exabeam™

The Ultimate Guide to Behavioral Analytics

eBook

Table of Contents

Introduction

User and entity behavior analytics (UEBA) was defined by Gartner as a category of cybersecurity solutions that analyze user and entity behavior on networks and other systems, and apply advanced analytics to detect anomalies and malicious behavior.

These can be used to discover security threats like malicious insiders and privileged account compromise, which traditional security tools cannot see.

This guide was created to help clarify the confusion about the UEBA market and addresses:

- **What UEBA is**
- **Why the market exists**
- **Why UEBA is needed**
- **How UEBA is different from other security tools**
- **The different types of UEBA solutions**
- **The benefits of UEBA**
- **Factors to consider when evaluating UEBA solutions**

We also consider some threat-centric use cases.

Our hope is that this guide will help organizations evaluating UEBA solutions better understand UEBA and how it can be adopted to improve their overall security posture with faster, easier, and more accurate threat detection, investigation, and response (TDIR).

What is behavioral analytics?

The UEBA software market was valued at $373.37 million in 2020, and is projected to reach $5.47 billion by 2028, growing at a CAGR of 40.5% from 2021 to 2028, according to Verified Market Research. Modern enterprise IT security solutions use this technology to detect and remediate complex threats that cannot be addressed by traditional rule-based solutions.

UEBA solutions ingest operational data from many sources and use analytics such as machine learning (ML) and behavior analysis to determine what is normal behavior by human users and non-human entities operating in an enterprise network. Entities may include IT assets such as hosts, applications, network traffic, service accounts, and data repositories. Over time, the solution builds standard profiles of behavior for these users and entities across peer groups to create a baseline for what is normal. When anomalous activity is identified, it is assigned a risk score. The score rises with increasing amounts of anomalous behavior until it crosses a predefined threshold. The UEBA solution then sends an alert to security operations center (SOC) analysts, and they use the data to investigate the incident. Some solutions can automate response actions if the incident reaches certain thresholds.

Why the behavioral analytics market exists

In 2012, Exabeam founder Nir Polak wondered why monitoring unusual behavior, which was being used to generate credit card fraud alerts, wasn't being used in cybersecurity. Credit card fraud alerts are triggered by unusual behavior — a suspiciously large purchase, buying shoes in another state, or shopping for jewelry in the middle of the night. Card issuers catalog a user's typical buying patterns, and when something abnormal pops up, the card issuer flags it. Knowing the daily routine of an employee to potentially spot unusual behavior — a likely sign of an insider or credential-based threat — seemed to be a logical cybersecurity measure.

A handful of cybersecurity startups had in the past tried to use credit card fraud detection techniques for cybersecurity, but all had failed. These companies had relied on "expert systems" using rules written by human experts. These systems were limited to what the experts knew (known knowns), were somewhat static, and were unable to adapt to frequently changing cybercriminal tactics.

These companies had relied on "expert systems" using rules written by human experts. These systems were limited to what the experts knew (known knowns), were somewhat static, and were unable to adapt to frequently changing cybercriminal tactics.

However, the market has changed in two ways since these companies failed:

- ML has advanced as a field from simple supervised and unsupervised learning into advanced hybrid modeling techniques. ML doesn't rely on experts; its routines are always watching and learning, and because of its ability to process huge amounts of behavioral data, it is much better than humans at determining what normal, or abnormal, behavior really is.
- Big data drove down the cost of data storage. Previous organizations that attempted to apply behavioral analysis to cybersecurity had been stymied by data storage that was so expensive it limited the number of behavioral signals the application could store. But by 2012, you could store massive amounts of data for a nominal amount of money.

Nir Polak and his team launched Exabeam in 2013. The "Exa" refers to an exabyte of data (one billion gigabytes), and "beam" to the ability of the Exabeam UEBA offering to shine a focused light on meaningful patterns in huge volumes of log data.

