exabeam

WHITE PAPER

EXABEAM SMART TIMELINES

Prebuilt timelines that automatically reconstruct the events underlying security incidents

INTRODUCTION

IMPROVING THREAT HUNTING, INVESTIGATION, AND REMEDIATION REMAINS A PRIORITY FOR ALMOST ALL SECURITY ORGANIZATIONS. THEY STRUGGLE WITH BEING ABLE TO RAPIDLY SYNTHESIZE DATA TO YIELD ACTIONABLE INFORMATION.

In a recent survey from the SANS Institute,[1] respondents identified security organizations' top three priorities as: better investigation functions, more staff with investigative skills to conduct searches, and an improved ability to search and discover data and information.

The primary roadblock to delivering on these priorities is a lack of skilled staff—the most common security operations center (SOC) challenge.[2] But the answer may not be to hire more skilled staff; rather, better tools can reduce the time, effort, and skill required for security management.

Exabeam Smart Timelines are prebuilt timelines that automatically reconstruct the events underlying security incidents so analysts can stop spending time combing through raw logs. Smart Timelines display the full scope of a user's or device's activity in an easy-to-use and graphical manner, identifying anomalous behavior and risk (Figure 1).

Smart Timelines are much more than a collection of logs sorted by their timestamp. They reduce the time and specialization required to detect, investigate, and respond to security incidents by taking machine-generated data and converting them into a narrative that makes sense to security analysts.

[1] SANS 2018 Threat Hunting Survey Results, SANS Institute, Robert M. Lee and Rob T. Lee, September 2018

[2] SANS 2018 Security Operations Center Survey, SANS Institute, Christopher Crowley and John Pescatore, August 2018.

