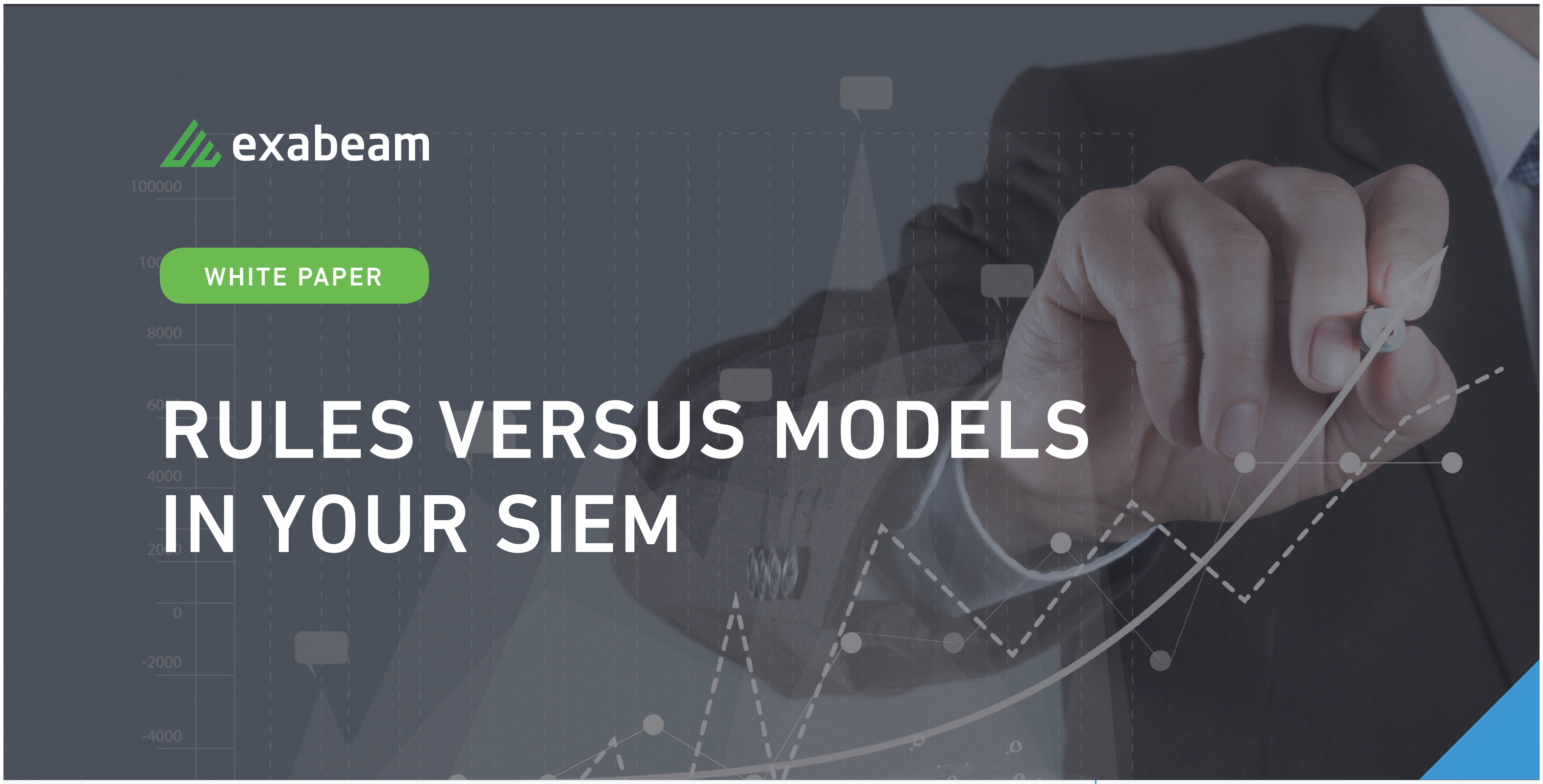


exabeam

WHITE PAPER

RULES VERSUS MODELS IN YOUR SIEM

INTRODUCTION

There has been a rapid increase in malicious insider threats, compromised insiders, and sensitive data exfiltration targeting enterprises today. Organizations need technologies to protect themselves from such threats and to be certain their SOC is compliant with regulations.

Security Information and Event Management (SIEM) technologies have been used for years to detect threats and to address compliance requirements for organizations. Many SIEM tools' detection methodologies are primarily based on correlation rules that look for known attacks at the points of entry. Such rules become increasingly ineffective as attacks become more complex, longer lasting, or more distributed. Next-gen SIEM tools are behavior and context aware, and models are used to track user behaviors, which makes it very effective to detect unknown threats and complex attack chains.