In 2014 Gartner coined the term "User Behavior Analytics" to describe what Exabeam was doing, and in 2015 updated its definition to include an "E" for "entity". **The User and Entity Behavior Analytics (UEBA) category now included behavioral analysis not only of human users, but also of entities such as routers, servers, and other network devices and endpoints. UEBA can analyze behavior across multiple users, IT devices, and IP addresses to detect anomalies.**

The "Exa" refers to an exabyte of data (one billion gigabytes), and "beam" to **the ability of the Exabeam UEBA offering to shine a focused light on meaningful patterns in huge volumes of log data.**

Why behavioral analytics are needed

Cyberattacks are increasingly sophisticated and often invisible to traditional rule-based security solutions. Security analysts respond as well as they can with the incumbent set of tools, but these tend to swamp analysts with alerts. Unfortunately, these alerts provide no context and are consequently useless when an analyst is trying to rapidly detect and remediate compromised credentials and the lateral movement of attackers.

The advanced analytics of a modern UEBA solution employ a different approach by using variations of AI and ML, data enrichment, and data science to effectively combat advanced threats. This modern UEBA solution integrates all the data sources and context for analysis and automatically synthetizes results. Analysts get fewer, more actionable alerts, and the organization gets a future-proof solution.

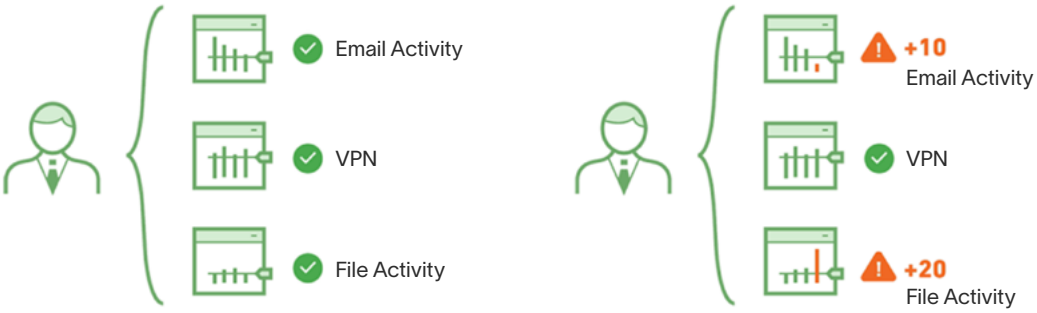

Figure 1.

Behavior analytics let you add risk scoring to a series of weak signals from other devices to aggregate into profiles for both users and entities or devices.

Augmenting or replacing your SIEM with behavioral analytics is the only way to effectively address the following use cases:

- Detecting compromised user credentials
- Detecting privileged user compromises
- Monitoring executive assets
- Detecting compromised systems, hosts, and devices
- Differentiating normal from malicious behavior
- Identifying and tracking lateral movement
- Detecting data exfiltration
- Providing context to failed login attempts and account lockouts
- Identifying service accounts and misuse
- Automating detection, triage, and investigations

How behavioral analytics are different from other security tools

Categories of current cybersecurity tools include:

- Penetration testing
- Packet sniffers
- Encryption and tokenization
- Scanning web vulnerability
- Network defenses
- Network security monitoring
- Detecting network intrusions

UEBA falls into this last category. But rather than relying on attack identifiers, which require an attack to have been detected elsewhere before the enterprise under attack can discover it is being attacked and respond, UEBA uses advanced analytics to look for incidents that are out of the ordinary for any given user or entity, and reports them to the enterprise's SIEM system when a combination of factors indicates the threat level is high. In this way, UEBA is able to discover previously unknown attacks and enable threatened enterprises to respond to the attacks.