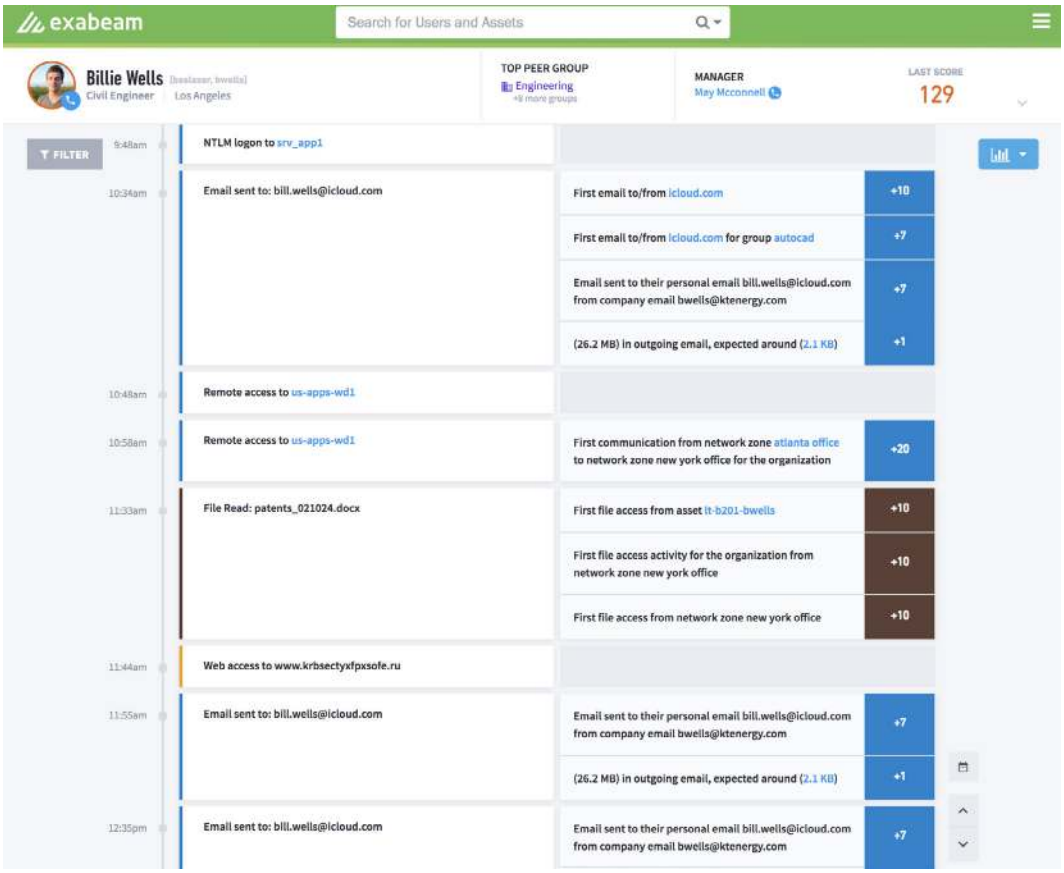

FIGURE 1 – AN EXAMPLE SMART TIMELINE SHOWING A LIKELY ATTACK INVOLVING LATERAL MOVEMENT (TO THE NEW YORK OFFICE) AND THE EXFILTRATION OF UNEXPECTEDLY LARGE FILES TO A NON-COMPANY EMAIL ADDRESS

This paper outlines the unique process Exabeam uses to convert IT and security logs into Smart Timelines. It also explains how each step in the process helps your analysts to work smarter.

WHY SMART TIMELINES?

They overcome the limitations of correlation rules

Most enterprise security analysts rely on correlation rules to detect potential threats. For them to work, analysts have to know in advance what they're looking for. For example, a rule for a virtual private network (VPN) log might raise an alert if the number of failed logins exceeds 20 times in five minutes. Or for an identity and access management (IAM) log, another rule might entail raising an alert if the same user account is created and deleted within a 24-hour period.

Smart Timelines provide analysts with a new method of threat detection (as well as investigation and response)—one in which they don't have to have prior knowledge of attacker tactics and techniques. With Smart Timelines, Exabeam preprocesses all logs and combines them with other data sources to detail user and asset activities. In this way it helps identify any anomalous behavior, and thus attacks.

Smart Timelines remove the need to build and maintain correlation rules. Yet, Exabeam also permits analysts to build rules to create alerts if desired—for example, if they are importing them from a legacy security incident and event manager (SIEM).

They let analysts quickly detect and respond to complex threats

Modern threats come in many forms and can be internal or external to an organization. Insider threats stem from the actions of someone within an organization who unintentionally, or with ill-intention, exfiltrates data or adversely affects your business.

Meanwhile, external threats have become more targeted; for example, malware can be installed inside your organization's network via a phishing attack, leading to a compromised user account.

Both insider and adversary-controlled activities can often be detected via anomalous deviations from users' and devices' historical patterns. By using a unique process to learn from all available log sources, Smart Timelines are assembled to help analysts discover such anomalous activities. With them, a prebuilt-incident timeline flags anomalies and displays details so analysts can fully scope an event and the severity of its risk (Figure 2).

