WHITE PAPER

UEBA and Machine Learning: Automating Data Security Analysis

Organizations seeking to accelerate, triage and mitigate user-driven data security risks

TABLE OF CONTENTS

Introduction

Securing electronic data in today's environment remains a daunting challenge, particularly as technologies including mobility and the cloud continue to increase the complexity of maintaining effective defenses.

For decades, organizations have invested significant resources in creating layered IT security policies and infrastructure. However, as those methods and tools have continued to mature, so has the complexity of related management. While existing methods are adept at determining where issues occur, practitioners are frequently challenged to pinpoint critical incidents and prioritize response, based on the requirement to analyze huge volumes of security data generated by a vast array of sources.

To overcome this hurdle, today's organizations require more effective analytical capabilities that calculate the precise intersection of sensitive data and user behavior, allowing them to focus on those responsive actions that will directly mitigate emerging data security risks.

User Behavior Analysis Holds the Key to Data Security

One of the most reliable methods of identifying potential data security risk is through application of comparative analysis to establish a baseline that can be used to isolate abnormal behaviors. Having created a contextual backdrop for future comparison, security practitioners are able to monitor ongoing activities to prioritize those issues that require further investigation.

For over three decades, numerous security tools, from endpoint protection agents to data loss prevention (DLP) platforms and cloud security systems, have leveraged this means of analysis to help unearth threats. However, as numerous breach incidents have proven, the ability to identify ongoing attacks is frequently thwarted by the need to evaluate large repositories of security data to efficiently escalate troublesome activities.

In fact, in nearly every situation where user activities or manipulated privileges have been cited as the origin of specific breach incidents, it has been discovered after the fact that existing security methods accurately identified those issues. As a result, addressing this susceptibility of security process and tooling, typically incurred by significant data analysis, remains a major focus for most organizations.

This prevailing environment stands as the backdrop for wider adoption of User and Entity Behavior Analytics (UEBA) within the context of enterprise security strategy. Via the addition of dedicated analytics that baseline user interactions and augment the input of human analysts to diagnose abnormalities that indicate threats, UEBA represents a significant breakthrough in addressing today's pervasive challenges.

Notably, those UEBA technologies empowered by purpose-built machine learning and designed to interface directly with existing security infrastructure to analyze complex data sets hold the key to achieving measurable progress. Symantec® Information Centric Analytics (ICA) is a software analytics platform that provides such an integrated, contextually enriched view of enterprise data security.

How it Works: Machine Learning for UEBA

As with most IT security technologies, UEBA solutions have matured over the course of many years to achieve widespread adoption. Similar to other types of analytics systems, in particular, early UEBA tools were widely perceived as highly technical capabilities best utilized by experienced data analysts.

However, today's packaged UEBA solutions, mirroring the path of endpoint, DLP and cloud security systems, among others, encompass a more straightforward and user-friendly approach. This improvement has been achieved largely through related development of underlying machine learning capabilities that speed delivery and interpretation of actionable results.

BY CUTTING THROUGH THE DATA, THIS FORM OF AUTOMATION GREATLY INCREASES THE ACCURACY AND PRIORITIZATION OF INCIDENT HANDLING, SAVING COUNTLESS HOURS OF HUMAN INVESTIGATION PREVIOUSLY REQUIRED TO GAIN CONTEXTUAL AWARENESS

These supporting technologies, specifically the combination of both unsupervised and supervised machine learning, have accomplished this feat by creating the required baseline for accepted behaviors and monitoring analyst inputs in a more automated fashion. By weaving together automatic and human-aided behavior analysis, this class of emerging UEBA platforms have established new benchmarks in processing large amounts of security incident data to effectively categorize user-based risks.

Automated Apprenticeship: Unsupervised Machine Learning

If the greatest challenge facing today's practitioners is finding the needle of abnormal user behavior within the proverbial haystack of existing security incident data, use of unsupervised machine learning represents a significant opportunity.

Clearly one of the most significant challenges involved in this process relates to building reliable models of accepted and non-malicious behavior, a pattern that varies greatly within every organization. For example, there are countless business processes involving handling of sensitive customer data within a large financial services provider, or a bank. As such, related policies, such as those implemented within DLP tooling, must be continuously refined to differentiate between both acceptable and potentially troublesome activities.

