervised machine learning.

ArcSight Intelligence's most significant differentiator is its extensive use of unsupervised machine learning, which is critical for detecting attacks such as advanced persistent threats and insider incidents.

Unsupervised machine learning finds new patterns in large sets of data, no matter how random that data seems. The algorithms collect data and cluster it to find structures and recognize emerging patterns. This type of machine learning is particularly well-suited to insider threat detection because inside(r) threat scenarios typically deal with limited, label-less datasets.

In contrast, deep learning—a type of supervised machine learning—is pretty good at malware detection, primarily because these use cases involve large datasets of labels (decades' worth of malware binaries). Most major antivirus and antimalware vendors tout the use of deep learning for their specific objectives, which is largely why it's become such a prolific topic in our industry.

Generative AI's role in threat detection is emerging. It has the potential to complement existing threat detection techniques by allowing analytical data to be interrogated and explained.

When External Threats Become Inside Threats

The Ponemon Institute* defines three types of insiders: the malicious insider—an employee or contractor who seeks to cause harm, the negligent or mistaken insider—an employee or contractor who inadvertently causes harm, and the credential thief—an outsider with stolen credentials obtained via social engineering. While all insider threats pose a risk, credential theft often leads to more significant damages.

In the case of a credential thief, the attacker can gain control over a legitimate user's credentials and move laterally within the targeted organization. The attacker is not an insider, but once penetrated, and with that level of access, the threat has moved 'inside.' Such an attacker often goes to some lengths to avoid raising alarms by steering clear of any activity not already permitted by IT policy and thus often 'flies under the radar' and, until detected, is in essence also an insider threat.

*Ponemon Institute—Cost of Insider Risks Global Report 2023

Regardless the type of insider, once a threat actor has access to internal systems they are more difficult to detect. This is due to the inherent trust a company must place in their employees to allow work to continue. Balancing security with usability is an ongoing struggle. As trusted employees, malicious insiders easily skirt past identity access management and two-factor authentication challenges and easily "blend in." Detecting these insider threats requires a powerful approach to threat detection.

The Importance of Measuring "Unique Normal"

When it comes to threats, every entity—person, machine, device, etc.—has a unique normal. It also follows that each entity's interaction with another entity is also unique with its own normal set of behaviors. No two threats are the same, making detection of such threats exceedingly difficult. Accurate threat detection requires precise measurement of how a unique entity behaves. This baseline of unique normal can then be compared with itself to see aberrations. Rules and thresholds don't work because they assume the same rules work for every entity, creating many false positives. While more nuanced, relying on models trained on labelled data requires knowledge of novel threats, before they are encountered, allowing zero-day threats to go undetected until uncovered by other means.

Accurate threat detection requires precise measurement of how a unique entity behaves. This baseline of unique normal can then be compared with itself to see aberrations. Rules and thresholds don't work because they assume the same rules work for every entity, creating many false positives.

A solution needs to reduce false positives, which is achieved through an architecture that can scale horizontally to accommodate the measurement of unique normal for thousands of entities.