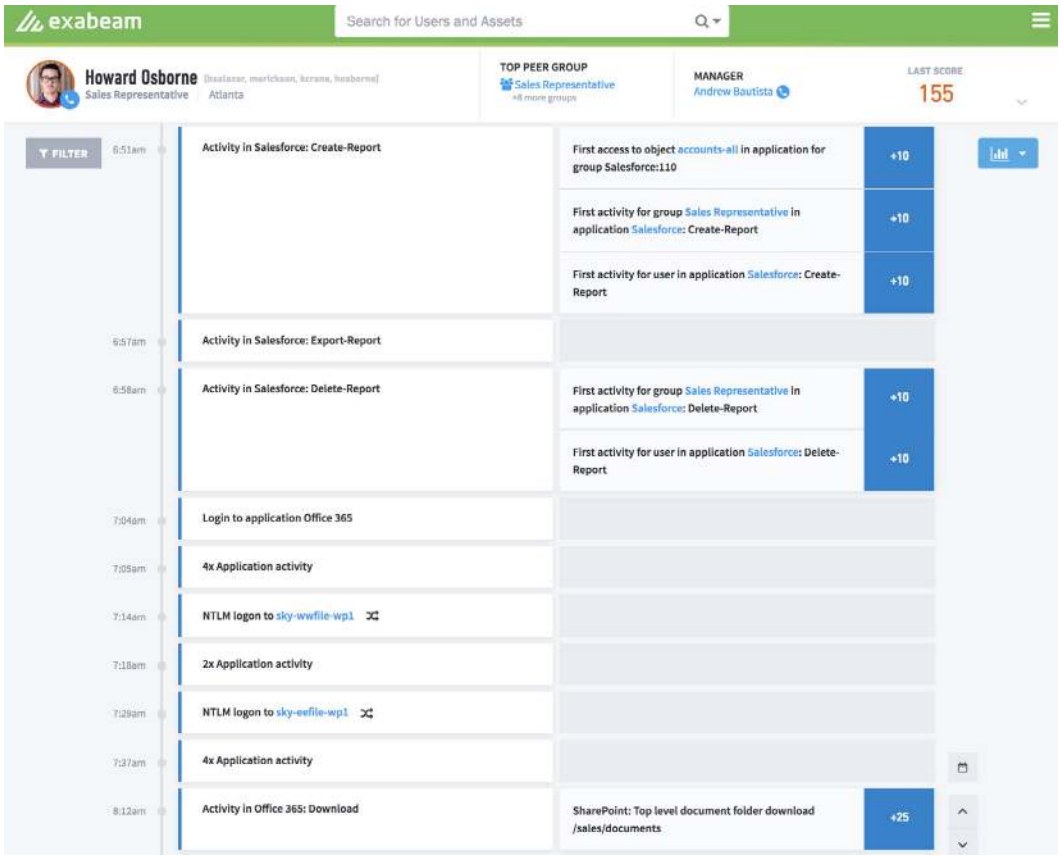

FIGURE 2 – AN EXAMPLE SMART TIMELINE FOR A USER SHOWING SUSPICIOUS FILE ACCESS, EXPORT AND DELETION

Here, time-ordered user or asset activities are displayed in plain language in the left column. Associated risk scores and explanations for the elevated risk are shown at the right. The overall risk score for the user or device is reported at the top right. Many elements (in blue type) contain additional information that may be needed during an investigation, such as members of groups the user may belong to and access history of applications and devices.

Smart Timelines let analysts detect insider and external threats and stop spending time combing through raw logs to investigate incidents. With them, what used to take weeks to investigate using a legacy SIEM can now be done in seconds.

Smart Timelines eliminate the need for analysts to build their own incident timelines, and reduces their need to "query and pivot" between IT and security applications to collect event details. Because they're a prebuilt timeline is automatically created for all users and devices, analysts can review the activities of a user or device anytime, whether there is an identified incident or not. For threat hunting, the dots have already been connected.

They track lateral movement

A specific challenge for security teams is handling lateral movement by attackers. The reason lateral movement is difficult to track is that, many times, logs alone don't contain all of the data analysts need to recreate an attack. For example, consider a user who switches between two accounts while logged onto a single machine (Figure 3). Without Smart Timelines, the "two users" would likely appear completely unrelated and a lateral movement attack likely would be missed.

Exabeam's patented technology follows attacks as they move through organizations. It's able to positively attribute relevant activity to the responsible user, even when lateral movement is involved. Such activity is then flagged as risky and displayed in Smart Timelines for rapid investigation (Figure 3).