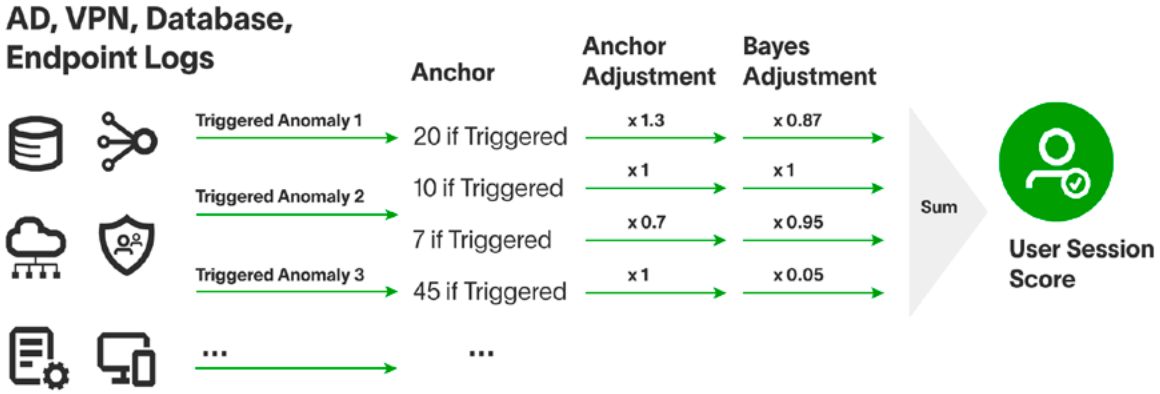

Figure 2.

A combined expert- and data-driven dynamic scoring framework used in modern UEBA solutions that reduces false positives by adjusting scores based on a variety of factors.

Advanced analytics

Advanced analytics involves several modern technologies that can help identify abnormal behavior even in the absence of known patterns:

Supervised machine learning

Sets of known good behavior and known bad behavior are fed into the system. The system learns to analyze new behavior and determine if it is similar to the known good or known bad behavior set.

Bayesian networks

Can combine supervised machine learning and rules to create behavioral profiles.

Unsupervised learning

The system learns normal behavior and is able to detect and alert on abnormal behavior. It will not be able to tell if the abnormal behavior is good or bad, only that it deviates from normal.

Reinforced/semi-supervised machine learning

This is a hybrid model where the basis is unsupervised learning, and actual alert resolutions are fed back into the system to allow fine tuning the model and reducing the signal-to-noise ratio.

Deep learning

Enables virtual alert triage and investigation. The system trains on data sets representing security alerts and their triage outcomes, performs self-identification of features, and is able to predict triage outcomes for new sets of security alerts.

Traditional analytics techniques are deterministic, in that if certain conditions were true, an alert was generated, and if not, the system assumed all is fine. The advanced analytics methods listed above are different, because they are heuristic. They compute a risk score which is a probability that an event represents an anomaly or security incident. When the risk score exceeds a certain threshold, the system creates a security alert.

Traditional analytics techniques are deterministic, in that if certain conditions were true, an alert was generated, and if not, the system assumed all is fine. **The advanced analytics methods listed above are different, because they are heuristic.**

An integrated view of multiple data sources

The true power of a UEBA solution is in its ability to cut across organizational boundaries, IT systems, and data sources and analyze all the data available for a specific user or entity.

A UEBA solution should analyze as many data sources as possible, some example data sources include:

- Authentication systems like Active Directory
- Access systems like VPN and proxies
- Configuration management databases
- Human resources data — new employees, departed employees, and any data that provides additional context on users
- Firewall, intrusion detection, and prevention systems (IDPS)
- Anti-malware and antivirus systems
- Endpoint detection and response systems
- Network traffic analytics
- Threat intelligence feeds

For example, a UEBA solution should be able to identify unusual login via Active Directory, cross-reference it with the criticality of the device being logged onto, the sensitivity of the files accessed, and recent unusual network or malware activity which may have enabled a compromise.