In the next few sections, will discuss rules and models, the pros and cons of using them, and how to design and build effective rules and models.

What is a Correlation Rule?

Correlation rules compare incoming events with predefined relationships between entities to identify anomalies. They are essentially a set of conditions that function as a trigger and they're created based on the existing knowledge of an attackessentially a series of known steps that can be detected. Think about correlation rules trigger such as "send an alert whenever A, B, and C happen within 30 minutes."

Here are a few more realistic correlation rule examples:

- If a user fails more than three login attempts to the same machine within an hour, then alert.
- If more than 10 failed logon attempts are followed by one successful logon, then alert.

Rules can be divided into

- Simple rules that look for a certain event type and trigger an alert. They can include the use of lookups against a watchlist. For example: if an event uses IGMP transport protocol and it's source IP address is 73.230.79.1, then alert.
- Complex/Composite rules that join two or more rules or nest rules within a rule. For example: detect the presence of events and looking for a pattern in a threat.

What is a Model?

A model profiles a given user or asset behavior on a particular aspect of interacting with the corporate or IT environment. Examples: "what time an employee badges in and out of the office," "the number of assets the user is normally logged on to," "users who VPN in after badge access," "operating system the user logs on to". When using models, you begin by baselining activities, then identify deviations from the baseline. This is a key benefit of using models.

Examples of models:

- If a user has switched to a privileged account and undertakes an abnormal data upload to external services, then alert.
- If a user VPNs from a location for the first time and accesses executive file shares, then alert.

WHEN DO YOU USE CORRELATION RULES VS. MODELS?

Here are use cases that describe when to use rules or models.

USE CASES	WHAT TO USE
REAL-TIME MONITORING OF KNOWN THREATS Many threats that infiltrate organizations are well documented; hackers use the same methods with a few small tweaks. Rules can easily detect such attack patterns.	Correlation Rules
COMPLIANCE VIOLATION CHECKS Many data security regulations, e.g., GDPR, PCI DSS, HITECH, require organizations to demonstrate effective controls. These regulations have well documented security controls and rules can help with compliance checks. Example: Alert if anti-virus is disabled on our PCI systems.	Correlation Rules
SIGNATURE-BASED THREAT DETECTION When a malware is detected, it's signature is added to a repository. These repositories contain hundreds of millions of signatures that identify malwares. For example, companies like VirusTotal aggregates malware signatures that can be accessed by vendors. Rules can be used to detect known malware signatures.	Correlation Rules
BEHAVIOR-BASED ANOMALY DETECTION Dynamic environments regularly face technological trends, e.g., bring your own technology (BYOT) and corporate data in the cloud. This means hackers discover more points of entry to your organization, making it increasingly important to track both normal and abnormal user behavior. Models can be used to track behaviors unlike correlation rules that are not designed to track user behavior.	Models
CORPORATE DATA EXFILTRATION DETECTION Data exfiltration is the unauthorized transfer of data from inside your organization to the outside. Hackers access targeted machines through remote applications or by installing a portable media device. Advanced persistent threats (APTs) are one form of cyber attack in which data exfiltration is often a primary goal. Data exfiltration might involve movements across assets, privileged account access, employees, peer-groups, and these are best detected using Models.	Models
ZERO-DAY THREATS Zero-day threats haven't been encountered yet so a rule cannot be written to identify them. They may involve an unknown mix of anomalous lateral movements, abnormal/remote logins, file access, and/or abnormal data uploads. A modeling approach is useful here because it can easily identify these threats based on deviations from behavioral baselines.	Models
LATERAL MOVEMENT DETECTION Lateral movement is widely used in cyber attacks to access hosts from a compromised system, then from there get access to sensitive data, shared files, and/or privileged credentials. The latter can be further leveraged to access more resources, further elevate privileges, and steal even more valuable credentials. Lateral movement evades detection via correlation rules because parts of the attack may be present in different IP addresses, identities, and machines. Models are able to identify and correlate the anomalous activity across different systems to detect attacks.	Models

PROS AND CONS OF CORRELATION RULES

PROS OF CORRELATION RULES

- **Decrease response times** for routine or known attacks.
- **Detect risks** by correlating relationships between resources.
- **Provide real-time monitoring** of known threat vectors
- **Easy to express rules** – Traditional SIEMs have made it easy to build and express policies through a rule builder.

Cons of Correlation Rules

- **High false-positive rate** – Constant IT environment changes require frequent rules updates by analysts/ consultants. Incorrect correlation rules and a lack of user or asset context(departments, job function, peer groups, roles) can trigger hundreds of false-positive security alerts.
- **Applicable only for known attack patterns** – Correlation rules are best for creating policies for known patterns; they're not suitable for unknown attack chains.
- **Long nested rule execution** – Rule execution can be time consuming. Nested rules further exacerbate the issue. Refer to this post for more information: https://blogs.gartner.com/augusto-barros/2017/03/31/siem-correlation-is-overrated/
- **Constant rule maintenance** – Enterprises introduce new IT products to their environments and patch new releases. To keep up-to-date, correlation rules have to be checked constantly and tailored accordingly.