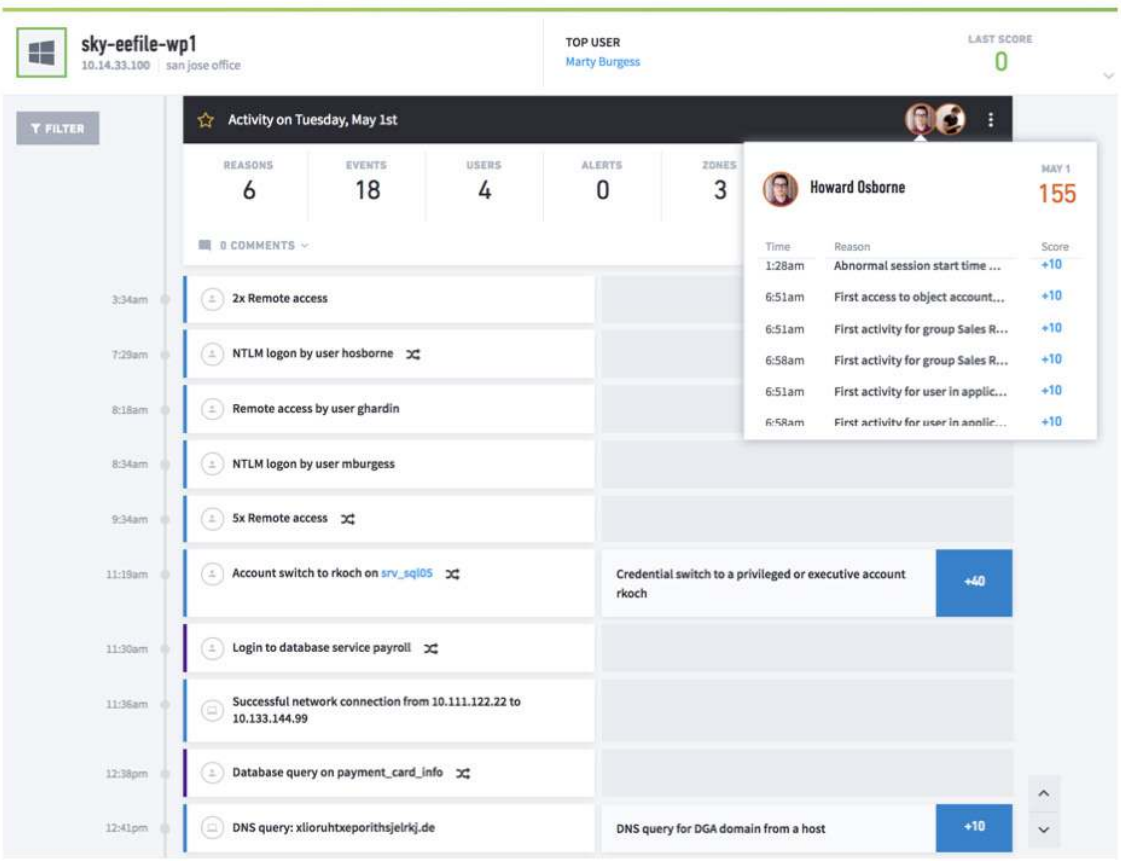

FIGURE 3 – AN EXAMPLE SMART TIMELINE FOR A DEVICE, SHOWING MULTIPLE LATERAL MOVEMENTS AS THE ATTACKER SWITCHES BETWEEN ACCOUNTS—INCLUDING A PRIVILEGED OR EXECUTIVE USER ACCOUNT.

SMART TIMELINES

The remainder of this paper outlines a nine-step process for building Smart Timelines. As we've seen, these machine-built timelines automatically reconstruct the events underlying security incidents so analysts can stop spending time combing through raw logs.

FIGURE 4 – THE STEPS TO BUILD SMART TIMELINES.

Step 1: Collecting logs

Security management improves as the amount of information available to analysts increases. This is analogous to a traditional crime investigation where the police benefit from having more clues.

Large companies and government agencies often log data sources that range from legacy on-premises systems to cloud applications and infrastructure. By enabling them to pull a large variety and volume of event data into a central repository, Exabeam helps analysts get the clearest picture of the activity of users and assets in their environment. Out of the box, it can ingest logs from over 300 vendors. Examples include:

- Microsoft Active Directory (AD) logs that record domain controller user-to-machine authentication events in a Windows environment
- Logs generated by firewall security products that track user activities crossing the network perimeter
- Endpoint products that track file actions, such as file deletion and creation

An aside: predictable flat-rate pricing

Most vendors use volume-based pricing for log management. This can result in excessive logging fees. It can also lead to organizations limiting the number of logs pulled into their SIEM to reduce costs or stay within a defined budget. This prevents them from having the broadest set of available information during security investigations.

Unlike most SIEM and log management vendors, Exabeam doesn't set its price based on log volume.

Instead it's based on the number of users and/or devices, making SIEM cost very predictable. This benefits organizations by letting them ingest as much data as possible without affecting costs. Furthermore, pricing is the same (and portable) whether Exabeam is deployed on-premises or in the cloud.

