Is It Normal or Suspicious? Detecting Anomalies via Market Basket Analysis

Zhuxuan (Nancy) Jin
Data Scientist | SplunkUBA

Forward-Looking Statements

During the course of this presentation, we may make forward-looking statements regarding future events or plans of the company. We caution you that such statements reflect our current expectations and estimates based on factors currently known to us and that actual events or results may differ materially. The forward-looking statements made in the this presentation are being made as of the time and date of its live presentation. If reviewed after its live presentation, it may not contain current or accurate information. We do not assume any obligation to update any forward-looking statements made herein.

In addition, any information about our roadmap outlines our general product direction and is subject to change at any time without notice. It is for informational purposes only, and shall not be incorporated into any contract or other commitment. Splunk undertakes no obligation either to develop the features or functionalities described or to include any such feature or functionality in a future release.

Ping Jiang

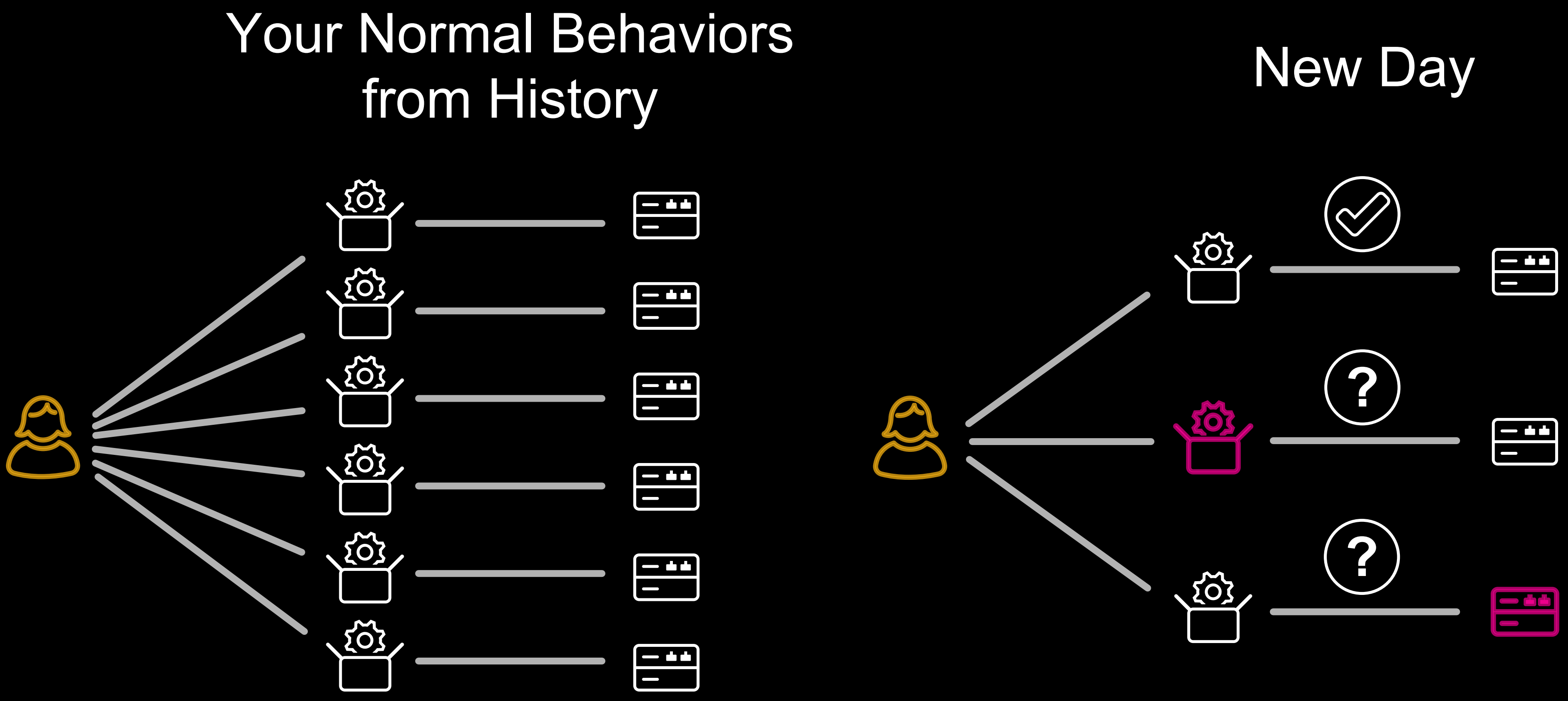
Senior Software Engineer | SplunkUBA

What Makes You You?