PROS AND CONS OF MODELS

Pros of Models

- **Baseline behavior to detect anomalies** – Tracking normal user or asset behavior for a duration establishes a baseline, enabling you to more easily recognize abnormal activity. An effective product predominantly uses models to baseline behavior, also factoring in expert knowledge and contextual information about your organization
- **Track lateral movement and look for abnormalities** – Models can track users who've abnormally accessed file servers, monitor their login activities, track administrative assets, and track service accounts to detect lateral movement.
- **Detect unknown threats** – Deviations from the baseline are tracked for abnormalities. Various models can be tied together in a user timeline to provide analysts with the complete story regarding an attack chain. Example: Detecting an advanced threat involving account privilege elevation, lateral movement, administrative asset access, and data exfiltration.
- **Ability to use contextual data for effective anomaly detection** – To effectively detect anomalies, models leverage contextual data, e.g., user hostname, peer group behavior data, user type (executive, administrator, service account), and departmental user data. And they can take in information around users and entities such as defacto asset owners, normal VPN access time zones, top network users, and folders containing source code, et al.

Cons of Models

- **User baselining takes time** – Baselining requires a certain amount of time so the analytics engine can model normal user behavior. A well constructed baseline helps to effectively and quickly detect threats.
- **May require lots of professional services if not well engineered within the products** – An unknown threat usually involves combination of abnormal behavior like anomalous lateral movement, abnormal remote logins, access of administrative assets, account switching, abnormal file access, and unusual data uploading. All relevant events have to be stitched together to have the complete story regarding an attack. Data structures also need to be in place to prioritize risks based on anomalies.

  Evaluate products that provide out-of-the-box features with effective use of models to detect unknown threats, solve your use cases and provides quick time to value, without customizations.

DESIGNING RULES AND MODELS

Correlation Rule Builder

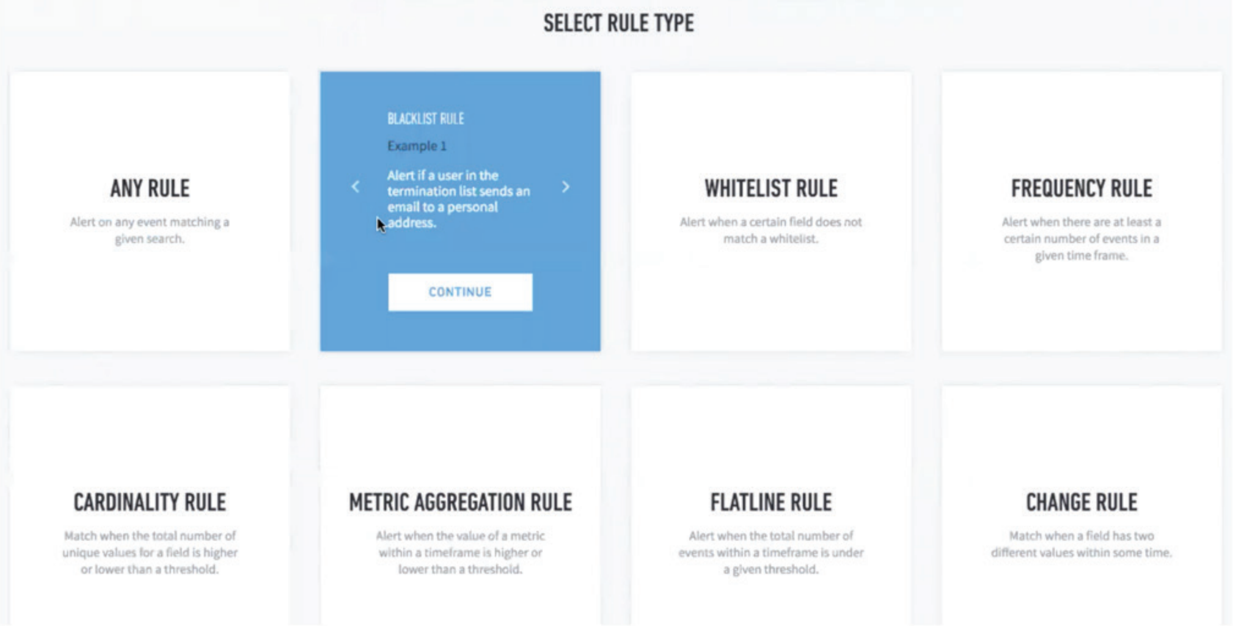
Many SIEM tools provide rule builders within their UI, making it easy for administrators to pick up key parameters from various events and build rules. A rule builder typically contains categorized lists of all components you can use during configuration.