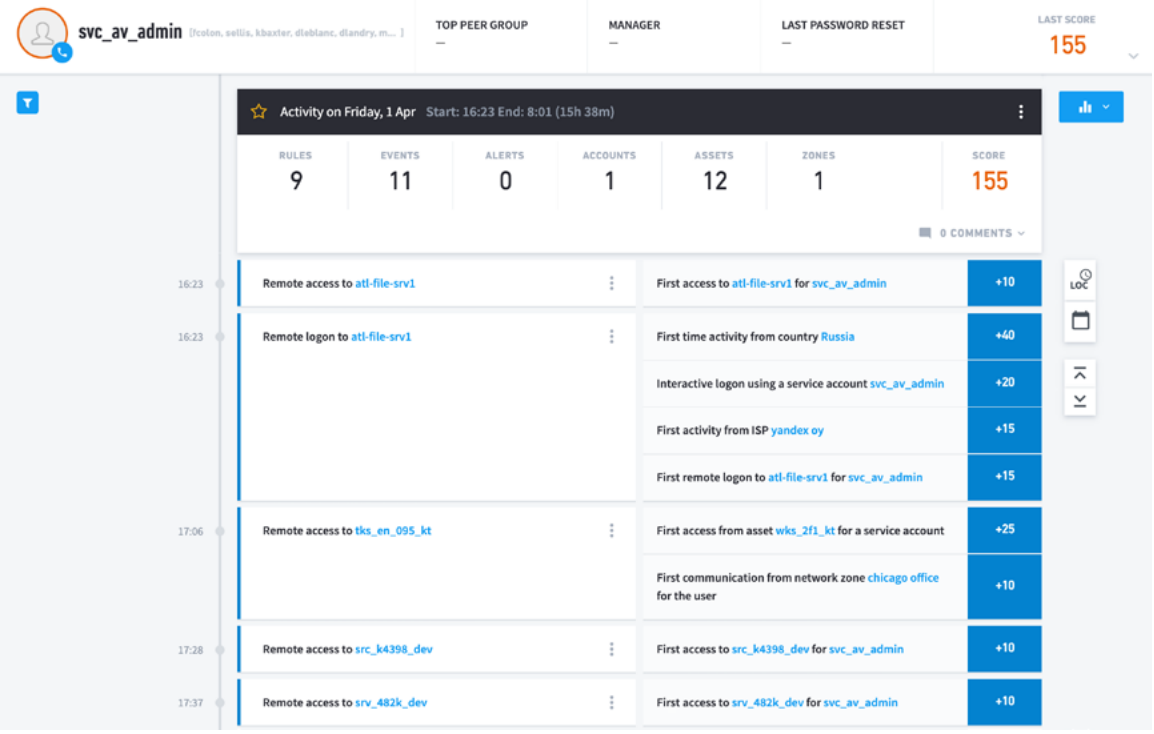

Figure 3.

A compromised service account showing unusual remote logons from new ISPs, new geographies, with new access permission

Learning normal to identify abnormal

A UEBA solution learns normal behavior to identify abnormal behavior. It examines a broad set of data to determine a user's baseline or behavioral profile. For example, the system monitors a user and sees how they use a VPN, at what time they arrive at work, which systems they log into, what printer they use, how often and what size of files they send by email or load to a USB drive, and many other data points that define the user's "normal behavior". The same is done for servers, databases, or any significant IT system.

When there is deviation from the baseline, the system adds to the risk score of that user or machine. The more unusual the behavior, the higher the risk score. As more and more suspicious behavior accumulates, the risk score increases until it hits a threshold, causing it to be escalated to an analyst for investigation.

Advantages

This analytical approach has several advantages:

Aggregation

The risk score is made up of numerous events, so there is no need for analysts to manually review large numbers of individual alerts and mentally combine them to detect a threat.

Reduced irrelevant alerts

One slightly abnormal event on its own will not result in a security alert. The system requires multiple signs of abnormal behavior to create an alert, reducing the number of irrelevant alerts and saving analyst time.

More context

Traditional correlation rules defined by security administrators may have been correct for one set of users or systems, but not for others. For example, if a department starts employing shift workers or offshore workers, they will start logging in at unusual times, which would trigger a rule-based alert all the time. UEBA is smarter because it establishes a context-sensitive baseline for each user group. An offshore worker logging in at 3:00 a.m. local time would not be considered an abnormal event.

Timeline analysis and session stitching

When analyzing security incidents, the timeline is a critical concept which can tie together seemingly unrelated activities. Modern attacks are processes, not isolated events.

Types of behavioral analytics

In behavioral analytics, security expertise meets ML to analyze and model activities driven by humans and entities on the system. Behavioral analytics model each dimension of user and entity behavior separately: from geographical location, access time of day, active days of the week, resources accessed. There are no assumptions about data distributions; the incoming logs are each scored against previous and expected behavior of the group, the individual, or the device, focusing on access to networks, systems, and assets.