Is your behavior normal or suspicious?
What is the real you?

Mining Your Behaviors

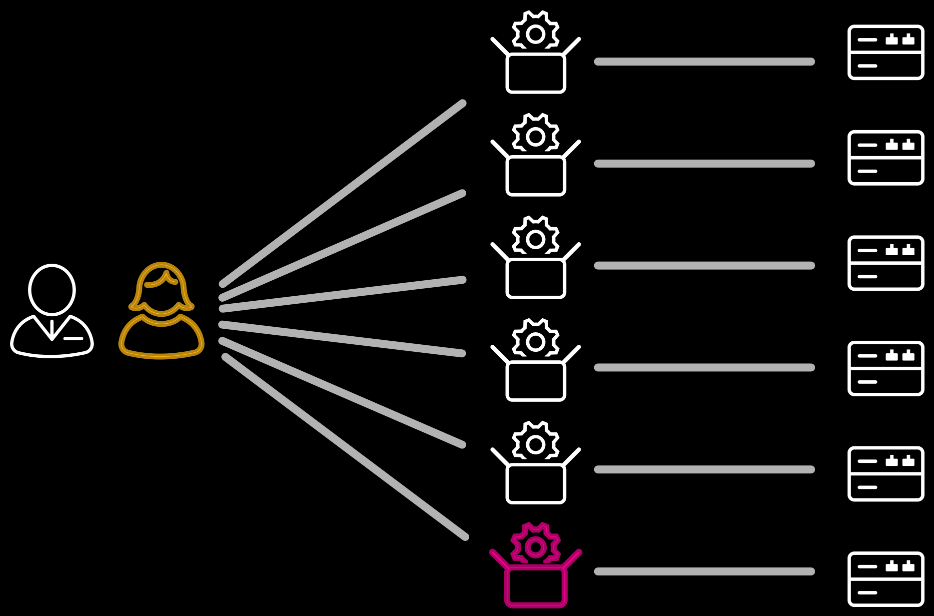
Compared to your normal behavior, what makes a new day suspicious?



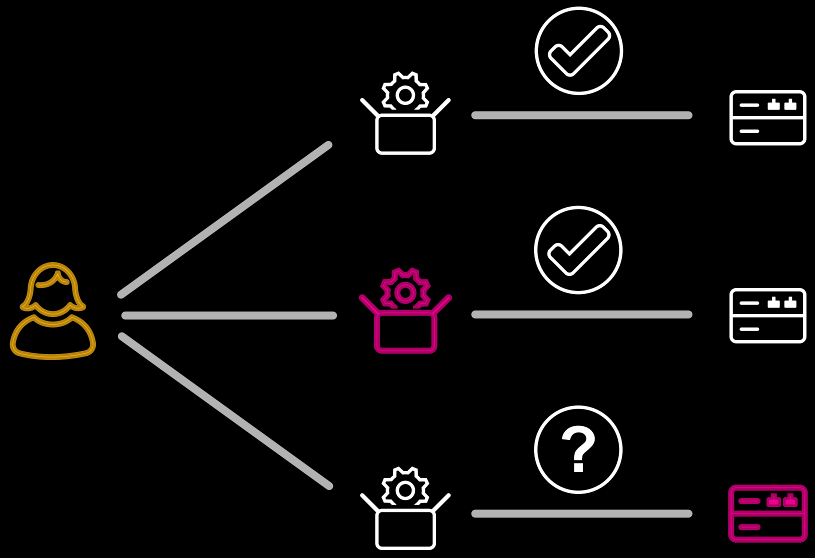
Mining Your Behaviors

What do your peers normally do?