Correlations can be created by dragging items from events and their associated alerts so as to prompt an action. For example, if a rule set pertains to a threat, you can either set up many correlations or nest all into one.

Modern IT environments being complex, building rules this way can be daunting and time consuming. But next-gen SIEM tools like Exabeam use natural language for correlation rule building. Its rule-building wizard makes it possible for even junior analysts to create complex rules.

Here is a typical rule building sequence:

1. Select a rule type.



SELECT RULE TYPE

ANY RULE
Alert on any event matching a given search.

BLACKLIST RULE
Example 1
Alert if a user in the termination list sends an email to a personal address.
CONTINUE

WHITELIST RULE
Alert when a certain field does not match a whitelist.

FREQUENCY RULE
Alert when there are at least a certain number of events in a given time frame.

CARDINALITY RULE
Match when the total number of unique values for a field is higher or lower than a threshold.

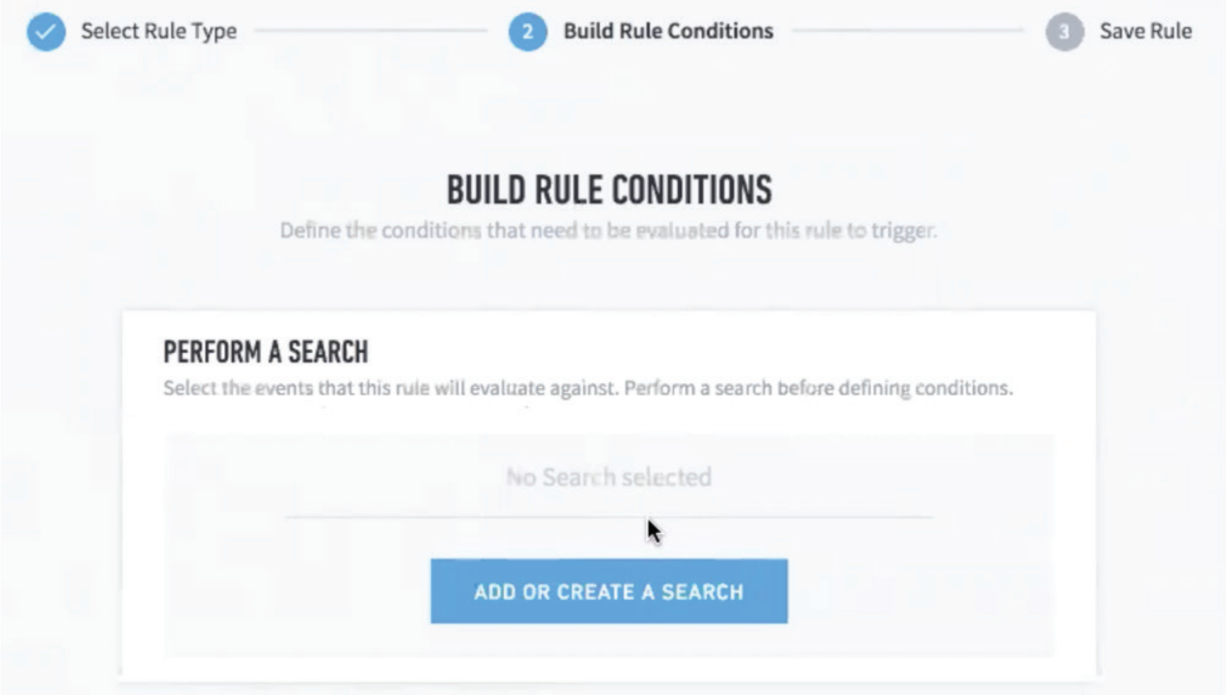
METRIC AGGREGATION RULE
Alert when the value of a metric within a timeframe is higher or lower than a threshold.

FLATLINE RULE
Alert when the total number of events within a timeframe is under a given threshold.

CHANGE RULE
Match when a field has two different values within some time.