Four principal ways behavioral analytics are used are: in the identification of network threats, automated application (authentication) security, email monitoring, and the next generation of antivirus detection.

1. Approaches to network threat identification

- UEBA analyzes baseline network behavior and identifies anomalies that have security significance
- Threat intelligence adds context to events, helping correlate network traffic with known attack techniques, known malicious IP or DNS addresses, or other IoCs from bad sources indicating that a known bad actor may be active on the network.

2. Automated application security

UEBA automated application security solutions can use machine learning to identify anomalous traffic, and even block it or respond automatically to an attack. They can identify malicious behavior such as unauthorized access and misuse of privileged accounts. They can also help automatically detect and avoid software vulnerabilities through static or dynamic code analysis.

3. Email monitoring

Machine learning can be used to improve the accuracy of existing approaches for detecting spam, malware or social engineering in email messages. Machine learning-based classification of images or other attachments to emails can help identify threats. Natural language processing (NLP) approaches analyze text within emails to see if the email may be part of a phishing campaign, and to analyze links within the email and decide if they are safe.

4. Next-gen antivirus

Traditional antivirus depends on signatures of known malware variants. This approach cannot deal with zero-day malware — new viruses that are not yet known to researchers or antivirus providers. It is also vulnerable to malware "mutations" where a virus is deliberately modified to evade detection. New AI-based approaches analyze malware source code or activity to understand if the software is legitimate or may be doing something malicious.

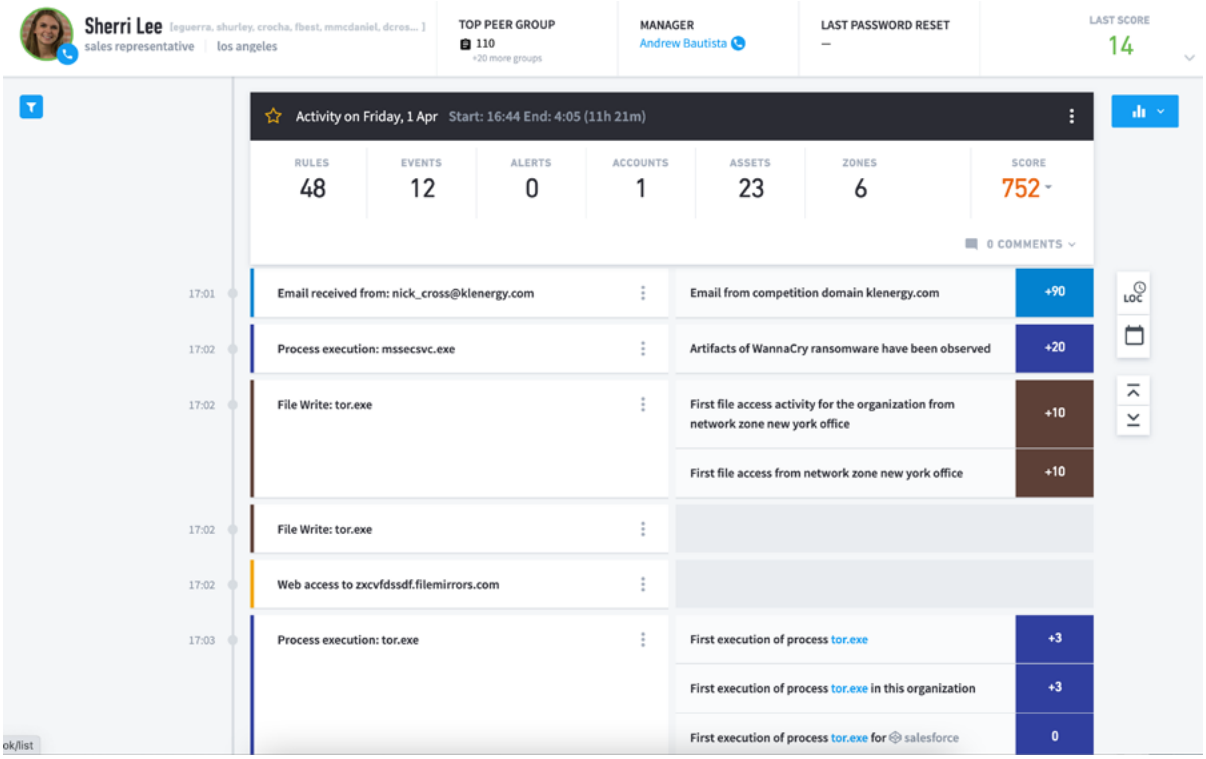

Figure 4.