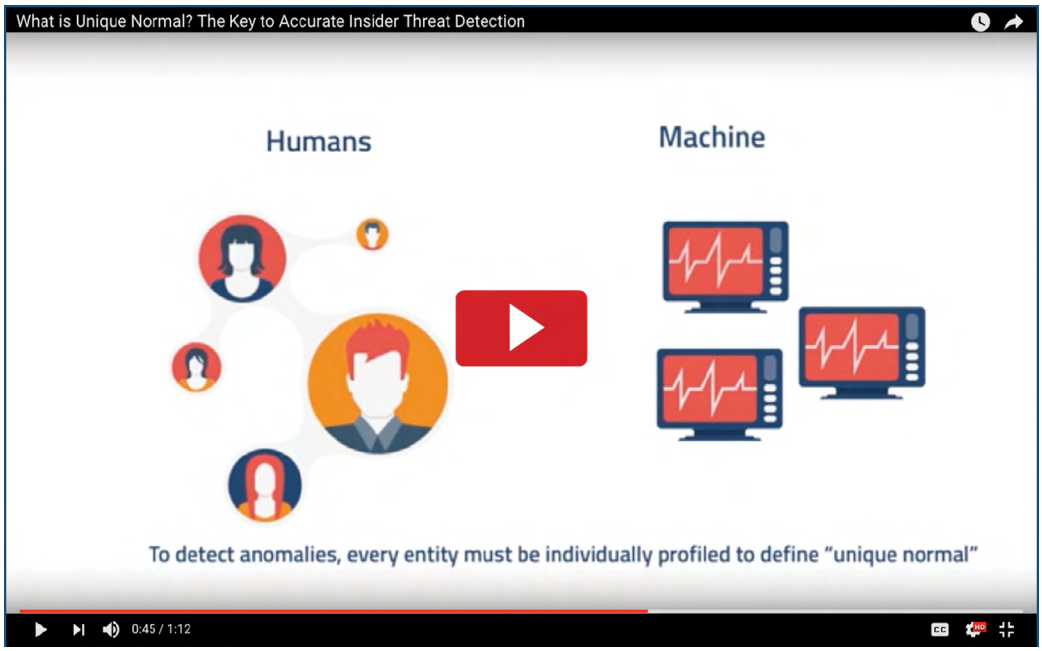


Figure 2. The key to accurate insider threat detection

ArcSight Intelligence creates unique behavioral baselines for every entity and its relationship to every other entity by applying unsupervised machine learning (a type of artificial intelligence) to mathematically discover patterns. The ability to identify unique baselines for peer groups, dynamically and mathematically, further reduces false positives.

ArcSight Intelligence scales horizontally to accommodate measurement of unique normal for millions of entities, and the platform comes with hundreds of built-in analytical models that automatically incorporate "live" data. This automated data analysis, powered by machine learning, enables flexible and effective threat detection as it removes the traditional dependency on rigid rules and thresholds-based systems, which generate many false positives and requires continuous, labor-intensive maintenance.

The Math Driving ArcSight Intelligence's Risk Scores

Because math is applied 24/7 to all data sources coming into the engine, it does not miss any events as a human might. Math's tireless ability to see relationships across huge amounts of data—something humans are not good at—allows it to detect threats that currently go unseen. ArcSight Intelligence leverages machine learning to measure both event risk and entity risk. This is important because one entity (e.g., a user) can participate in multiple anomalous activities (e.g., authentication, access time, data movement). For example, a single erroneous authentication attempt is suspicious but not alarming.

Next, ArcSight Intelligence's threat detection creates a meaningful risk score by compressing all the various anomaly probabilities into a single entity risk score, essentially normalizing each single entity risk score against each other. This gives security practitioners a meaningful, rank-stacked list of threat leads out of millions of entities continuously measured over time.

ArcSight Intelligence's threat detection creates a meaningful risk score by compressing all the various anomaly probabilities into a single entity risk score, essentially normalizing each single entity risk score against each other. This gives security practitioners a meaningful, rank stacked list of threat leads out of millions of entities continuously measured over time.

How Data Science Creates Risk Scores

To create accurate risk scores, ArcSight Intelligence's analytics engine utilizes artificial intelligence techniques that include probabilistic methods for uncertain reasoning, clustering algorithms, classifiers and statistical learning methods, and neural networks. It also employs a statistical approach to compress all the various anomaly probabilities into a single entity risk score, a critical aspect of meaningful, actionable security analytics. A score is computed for each event to quantify the anomalies. ArcSight Intelligence aggregates these event probabilities to their associated entities, considering the entities' previous risk scores. This produces a risk score that considers all entities related to an event based on all the context ArcSight Intelligence can gather.