FIGURE 1 – RULE TYPES TO CHOOSE FROM OUR CORRELATION RULE BUILDER

2. Add a search query to fetch all events against which the rule can be evaluated.



BUILD RULE CONDITIONS

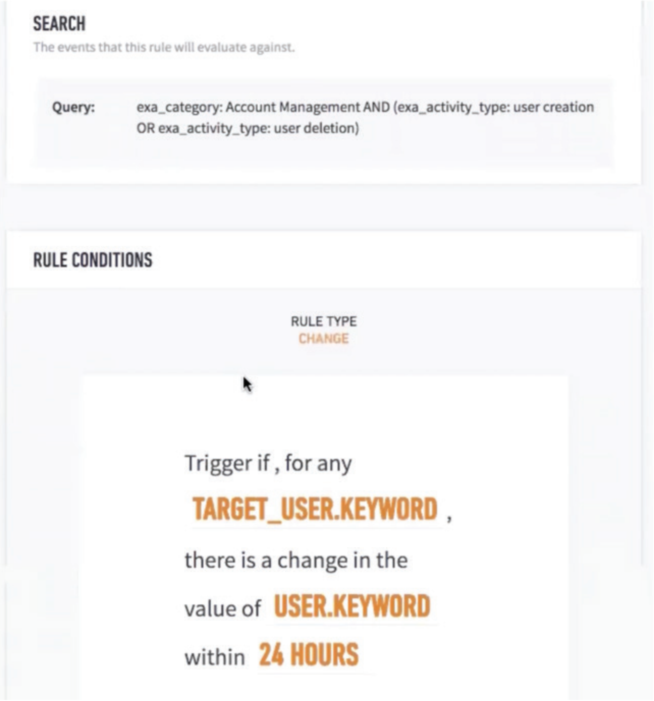
Define the conditions that need to be evaluated for this rule to trigger.

PERFORM A SEARCH

Select the events that this rule will evaluate against. Perform a search before defining conditions.

FIGURE 2 - ABILITY TO ADD OR CREATE SEARCH CRITERIA TO FETCH THE EVENTS

3. Add rule conditions. Here is an example of a change rule type and the conditions that would trigger it.



SEARCH

The events that this rule will evaluate against.

Query: exa_category: Account Management AND (exa_activity_type: user creation OR exa_activity_type: user deletion)

RULE CONDITIONS

RULE TYPE
CHANGE

Trigger if , for any TARGET_USER.KEYWORD , there is a change in the value of USER.KEYWORD within 24 HOURS

FIGURE 3 – EXAMPLE OF A SEARCH QUERY TO SELECT THE EVENTS AND RULE CONDITIONS TO APPLY TO THE EVENTS

Design Considerations for Models

The key reasons for creating models are to track users behavior and associated context. But before addressing user behavior, let's first examine context.

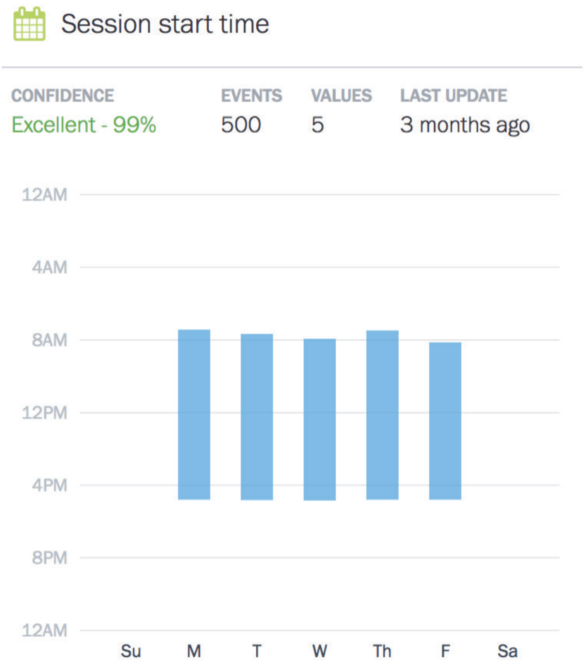
CONTEXT - WHY IS IT IMPORTANT?

Security logs are sourced from various entry points, e.g., Windows servers, VPNs, firewalls, endpoint devices, DNS servers, and are aggregated to enable threat discovery. Logs tell you what users and entities are doing (actions), while context informs who the users and entities are (identity), what their roles are, and how they normally behave. For example, you know the source IP address of a user's workstation from the logs, but if the user VPNs from a different location, you can track the user by mapping the IP address to the hostname.

Adding contextual data enriches security events, making it easier to uncover threats. Other context enrichment examples include classifying service accounts in addition to servers and workstations, as well as the identification of user peer groups, contractors, and privileged accounts.