Exabeam pricing is based on the number of users and devices, not data volume.

Step 2: Parsing logs into fields

Being a text-based string of characters, logs are often very cryptic. For each log available in the Exabeam Security Management Platform, Exabeam security specialists have analyzed whether it contains information that is useful for security management. Its components are then identified (tokenized) and broken into fields to permit better searching, alerting, and reporting—part of the data normalization process. For example, a "2H3M4S" log field refers to a duration of 2 hours, 3 minutes, and 4 seconds.

As of December 2018, Exabeam provides its customers with over 2500 built-in parsers from over 200 vendors to help them quickly make sense of logs. If a parser doesn't already exist for a particular data source, Exabeam customers can quickly get one—usually within days—simply by sending log samples to Exabeam.

Exabeam provides its customers with over 2500 built-in parsers from over 200 vendors

Once a field is identified as useful, Exabeam uses a common classifying taxonomy to normalize it—making log fields easy for analysts to use by sort. For example, if an internet protocol (IP) address indicates the destination of a connection, it's given a "Dest_IP" label. This classification is available across the Exabeam Security Management Platform, including our Data Lake product.

Step 3: Creating events by combining log fields

The next step to build Smart Timelines is to identify events—actions that occur at points in time. Exabeam identifies nearly one hundred commonly occurring event types, such as, VPN access, local access, etc. in Smart Timelines. While other vendors provide timelines, they simply organize ingested logs into chronological order. Only Exabeam provides this event abstraction layer to normalize log information and make it easy for analysts to quickly understand what a user or device did.

To create an event, Exabeam combines relevant parts of a log or multiple logs. This process is made easier for analysts by the Exabeam team's familiarity with the fields in the logs generated from IT and security services. Exabeam performs this data manipulation so security analysts can spend their time on investigations and incident remediation, not analyzing logs.

Step 4: Adding context while creating events

Building events doesn't stop there. Often data from log fields tell only part of the story. Exabeam further enriches the event data in thousands of ways by adding information that the log source didn't have originally. Again, this helps improve analysts' understanding of what took place.

For example, most systems identify new users by their Active Directory (AD) account. Some systems use an email address to identify new users; others may use an organizational role to identify new users—for example, denoting someone as an IT administrator. To assist analysts, Exabeam maps emails and user hierarchy to AD accounts, adding a new "user" field to augment the original log.

The information to do this simple mapping often isn't readily available. In such cases, Exabeam infers user information using a data graph underpinned by machine learning models. To oversimplify, if a user always logs into a device assigned to Barbara Salazar, that person could be identified as Barbara Salazar, assuming there isn't contradictory evidence to suggest otherwise.

By building a data graph, Smart Timelines automatically fill in the missing holes in log data.

By building a data graph, Smart Timelines automatically fill in the missing holes in log data. This happens in real time, often using millions of logs from thousands of users and machines which are constantly changing their IP addresses. As a result, Exabeam

works from complete datasets—using them to track all user and device activity – while other SIEMs are building timelines from incomplete datasets. This data graph is especially critical for identifying lateral movement.

Exabeam also adds context to dynamic activity to characterize events. For example, Exabeam models the source from which a user account was created by identifying which machine was used to make a request. It denotes whether an account was created by a machine an admin typically uses (a common occurrence), or if it was done by a machine in the network DMZ (an unusual situation). This context is not provided in a log itself.

In instances like this, Exabeam uses the Marcov model, a model for random processes, to model these changing systems that otherwise appear to vary in a random manner and be an intractable problem when it comes to defining associations.

Said another way, logs are not enough.

Exabeam's process is radically different from a common approach that simply places all the raw logs into an environment like Hadoop. That approach not only limits how much event information is portrayed to analysts, but also leads to many false positives.

Exabeam uses a lot of machine learning coupled with a lot of researchers' time and effort to enrich log data. Exabeam does the enrichment once, so each organization's security analysts do not need to expend the effort themselves.