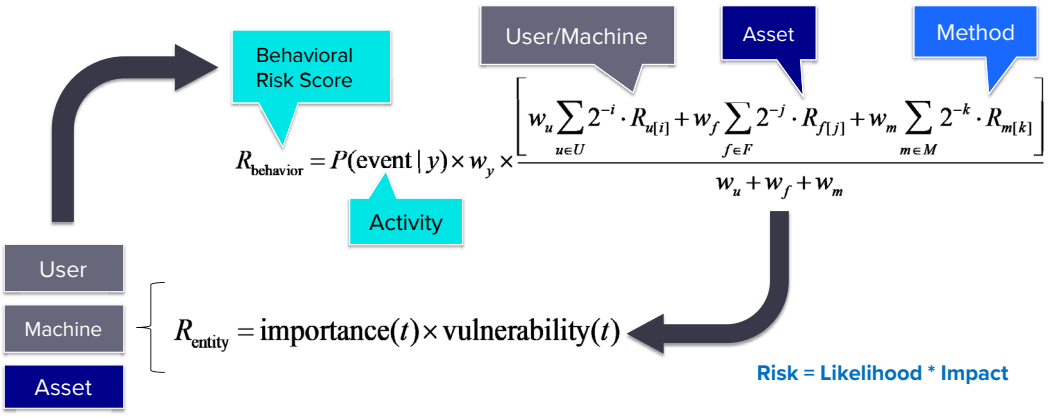


Figure 3. Statistical approach to compress all the various anomaly probabilities into a single entity risk score

As the analytics engine refines anomaly patterns, it learns which users create more risk, which files are the most at risk, and which machines are most often part of risky activities. Through statistical analysis, the engine also quantifies just how anomalous an observed behavior is. The more an entity is involved in high-risk anomalous activities, the more its risk score increases. Conversely, a risk score will decrease over time for an entity that is not involved in high-risk activities, but instead behaves normally compared to itself and other similar entities.

This risk scoring is calculated for every single entity—every single user, machine, IP Address, web server, files hare, etc. From these millions of individualized risk scores, the machine learning normalizes them such that they can all be compared with each other accurately, creating a single rank-stacked list of threat leads for security teams to prioritize time and effort.

ArcSight Intelligence's user experience is optimized to give security analysts the quickest path to the events that matter. It simplifies the advanced multi-level analytics from our machine learning platform to make the results easy to understand and actionable.

Ease of Use

Intuitive UI Design

ArcSight Intelligence's user experience is optimized to give security analysts the quickest path to the events that matter. It simplifies the advanced multi-level analytics from our machine learning platform to make the results easy to understand and actionable.

Security analysts have the unenviable job of identifying threatening activities within millions or billions of events, which is no small feat. ArcSight Intelligence's user experience is optimized to alleviate that pain through its risk-based aggregation and easy-to-understand, drill-down functionality. This provides a top-down view that focuses on a small number of risky persons or objects, as well as an easy drill-down click path that is always prioritized based on risk.

The UI guides the user with an easy-to-understand list of top risky entities based on machine learning.