Your Peers
Normally Behave like This

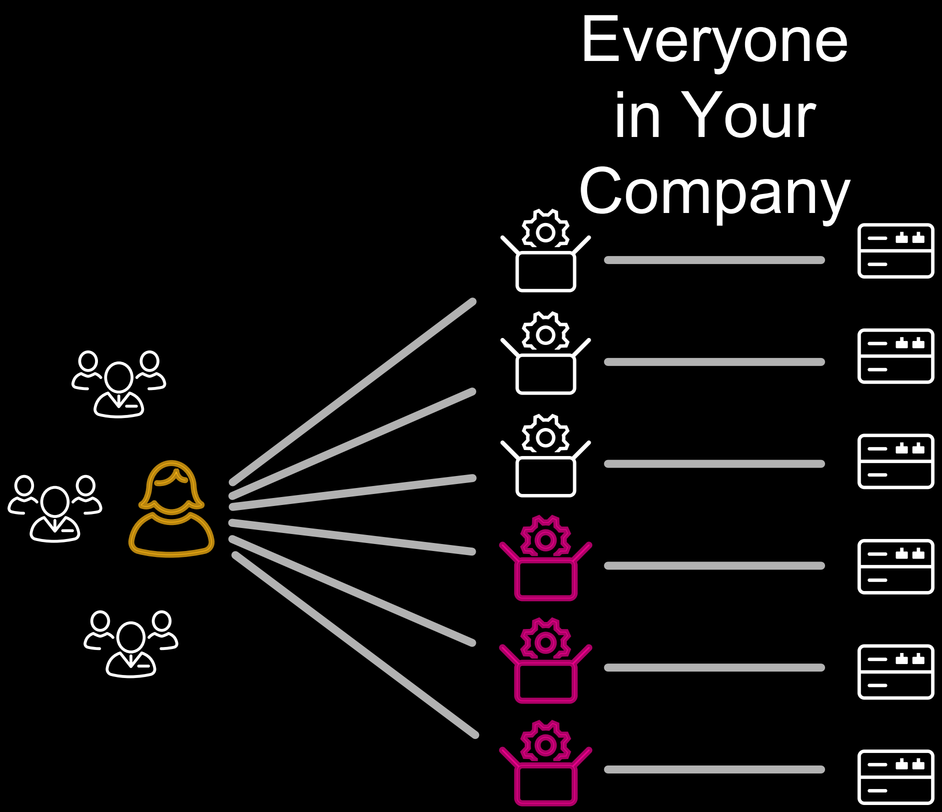


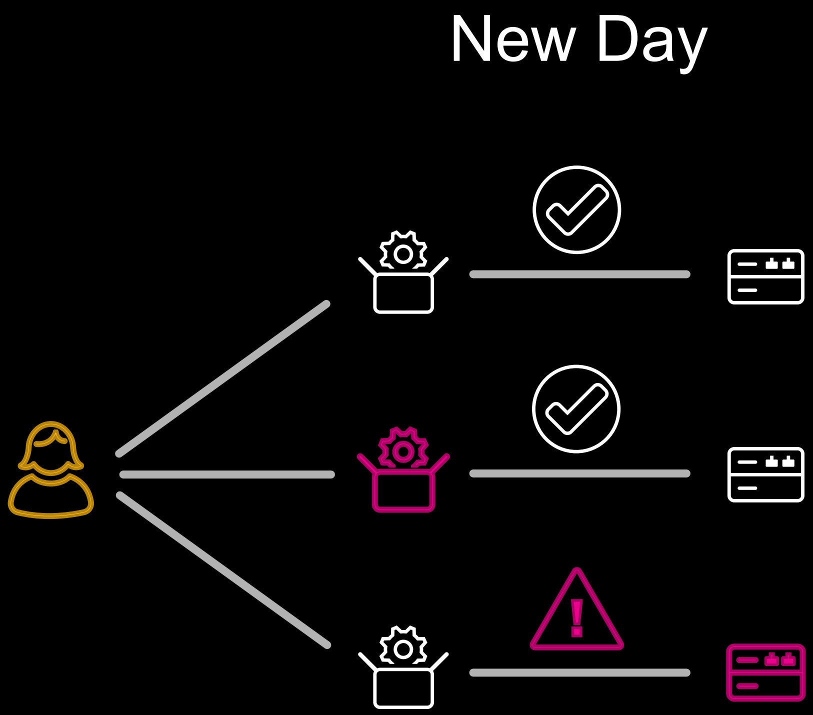
New Day



Mining Your Behaviors

How’s everyone?





Normal vs. Suspicious

Does it follow history or far from history?

New Events

- Normal = what follows entity's historical behavior
  - Routine / pattern / frequently
  - Entity can be account, device
  - Various scopes of history: entity's own, peergroup, everyone else