Information available to analysts through enrichment includes:

- Annotating whether an account belongs to a real user or a computer program
- Determining if a host/IP is a workstation or a server by analyzing the manufacturer's device naming convention, an IP range, and machine behavior (if it has a lot of users, it's probably a server)
- Dynamic peer grouping (automatically identifies peer groups based on user behavior and interactions with the IT environment—for use in detection rules, risk scoring, and context enrichment)
- Automated asset ownership association (determines asset owners based on users' behavior pattern and interactions)
- Automated host-to-IP mapping (automatically associates host names to asset IP addresses to detect and track lateral movements of users across an organization)

This is a short list of examples. Customers can also use Exabeam to enrich data themselves—for example, by adding industry-specific use cases.

In addition to the aforementioned user creation event, other examples of events include remote logons, data loss protection (DLP) alerts, and inbound/outbound emails.

Describing events in plain language in Smart Timelines

There isn't a common event code identification system used by the organizations that create applications. For example, Windows, RSA, UNIX, and Proofpoint each have unique codes for remote logons. In Microsoft AD logs, event D 4720 or 624 represents an account-creation event. This is unique to Microsoft. The lack of uniformity requires analysts to know event codes for each system.

To make life simpler for analysts to understand at a glance what took place in events, Exabeam takes the additional step of describing them in plain language as it builds Smart Timelines.

Step 5: Creating sessions

These named events are then grouped into user sessions that form the Smart Timelines. Sessions can be opened and closed based on behavioral conditions (such as logging into a laptop), or by a set of configurable set of logical conditions. The value of sessions is that they organize events in sequence; they also put time-based boundaries around behavior to establish a normal behavior baseline. A typical session might comprise a work day, during which a user's or device's activities are tracked through the network environment. Organizing events into sessions provides analysts with considerably more context than if they were to look at a single event in isolation.

Step 6: Modeling behavior and normal activity

Once events have been grouped into Smart Timelines, a combination of machine learning models and rules are applied to detect normal versus anomalous activity. Highly scalable, underlying machine-based models—used by organizations with over 200,000 users—describe how each user and machine behaves. They are what empower Exabeam to reveal anomalous behavior.

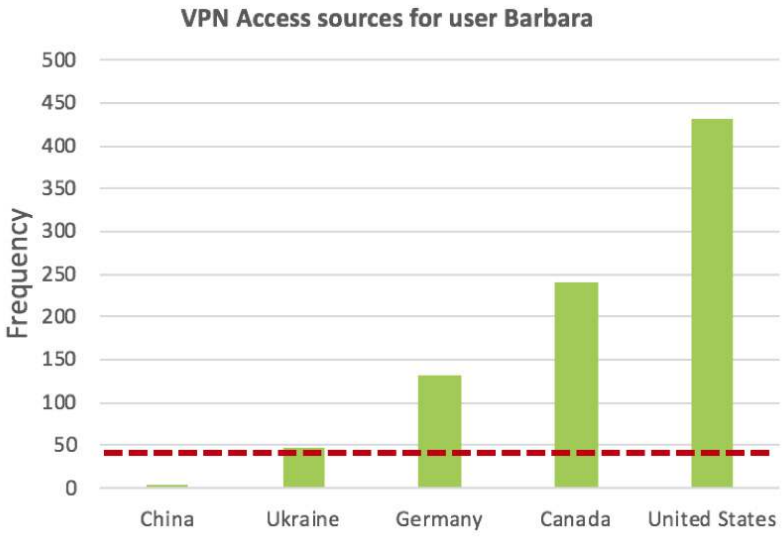

FIGURE 5 – VPN ACCESS LOCATIONS FOR USER BARBARA.

A simple example is to look at whether it's normal for a user, say Barbara Salazar, to have logged in from China. As seen in Figure 5, she regularly connects from the US, Canada, Germany, and Ukraine. Prior to this present connection, she had never connected from China. On the surface, this might be considered anomalous activity. But the models might also take into consideration Barbara's role, her peers and their normal behavior, whether she's an executive or privileged user, the time of day she logs in, when she last logged in, and many other variables. All can be done to determine whether this behavior is notable, as well as to assess its severity of risk.

Models might also analyze whether she is accessing networks she has never accessed before, if she's accessing servers that no one in her peer group uses, and whether Barbara normally only reads files that she is now editing.