MODELLING USER BEHAVIOR (TIMELINE)

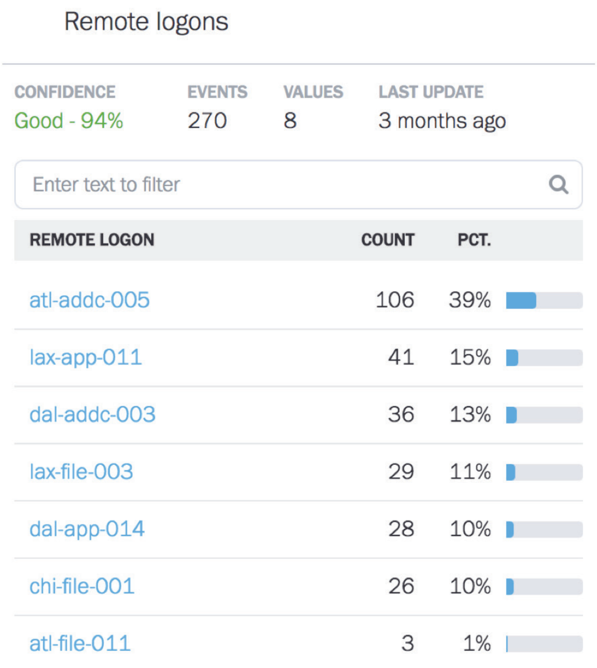
Enriched security events makes it easier for your SOC team to track user behavior. For example, you can model a user's working hours, the remote logon servers they access, the zones they log in from, their VPN locations, and more. This assists in baselining normal user behavior over time; deviations from the baseline help detect abnormalities.



Session start time

CONFIDENCE	EVENTS	VALUES	LAST UPDATE
Excellent - 99%	500	5	3 months ago

FIGURE 4 - USER MODEL SHOWS NORMAL WORKING HOURS



Remote logons

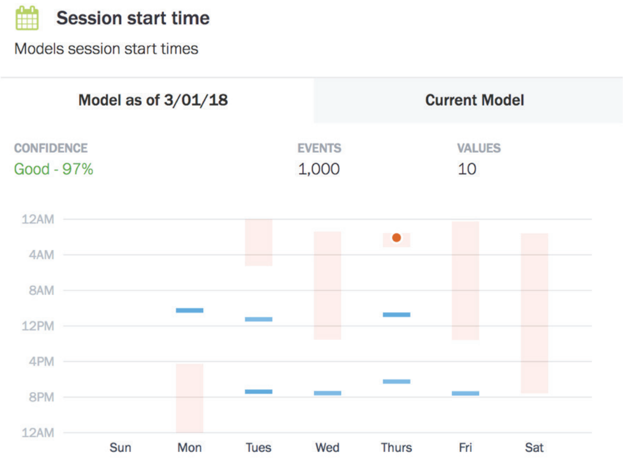
CONFIDENCE	EVENTS	VALUES	LAST UPDATE
Good - 94%	270	8	3 months ago

REMOTE LOGON	COUNT	PCT.
atl-addc-005	106	39%
lax-app-011	41	15%
dal-addc-003	36	13%
lax-file-003	29	11%
dal-app-014	28	10%
chi-file-001	26	10%
atl-file-011	3	1%

FIGURE 5 – USER MODEL SHOWS TYPICAL REMOTE LOGIN LOCATIONS (ATLANTA, LOS ANGELES, DALLAS AND SO ON)

USING MODELS TO DETECT THREATS

Any abnormal activities are tracked and given a risk score. When risk scores cross a pre-defined threshold they are escalated to a security analyst for investigation. This provides the natural benefit of combining many discrete alerts into a single object that an analyst can review at once, instead of reviewing many individual events and trying to piece them together.