Even for new infections and malware that the endpoint might not recognize, there are always other signs of anomalous activity that behavioral analytics can see and escalate. WannaCry traces may be known. How about the next version?

Benefits of using behavioral analytics solutions

Using ML with UEBA provides the ability to learn a behavior and integrate it into the detection engine. This saves analyst time writing and modifying complex correlation rules. Correlation rules are static and require analysts to create multiple iterations of the same rule to account for every possible scenario. This, in turn, leads to fewer unnecessary alerts.

UEBA dynamically adapts to an environment and can detect subtle changes in behavior in a way that is difficult to do with static correlation rules. The dynamic nature and detection capabilities of UEBA benefits your cybersecurity in many ways, including:

Detect breach of protected data

If you have protected data, it is not enough to just keep it secure. You need to know when a user with no legitimate business reason to do so accesses, alters, copies, or deletes this data. The UEBA system can detect this data through integration with other security tools and alert you when it happens.

Detect insider threats or credential compromises

An employee could go rogue, stealing data and information by using their access or by selling their login credentials to another malicious actor. UEBA can help you detect data breaches, sabotage, privilege abuse, and policy violations made by your staff. Alternatively, if an adversary compromises a system administrator's credentials, the adversary, using legitimate credentials, could escalate their privileges and move data within the environment, including offline storage (OST) files, documents, and presentations containing sensitive or proprietary data. Without UEBA, SOC analysts will likely find one or two true positive DLP incidents out of 100 or more alerts. UEBA helps reduce that number to identify true insider threats.

Flag changes in permissions and creation of privileged users

Some attacks involve the use of privileged users. If an attack doesn't start with privileged user credentials, it will often attempt to escalate the privileges or move laterally to acquire systems with higher-level privileges, such as an admin or service account.

UEBA alerts you when privileged users are created, or if there are accounts that were granted unnecessary permissions. According to the MITRE ATT&CK® framework, one of the common tactics, techniques, and procedures (TTPs) adversaries use is to establish persistence through Create Accounts (T1136). UEBA helps identify abnormal account creations based on the user's or role's baseline. For example, suppose a system administrator regularly creates new user accounts from 9 a.m. to 6 p.m. ET. When an adversary compromises the admin account and begins to create accounts outside that time frame, a UEBA solution will flag the activity as suspicious. The UEBA solution can also identify other anomalies such as the privileges granted, which system the privileges were granted from, the user's location, the network zone of the system, where each account moved or logged into, and other factors.

Detect brute force attacks

Cyberattacks sometimes target your cloud-based entities as well as third-party authentication systems. With UEBA, you can detect identity-based attacks such as phishing, brute force, dictionary, and others; this allows you to block access to these entities. For organizations that regularly monitor failed logins, there is not enough time in the day to look through a list of 200 accounts that generated a failed login event and identify which ones are potentially malicious. A UEBA solution can help prioritize the accounts that generated an abnormal number of failed logins based on the account profile and provide the contextual information to make a decision or trigger an automation.

Reduce unnecessary alerts

The UEBA system is constantly learning how to be more accurate. This approach reduces noise in the system because multiple abnormalities must occur before an analyst is alerted. ML and UEBA prevent getting a profusion of unnecessary alerts.

Some attacks involve the use of privileged users. If an attack doesn't start with privileged user credentials, it will often attempt to escalate the privileges or move laterally to acquire systems with higher-level privileges, such as an admin or service account. **UEBA alerts you when privileged users are created, or if there are accounts that were granted unnecessary permissions.**

Focus on threat-centric use cases

User behavior analytics solutions can help you discover security threats that traditional solutions, which are based on signatures, correlation rules, or simple statistical analysis masquerading as ML, cannot see. **Many vendors present their solutions as UEBA when they are not.**

Discovering compromised accounts

UEBA can identify user accounts taken over by attackers, because they exhibit anomalous behavior compared to the real business user or service account.