This is where unsupervised machine learning offers a major breakthrough, allowing organizations to aggregate and analyze large volumes of security incident data, without requiring human analyst interaction, to create a necessary baseline. By automatically compiling and modeling existing data using behavioral analytics, organizations are rapidly empowered to identify both those events that represent truly problematic user activities, along with broken business processes and non-malicious violations that merely require further policy adaptation to limit alerts.

By cutting through the data, this form of automation greatly increases the accuracy and prioritization of incident handling, saving countless hours of human investigation previously required to gain contextual awareness. Using such automation to refine and simplify analysis also allows incident handling to be addressed by a far greater breadth of security staff.

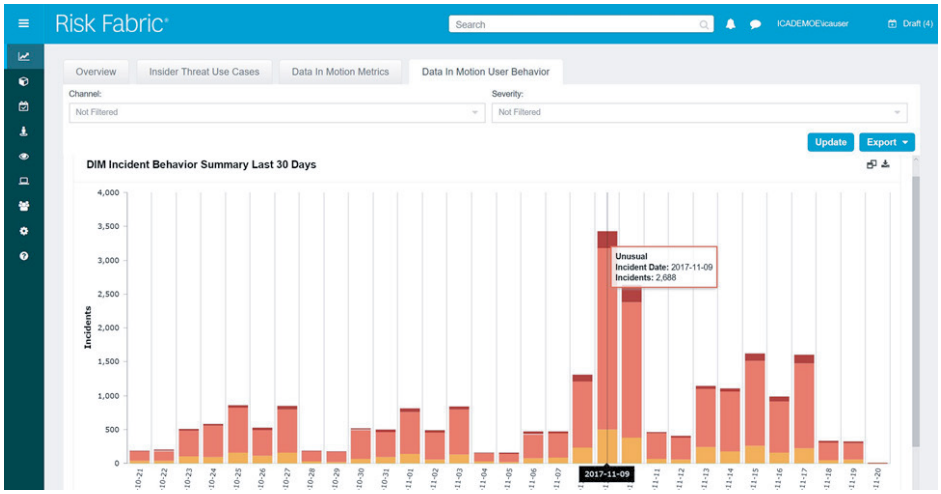
Learning from the Master: Supervised Machine Learning

Just as advancement of unsupervised machine learning has accelerated analysis and empowered use of UEBA among a wider set of practitioners, maturation of supervised machine learning has enabled the tools to become increasingly effective in the hands of those workers.

SUPERVISED MACHINE LEARNING OBSERVES AND INCORPORATES THE INPUT OF HUMAN SECURITY ANALYSTS, ALLOWING UEBA SYSTEMS TO BECOME MORE ACCURATE OVER TIME

By comparison, supervised machine learning, most often reinforcement training, observes and incorporates the input of human security analysts, allowing UEBA systems employing this capability to become more accurate over time. While unsupervised machine learning helps form a baseline through comparative analysis, supervised machine learning builds a historical reference through monitoring and incorporation of analyst workflows to make recommendations informing future investigation.

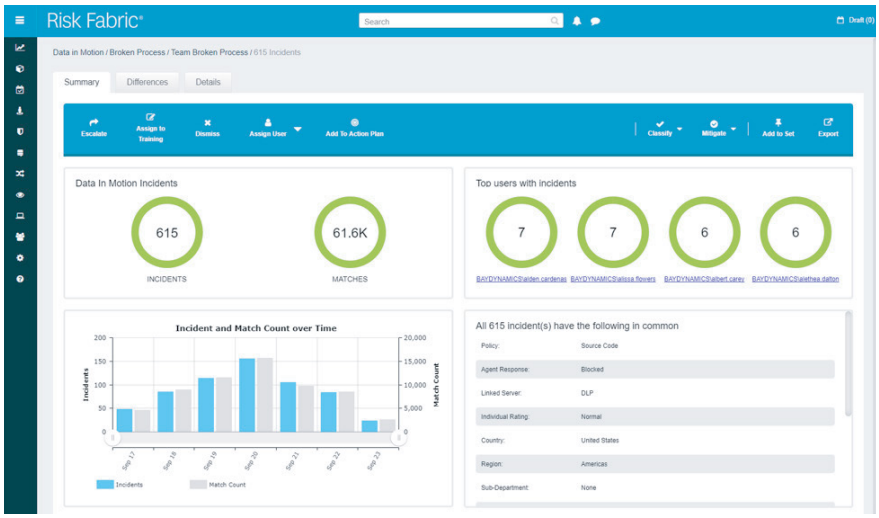
Figure 1. ICA's Unsupervised Machine Learning Allows for Automated Identification and Prioritization of an Organization's Most Problematic Security Incidents