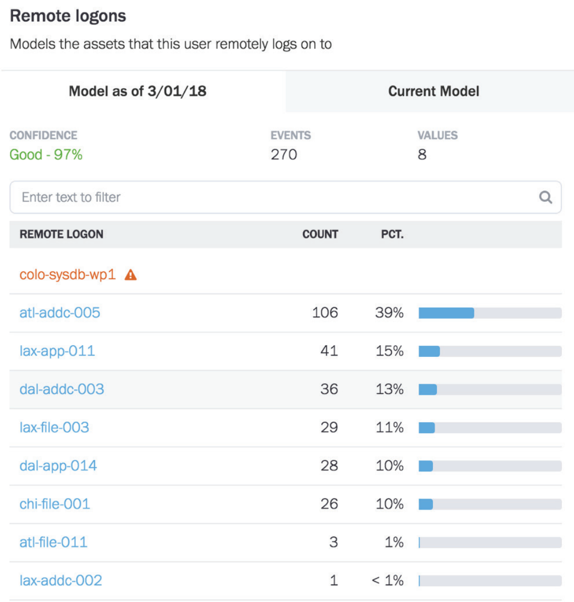
Session start time

Models session start times

Model as of 3/01/18	Current Model

CONFIDENCE	EVENTS	VALUES
Good - 97%	1,000	10

FIGURE 6 – USER MODEL FLAGGING 4 AM LOGIN TIME AS ABNORMAL



Remote logons

Models the assets that this user remotely logs on to

Model as of 3/01/18	Current Model

CONFIDENCE	EVENTS	VALUES
Good - 97%	270	8

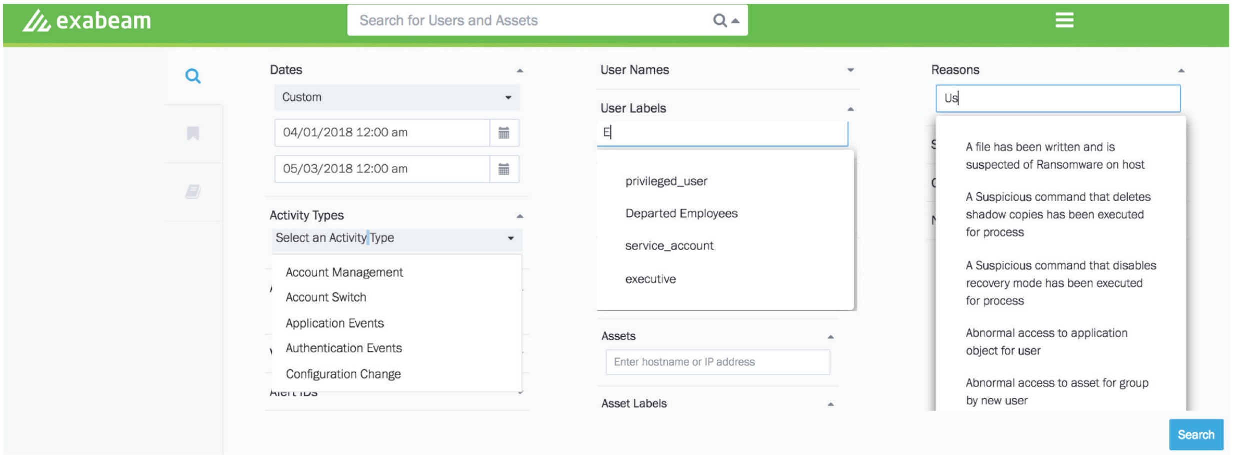
Enter text to filter

REMOTE LOGON	COUNT	PCT.
colo-sysdb-wp1 ⚠		
atl-addc-005	106	39%
lax-app-011	41	15%
dal-addc-003	36	13%
lax-file-003	29	11%
dal-app-014	28	10%
chi-file-001	26	10%
atl-file-011	3	1%
lax-addc-002	1	< 1%

FIGURE 7 – USER MODEL SHOWS FLAGGING FIRST-TIME REMOTE LOGIN FROM COLORADO AS A RISK

Ad-Hoc Threat Hunting

By tracking normal/abnormal user behavior, being able to hunt for potential threats across your IT environments becomes far more effective. You're able to apply the context data, risk reasons, and activity types to your threat search criteria, essentially creating a very complex query structure that remains simple for analysts to use.



Dates

Custom

04/01/2018 12:00 am

05/03/2018 12:00 am

Activity Types

Select an Activity Type

- Account Management
- Account Switch
- Application Events
- Authentication Events
- Configuration Change

Alert IDs

User Names

User Labels

E

- privileged_user
- Departed Employees
- service_account
- executive

Assets

Enter hostname or IP address

Asset Labels

Reasons

Us

- A file has been written and is suspected of Ransomware on host
- A Suspicious command that deletes shadow copies has been executed for process
- A Suspicious command that disables recovery mode has been executed for process
- Abnormal access to application object for user
- Abnormal access to asset for group by new user

FIGURE 8 – THREAT HUNTER SEARCH QUERY EXAMPLE SHOWING ACTIVITY TYPES, USER TYPES, AND RISK REASONS

Leveraging Models for Custom Use Cases

Machine learning algorithms can leverage models for threat detection. Here is an example of how you can extend models to customize your algorithms.

- **Daily Activity Change Detection** – In this model, Exabeam's User Behavior Analytics engine tracks overall daily user activities (e.g., Windows login, VPN login, web activity, file sharing) looking for activity pattern anomalies. If there is a significant change in any user pattern, alerts are triggered and added to that user timeline for further investigation.