Identifying malicious insider threats

Insider threats are a major, growing threat and are extremely difficult to detect via traditional security solutions, because these attacks use legitimate credentials, services, and entities, and access privileges. UEBA solutions can identify malicious insiders by analyzing their behavior compared to the old baseline as well as to similar, non-malicious users within the same workgroups.

Identifying privileged account abuse

UEBA can help monitor accounts with administrative or escalated privileges to ensure they are not being misused by their designated owner or by others. Privileged account issues include policy violations or neglectful acts, which may not be malicious activity but can still have damaging results.

Cloud security monitoring

Cloud assets are provisioned dynamically and used remotely. This makes monitoring security more challenging when using only traditional network perimeter tools. UEBA can look at cloud-based assets and discover if, as a group, they are acting normally or abnormally. This includes coordination with cloud access security broker or data loss prevention tools, which can alert on unusual file size movement, file alteration, or inappropriate sharing.

Entity monitoring

UEBA can be used to monitor IoT devices, such as critical medical equipment, facility access points, or sensors deployed in the field. Behavior analysis can be used to establish a baseline for these groups of similar IoT devices and identify when a device exhibits anomalous behavior. For example, if an industrial control system's service account attempts to log into the Active Directory or web server, this highly unusual behavior will generate an alert.

Factors to consider when evaluating UEBA solutions

Many vendors claim to offer UEBA capabilities, but a variety of implementations make comparative evaluations difficult. **The following list can help your organization evaluate and select an effective UEBA solution:**

☐ Shows normal activity as well as anomalies

Normal activity is the typical behavior for a given user and that user's peers, and it is required to show context. For example, does this user normally access a particular server or that sensitive database? Does this user normally access the network via a VPN from Canada or other countries? Does this user normally upload large files to Dropbox? Also, do the user's peers normally do the same things? The UEBA should demonstrate the ability to show both normal and anomalous activity within a user session timeline. This lets investigators understand the situation in a broader context, which significantly reduces the amount of time spent on data gathering, validation, and subsequent investigation.

Normal behavior is also proof that the detection results aren't from a static correlation rule, because rules only fire when the bad condition is met. There would be nothing to show under non-malicious conditions. For example, static correlation rules do not tend to include time as a variable or consider normal traffic patterns when looking for abnormalities. Suppose the IT administrator takes care of Okta request tickets mid-morning after a meeting. Then one night, a set of brand-new users are added or granted Okta permissions of super users on a set of internal resources. All this activity is normal, passes no thresholds, and breaks no rules, but the time variance on the behavior is abnormal enough that a UEBA will create an event and make it easy to look at all the affected entities — from devices to users to resources touched.

☐ Connects a host to IP-to-user for establishing identity automatically

Hostnames are numerous and typically don't provide much useful identifying information. IP addresses are reassigned continuously, and account credentials are often shared. (This is especially true of administrative accounts, although this practice is forbidden by most compliance and security managers.) Host to IP-to-user credential mapping positively attributes activity on a specific host to a specific person, using a specific IP address, at a particular time. The UEBA should demonstrate the ability to link a host positively and automatically to an IP and an identity, even when shared accounts are used. Good UEBAs will also have device views showing all the users with permissions and privileges as well as user views showing all the entities and devices they access. This makes threat detection much more effective. Connecting these dots manually can take hours of work, so performing it automatically drastically reduces the amount of time to investigate and take remedial action.

☐ Detects lateral movement

A key indicator of a compromised account is lateral movement. During reconnaissance, an attacker will move throughout the network to hunt for valuable information, hopping between user credentials and devices to obscure their movements and avoid detection. The UEBA should demonstrate the ability to track lateral movement even as the user changes accounts, machines, or IP addresses. This capability ensures both effective detection and more accurate incident investigations, as the UEBA ties weak signals from multiple sources into a single events timeline and assigns risk for each change and permission enhancement.