For instance, referring to the same hypothetical financial services company, a set of incidents that at first appear to be malicious user interactions, may in fact turn out to represent a legitimate set of processes improperly accounted for by existing data security policies. As a result, analysts tasked with investigating those alerts repeatedly deescalate the incidents as they do not require immediate attention.

Over time, better informed by supervised machine learning of those analysts' actions, these same events can be automatically deprioritized, or assigned for policy reviews, allowing the UEBA system and its operators to become far more accurate in isolating real-world risks. In addition to saving time and focusing resources on more tangible risks, the system can flag behaviors that represent non-malicious or unintended data security risks and automatically refer the involved users for additional security training.

Figure 2. ICA's Supervised Machine Learning Tracks Analyst Classification and Remediation Activities to Make More Effective Recommendations over Time



Immediate impact: Use Cases

UEBA, BACKED BY ADVANCED MACHINE LEARNING, HAS PRODUCED A DIRECT IMPACT ON THE IMPROVEMENT OF MANY EXISTING SECURITY METHODOLOGIES

In practical terms, this manner of UEBA, backed by advanced machine learning, has produced a direct impact on the improvement of many existing security methodologies. By deploying Symantec ICA alongside other solutions, adopting customer organizations have seen significant benefits across three specific manners of workflow including:

- **DLP simplification:** Leveraging comparative UEBA capabilities to boost analysis and prioritization of DLP alerts to highlight those events that truly represent potential risks.
- **Hunting malicious insiders:** Creating a persona-based approach to investigate user-based incidents through integration and analysis of both alerts and human resources data.
- **Cyber-breach analytics:** Integrating incident data from a broad swath of tooling including DLP, Endpoint Protection and CASB systems to centrally analyze problematic activities.

While there are numerous additional use cases that have already emerged, or can be easily deployed based on these evolving UEBA and machine learning capabilities, these primary cases alone have produced measurable gains among existing customers.

For instance, in one case a global media and telecom provider that integrated ICA's UEBA capabilities alongside its Symantec DLP implementation was able to assign a full 80% of reported and ultimately non-malicious end-user policy violations for automated mitigation, directing affected parties to video-based security training and greatly increasing the efficiency of its analyst teams.

In another example, a global electronic payments leader implemented Symantec ICA and Symantec DLP in concert, leveraging the combined solutions' abilities to prioritize incident response to reduce its team of dedicated analysts from 35 to 5, reassigning the involved staff to address other strategic risk-management initiatives.

With plans to integrate numerous additional data feeds from across their existing security infrastructure using ICA's out of the box integrations, these two organizations clearly illustrate how ICA's payload of UEBA capabilities represent a huge leap forward in maturing data security practices.

Figure 3. ICA's Automated Machine Learning Allows for Continuous Analysis and Visualization of Peer-Based User Risk across Numerous Platforms

Symantec Delivers Today's Integrated UEBA

Every security organization wants to become more efficient at filtering existing security data to make accurate decisions. Symantec Information Centric Analytics (ICA), integrated with Symantec DLP, Symantec Endpoint Protection, Symantec CloudSOC Cloud Access Security Broker (CASB) and many other solutions including Active Directory and other human resources systems, enables rapid identification of malicious insiders and cyber breaches.

ICA ESCALATES ISSUES THAT MIGHT OTHERWISE GO UNNOTICED

Through centralized UEBA, extensive dashboards and in-depth metrics, ICA escalates those issues that might otherwise go unnoticed or demand complex manual analysis. With automated remediation recommendations, ICA provides organizations with the end-to-end workflows and action plans necessary to directly reduce exposure to sophisticated threats.

Backed by patented unsupervised and supervised machine learning, ICA not only provides tactical ability to isolate and mitigate those incidents that represent an organization's leading risks, but also represents a powerful set of capabilities for improvement of enterprise security management, allowing for continuous measurement and improvement of the overall reach and precision of security infrastructure.

SYM-UEBAML-WP100 March 18, 2025