Implicit in identifying an anomaly is that previous normal behavior has been captured over time. The models used in Smart Timelines can automatically determine when they have sufficient relevant data to accurately classify a given behavior as anomalous. Rather than train them over fixed amounts of time to predict behavior, user and device norms can typically be set with only a few weeks of data.

Exabeam comes preconfigured with hundreds of prebuilt machine learning-based models, all of which are highly customizable. Customers can also easily create their own or modify existing ones—all without having data science expertise.

Step 7: Detect anomalies

A single anomalous model usually doesn't trigger an alert. Instead, various anomalous models in different dimensions, such as time, location, and activity level, are usually required to trigger an alert. This approach reduces the number of false positives.

Instead, each anomalous model the system applies adds to the overall risk score for a given session. If the total exceeds a predetermined value in Smart Timelines, the system triggers an alert. For example, an alert could be triggered as a result of a user logging in at an unusual time, from an unusual place, and being active on a particularly large or small number of machines.

Instead of static rules that must be maintained, Smart Timelines are heavily weighted toward models to improve analysts' ability to detect unknown threats. But in addition to the underlying Smart Timeline models, you can also use correlation rules. These are typically applied in straightforward situations, such as in monitoring well-known threats, compliance violations, and signature-based threat detection.

An aside: a word about false positives

Mitigating and managing false positive alerts has become a priority for most security teams. Each false positive consumes analysts' time and distracts them from investigating real security incidents.

Smart Timelines dramatically reduces the number of false positives by requiring anomalies in multiple dimensions before it triggers an alert.

The underlying Smart Timelines models also calibrate the score of anomalous events based on contextual factors—including an understanding of roles, groups, and normal behavior. For example, having learned from historical access patterns and other contextual information, the models can reduce false positives when users access a resource for the first time.

Yet system control remains flexible. For example, analysts can mute certain anomaly alerts, such as policy violations they don't consider to be security threats. Smart Timelines also let them mute an entire Smart Timeline session, or particular incidents within a session.

Step 8: Scoring risk due to anomalous activity

With Smart Timelines, user-specific models assign risk scores for those whose observed event patterns sufficiently differ from their own past patterns. Cross-user comparisons are also made to normalize behavior.

Not all event anomalies are given equal importance. Exabeam applies its expertise in providing reference scores for anomalies based on how much security exposure they have. In addition, it uses Bayesian statistics to further highlight unusual anomalies and de-emphasize those that occur more commonly. (Anomalies due to changes in common events, such as a user logging into a device, are typically less interesting than rarer events, such as an account password change or a new user being added.)

Thresholds are established for different types of events. For example, a higher risk score would likely be generated if a user logs into a device (such as a server) for the first time, coupled with a high volume of privileged activities (such as account switching or obtaining privileged access). In this example, the risk score increases if these actions are performed on a critical server.

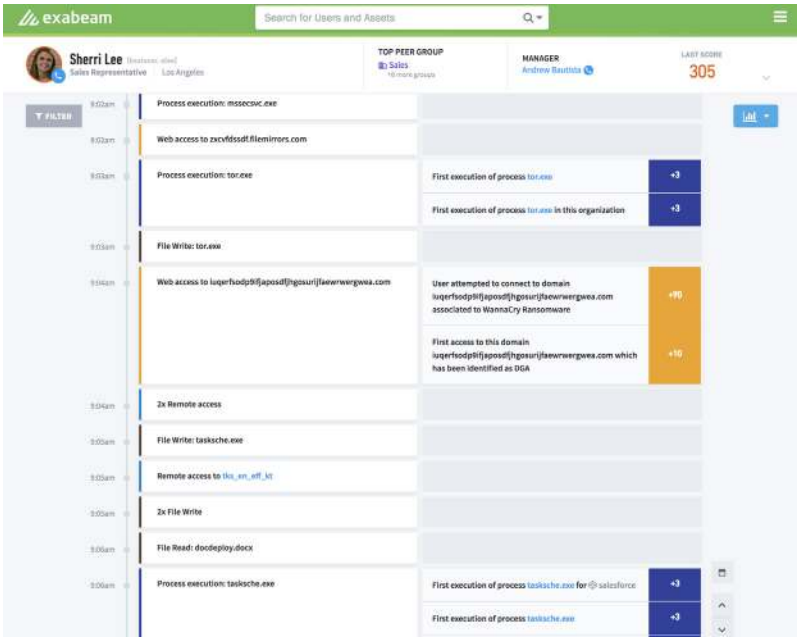

FIGURE 6 – AN EXAMPLE SMART TIMELINE SHOWING A USER'S COMPUTER BEING INFECTED BY RANSOMWARE.