☐ Creates timelines of all incidents automatically

Activity data is produced in the form of events, but detection and response require timelines. Stitching events into coherent timelines typically requires hours or days of significant manual effort. A full timeline should include every activity by the user and any other entities interacted with during a session from log-on to log-off, using data from all related endpoints, including network, DLP, physical security, and other systems. The UEBA should demonstrate the ability to produce coherent timelines of user activity, quickly and automatically. Many UEBA solutions do not provide a timeline for incident investigation; some provide a partial one at best. A machine-built timeline offers a better interface that is easily used by a junior analyst. Instead of presenting discrete events, a machine-built timeline presents the results with context and risk scoring to help rapidly distill the essence of a threat — and how to fix it if needed.

☐ Deploys and shows value quickly

As your organization considers UEBA options, look for those that can be deployed in a day or less and do not require extensive professional services for configuration, customization, and deployment, and provide built-in use cases. The UEBA should be able to deploy and begin operating with basic event feeds such as Active Directory and endpoint events in less than 48 hours and show clear value within one week. All of the models in operation should be visible, accessible, and transparent.

☐ Evolves to meet future needs easily and without additional costs

The UEBA should be able to scale and extend to new functionality without professional services or new engineering engagement from the vendor, e.g., needing new log sources introduced or correlation rules built. Vendors requiring multiple services to set up and tune their deployment will often require the same level of effort and cost when customers change their environment, add new data sources, or tackle new use cases. Customers shouldn't be penalized for evolving business initiatives. An effective UEBA makes it easy to show value as needs change over time — and can grow with you with clear visibility on the best log sources to meet your security and compliance requirements.

☐ Deploys without giving VPN access to the vendor

Many UEBA solutions require extensive service and heavy customization during deployment and after production cutover. This work is usually performed by offsite or offshore engineers and requires VPN access to your network. For many firms in regulated industries, this is a problem involving extensive change control, permission, and meetings. The UEBA should be able to deploy and be supported without external VPN vendor access. The UEBA shouldn't introduce new security risks, it should minimize them. Moreover, you want the evaluation to mirror those in your production environment, i.e., without vendor VPN access, so you know what you're getting into.

☐ Does not require agents or network taps to be deployed

Some UEBA solutions require that additional infrastructure be deployed to collect required data, typically via either endpoint agents (installed on every device) or network taps. This imposes a dependency that can extend the pilot by months. The UEBA should be able to operate without installing additional external agents or taps beyond getting the logs and APIs from your existing endpoint and other security solutions. While taps and agents can provide useful additional information, there are UEBA solutions, such as those that analyze log data, that provide value without requiring the deployment of additional infrastructure.

☐ Provides proactive threat hunting capabilities

Threat hunting with a UEBA solution entails performing simple and complex searches of collected security data. It should not require an in-depth understanding of a proprietary query language, rigorous attention to syntax, or the need to stitch together results from multiple, simpler searches. Nor should it take hours to return a result if you need to know what happened four months ago. Threat hunting capability driven by machine-built timelines with an extensive set of dropdowns covers a wide variety of potential arguments, operates on highly indexed data, and returns complete incident timelines.

☐ Integrates with SOAR for automation

Security orchestration automation and response (SOAR) is a big topic in the context of UEBA. After detecting a threat, typical security analysts' workflows require multiple products with multiple interfaces and credentials. This wall of glass panes reduces visibility, response speed, and productivity with a flurry of manual actions sometimes referred to as "swivel chair incident response." SOAR is an additional capability that augments a UEBA solution with a centralized approach and single console to pull in data, create and assign cases with prescriptive guidance, and automate some responses while pushing actions to other systems as required by your specific IR needs. Essentially, the integration of UEBA technology with SOAR tools automates incident response. Look for a UEBA solution with semi-automated or fully automated incident playbook actions. It will help your SOC analysts be more productive and accelerate incident response for better enterprise security.

About Exabeam

Exabeam is a global cybersecurity leader that created the New-Scale SIEM™ for advancing security operations. We Detect the Undetectable™ by understanding normal behavior, even as normal keeps changing – giving security operations teams a holistic view of incidents for faster, more complete response.

Learn more about Exabeam today

Get a Demo Now →