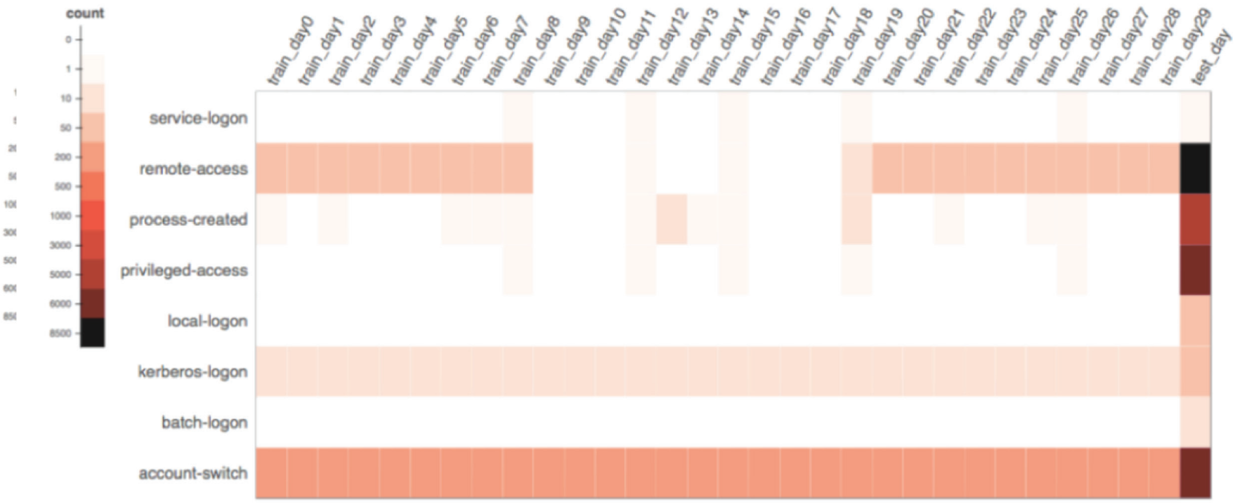


FIGURE 9 - DAILY ACTIVITY TRACKING SHOWING ABNORMAL REMOTE ACCESS, PRIVILEGED ACCESS, AND ACCOUNT SWITCH ON DAY 30

WHEN SHOULD YOU RETIRE RULES?

Many organizations have difficulty supporting traditional SIEM tools, as they might not have the time, funds, resources, or processes to support rules. For them, they need to reassess the long-term value of traditional SIEM maintenance and support.

If your organization faces any of the rules-based tool challenges listed below, it is time for them to retire rules and consider modern behavior and context-aware SIEM tools.

- **Time and resources spent on false positives**
  If a correlation rule is wrong or produces too many false positives, analysts have to reanalyze the data and tune the rule. This consumes support, resources, and time.
- **Rule maintenance becomes burdensome**
  Organizations rely on data collection and retention for correlation purposes. To ensure the fidelity of its correlation logic, it has to verify its custom correlations every time there is an environment change. Generally, this requires engaging with professional services and increases operational overhead.
- **Ineffective alerts and reporting due to dynamic environment**
  Today's attacks are more advanced. And infrastructure is dynamic with BYOD, semi-managed devices, and corporate data residing in the cloud. If your rules aren't effective in detecting threats, it's time to evaluate model-based tools that are behavior and context aware.

HOW DO YOU MIGRATE RULES?

Follow these steps to successfully migrate your rules.

- **In-house SIEM expertise** – Use internal cybersecurity experts who have extensive knowledge of your current IT environment and SOC operations. Review your use cases, log sources, and security gaps with current SIEM deployments.
- **Validate your current rules** – Review your current rules set. List all rules that need to be retired vs. those that remain valid.
- **Engage with migration or professional services** – Map current use cases and document new ones, if any. Enlist SIEM consultants to help migrate valid rules

DECIDING BETWEEN RULES AND MODELS

Behavior and context-enriched analytics tools that use models are effective to detect threats across your IT environments. To recap, model benefits include the ability to detect previously unknown threats, detect user behavior abnormalities, detect lateral movements across your environment, and the ability to perform threat hunting armed with context-aware data.

Models can also be used to detune a useful, but potentially noisy, correlation rule, e.g., an interactive logon by a service account—a rule provided by many SIEMs. But when deployed at at an organization, there may be some scripts that require interactive logon to function and that might potentially trigger thousands of alerts—flooding the SOC. The logic can be implemented within a model—to first learn which service accounts normally perform interactive logons to specific assets, then trigger on anomalous occurrences of the pattern.

Correlation Rules are useful in detecting known threat vectors; they can keep critical systems in check and are valuable. But if your rules have missed anomalies that should have been detected, or some are disabled for reasons unknown, you might assess whether those anomalies detection could have been caught by models.

exabeam

ABOUT US

Exabeam provides security intelligence and management solutions to help organizations of any size protect their most valuable information. The Exabeam Security Intelligence Platform uniquely combines a data lake for unlimited data collection at a predictable price, machine learning for advanced analytics, and automated incident response into an integrated set of products. The result is the first modern security intelligence solution that delivers where legacy security information and event management (SIEM) vendors have failed. Built by seasoned security and enterprise IT veterans from Imperva, ArcSight, and Sumo Logic, Exabeam is headquartered in San Mateo, California. Exabeam is privately funded by Lightspeed Venture Partners, Cisco Investments, Norwest Venture Partners, Aspect Ventures, Icon Ventures, and investor Shlomo Kramer.

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