Step 9: Using Smart Timelines for investigations and response

Smart Timelines are designed with analysts in mind. They have a very intuitive user interface (UI) and require no knowledge of data science.

Smart Timelines are designed with analysts in mind. They have a very intuitive user interface (UI) and require no knowledge of data science (Figure 6). This lets more junior analysts do investigations that might have otherwise required a senior analyst's attention.

To conduct an investigation, an analyst can review Smart Timelines for a user or device, clicking on specific events to glean additional information, all the while staying within the single Exabeam UI. An analyst of any level can simply enter the name of a user and a date to instantly see all that user did that day. Moving away from manual investigations reduces their time from days to just seconds. Without any additional effort, analysts can also look at timelines for adjacent days and related users or assets.

Smart Timelines are only the beginning of how Exabeam helps analysts work smarter when it comes to investigations and incident response. If the risk score of a user or device crosses a preset threshold, it's flagged as notable and is prominently listed on the analyst dashboard to help prioritize investigations. An incident involving one user could also prompt an analyst to perform threat hunting for a similar incident across the entire workforce or a group of users.

Smart Timelines also enable Exabeam's security orchestration and automation response (SOAR) solution, Incident Responder, to run playbooks to respond to an incident and automate a response. This means that the importance of Smart Timelines extends to remediation; analysts can only respond to incidents they've detected.

Hunting for complex threats made easier due to Smart Timelines

Analysts can use Smart Timelines to detect attackers' tactics, techniques, and procedures (TTPs) by hunting behavior anomalies. And they're able to search for a specific TTP. They also don't have to write complex queries to look for indicators of compromise (IOCs). An example of threat hunting search criteria might be:

Lateral movement tactics:

- A possible pass the hash attack from the source
- The first account management activity from an asset
- The first account management activity from an asset for a specific user
- The first remote logon to an asset
- A service account that has logged into more than 30 assets

Privilege escalation tactics:

- An account switch to a privileged or executive account
- A non-executive user logon to an executive asset
- An abnormal addition to a privileged group by user

SUMMARY

Not all timelines are created equal. Smart Timelines are not just a collection of logs sorted by their timestamp. For a timeline to be "smart," it needs several key attributes:

- It must be able to merge and normalize raw, often cryptic logs into easy-to-understand, human-readable events. This is particularly true for AD events.
- It must show user and machine events regardless of log sources. To do so. the underlying process must convert logs fields into user activities. This is usually not a one-to-one translation of a log field into an event, but involves combining logs and adding context.
- It must be able to determine the beginning and end of user or entity activity, mapping it to sessions so as to normalize behavior.
- It must annotate each event with possible anomalies associated with it (relative to the user or entity history), and do so in a way that minimizes false positives.
- It must be able to identify and group all events belonging to a user regardless of whether the person moved laterally through various devices or switched identities.
- They must present and score events that exhibit anomalous user behaviors.

exabeam

ABOUT EXABEAM

Exabeam is the Smarter SIEM™ company. We empower enterprises to detect, investigate, and respond to cyberattacks more efficiently so their security operations and insider threat teams can work smarter. Security organizations no longer have to live with excessive logging fees, missed distributed attacks and unknown threats, or manual investigations and remediation. With the Exabeam Security Management Platform, analysts can collect unlimited log data, use behavioral analytics to detect attacks, and automate incident response, both on-premises or in the cloud. Exabeam Smart Timelines™, sequences of user and device behavior created using machine learning, further reduce the time and specialization required to detect attacker tactics, techniques, and procedures.
https://www.exabeam.com.

TO LEARN MORE ABOUT HOW EXABEAM CAN HELP YOU, VISIT EXABEAM.COM TODAY.