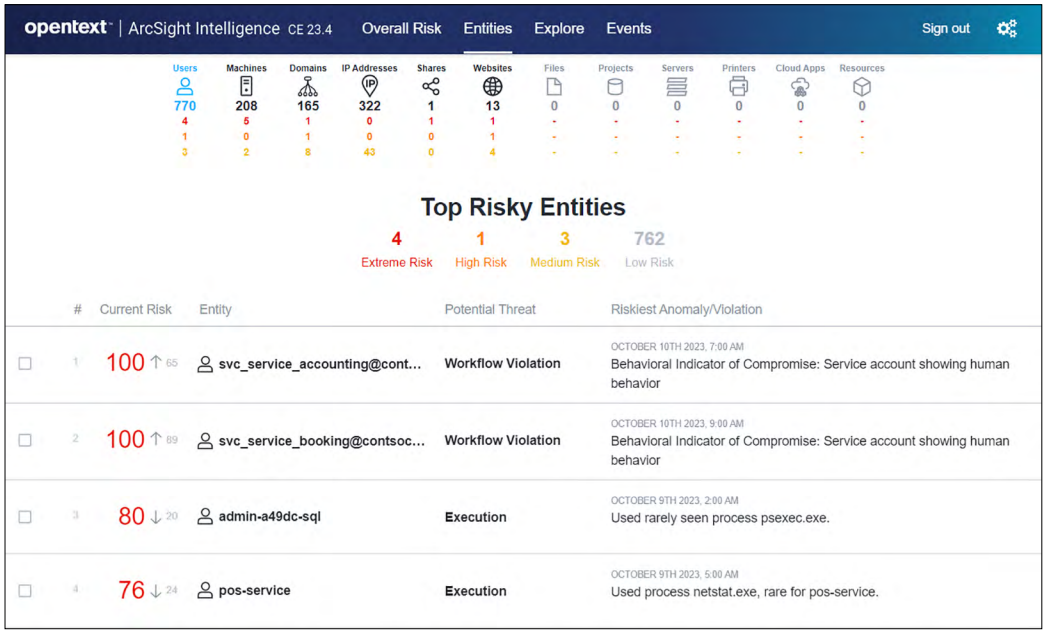


Figure 4. Top Risk Entities report

This easy-to-understand list of riskiest entities provides analysts with the ability to quickly understand why a user or other entity has a high-risk score and provides contextual evidence for why the risk score is high (or low) and how it compares to other threats.

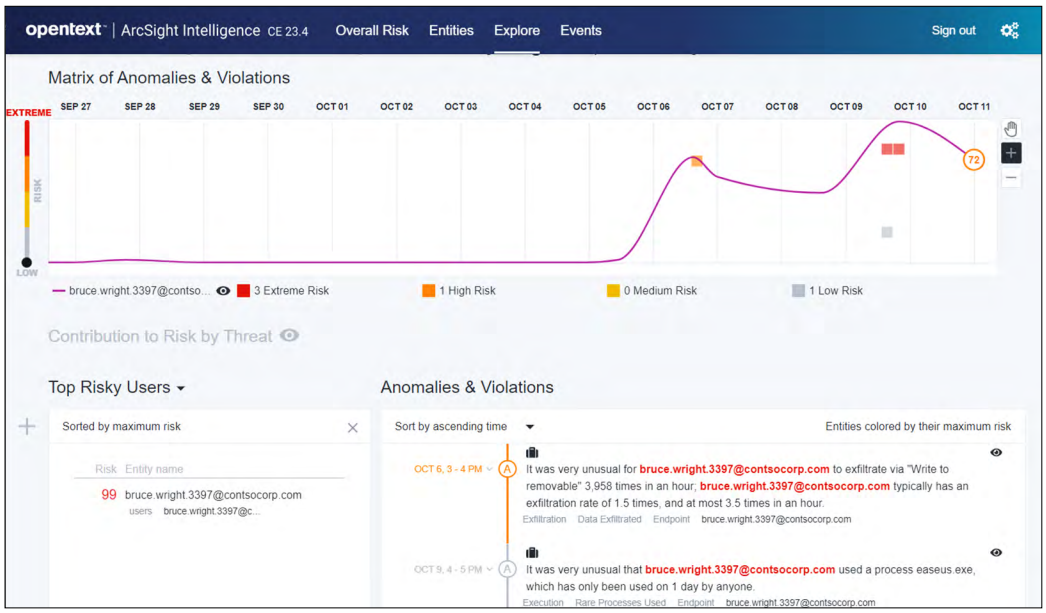


Figure 5. Matrix of Anomalies and Violations

An entity-focused risk view provides the analyst with intuitive controls to navigate the data. For example, simply clicking and dragging across the timeline view reduces the field of vision to a period of interest, dynamically updating all associated panels accordingly. The anomalies and violations shown in the timeline will respect the selected time horizon, empowering the analyst to get as granular as required to successfully investigate risky behaviors.

Figure 6. Detailed view of an anomaly

...analysts can very quickly identify the events that matter out of millions or billions of events and understand the full narrative of an attacker's path and sequence of activities.

In the above interface, the text and visual representations clearly show the difference between the observed anomaly and the parameters for normal behavior for an individual, as well as for a statistical peer group. In this specific example, the average number of file writes to a removable drive for Bruce Wright is 1.5 per hour, and his expected highest is 3.5 per hour. At the same time, the image shows that the average of other users like Bruce is similar at 1.3, but with an expected high of 3.3.

The observed anomaly is noted as being much higher than both individual "unique normal" as well as the peer groups' "unique normal." This type of visual design makes it much easier and faster for security teams to understand the full context of a threat. Additionally, directly from this view, there is an easy-to-follow click path and a series of intuitive visualizations. Through this, analysts can very quickly identify the events that matter out of millions or billions of events and understand the full narrative of an attacker's path and sequence of activities.

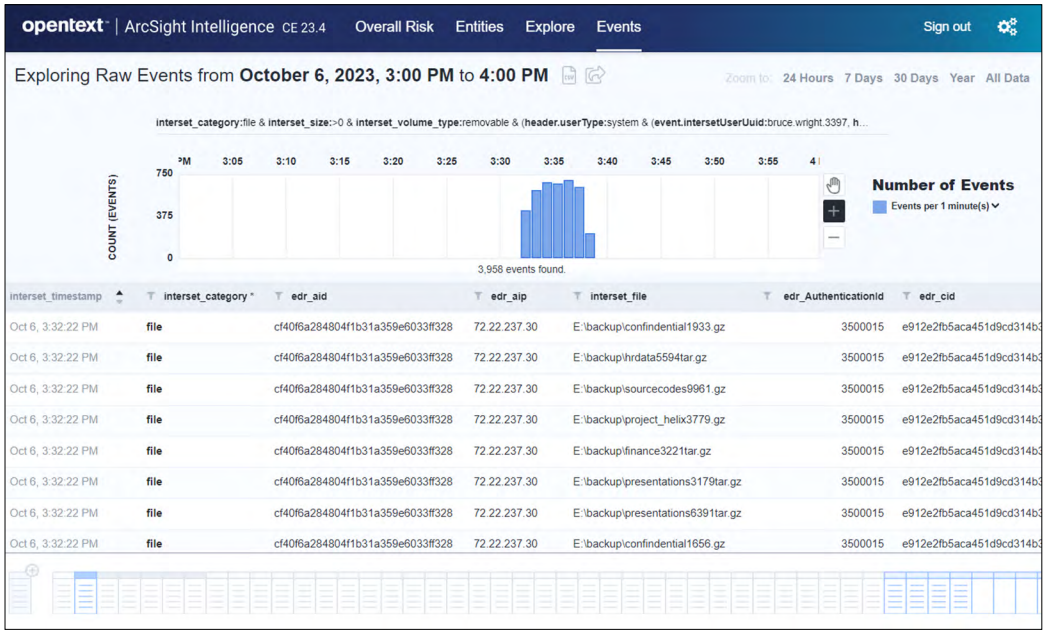


Figure 7. Raw events report

ArcSight Intelligence's UI is supported by a reporting engine and full REST API, providing the flexibility to leverage ArcSight Intelligence's rich set of results in any operational environment.

Conclusion

ArcSight Intelligence's extensive use of unsupervised machine learning allows it to determine unique normal behavioral baselines for all entities it encounters. This allows it to uncover unusual behavior effectively and create meaningful entity risk scores that are vital for effective risk detection. Unlike other analytical methods (including supervised machine learning), ArcSight Intelligence does not require any prior knowledge of how the threats will unfold. It can therefore detect unknown, zero-day threats. Attack techniques evolve but the one constant is that entity behavior always changes as a result.

By concentrating on "insider" behavior, ArcSight Intelligence helps detect threats to internal systems and data, regardless of whether the threat actor is an actual insider or an external party with insider access.

By concentrating on "insider" behavior, ArcSight Intelligence helps detect threats to internal systems and data, regardless of whether the threat actor is an actual insider or an external party with insider access.

In part two of our "What Makes ArcSight Intelligence Different" series we dive deeper into how our mature unsupervised machine learning empowers our unique approach to threat detection. Interested in seeing what ArcSight Intelligence can do for your organization, schedule a demo.

Connect with Us

www.opentext.com

opentext™ | Cybersecurity

OpenText Cybersecurity provides comprehensive security solutions for companies and partners of all sizes. From prevention, detection and response to recovery, investigation and compliance, our unified end-to-end platform helps customers build cyber resilience via a holistic security portfolio. Powered by actionable insights from our real-time and contextual threat intelligence, OpenText Cybersecurity customers benefit from high efficacy products, a compliant experience and simplified security to help manage business risk.