NO Clear Separation

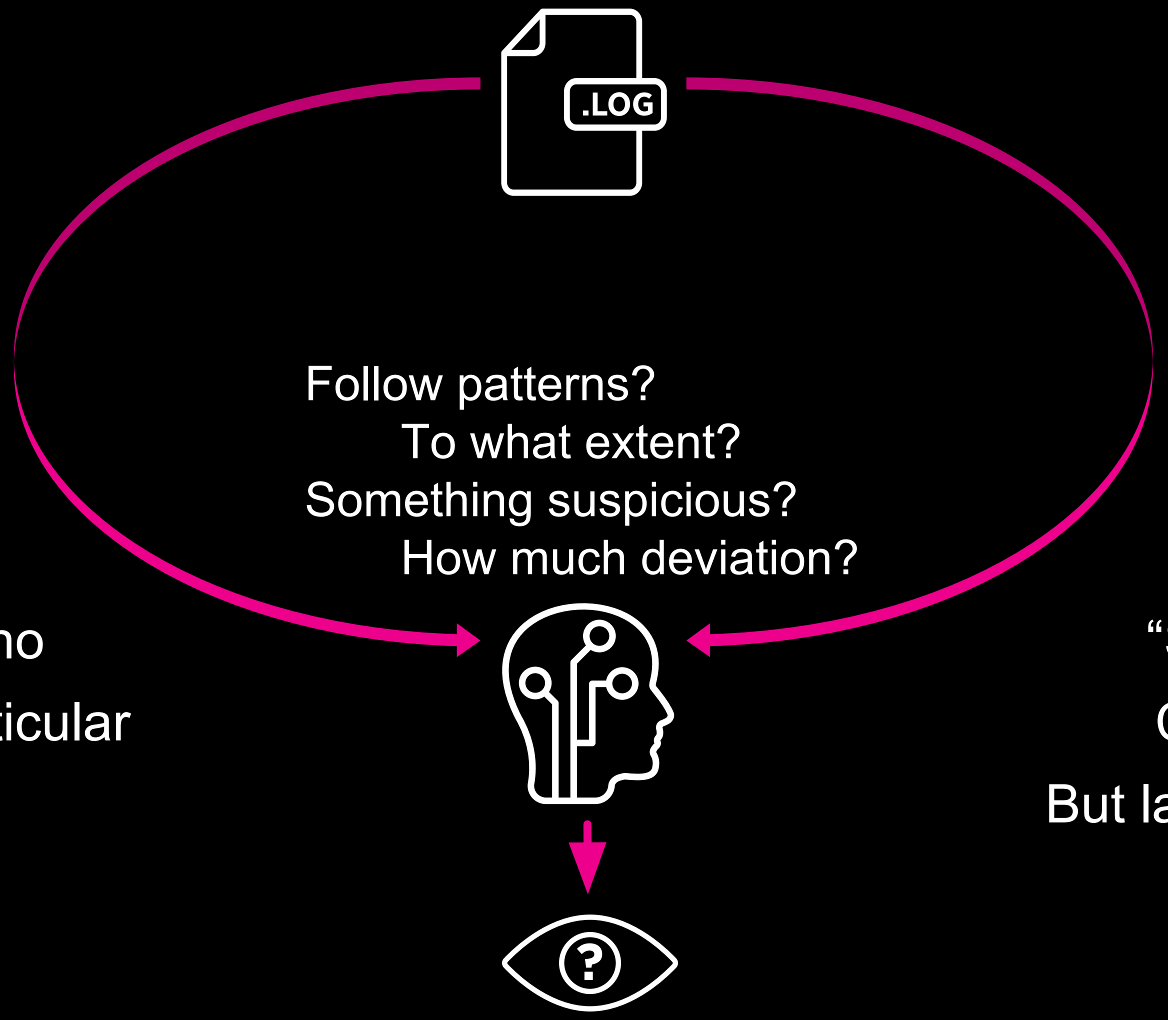
- Suspicious = what is far from history
  - Anomalous / unusual / rarely observed or never happened
  - Less suspicious / more suspicious

Frequent Patterns

Combine Weapons

Let’s create an engine to quantify the ‘suspiciousness’

Rule Engine



Machine Learning

”Anomaly?” yes or no

Simple, fast, target particular use case;

But cannot detect

“unknown”s

“Suspicious?” maybe

Complex, intelligent;

But lack of interpretability and expensive

The Model

Flexible framework for pattern mining, event scoring and anomaly detection inspired by Market Basket Analysis

A Motivating Example

What are frequent and what does a score look like?

- The Market Basket Problem[1]
  - Given transaction logs, mining through a customer's purchase behavior

History Transactions (Itemset)
{Bread, Milk}
{Bread, Milk, Diaper, Beer, Eggs}
{Milk, Diaper, Beer, Coke}
{Bread, Milk, Diaper, Coke}
{Bread, Milk, Beer, Coke}
...

Frequent Itemset Mining

Frequent Patterns
{Bread}
{Milk}
{Milk, Diaper}
{Bread, Milk}
{Bread, Milk, Diaper}
{Milk, Beer, Diaper}
...

New Transactions (Itemset)
{Bread, Milk, Diaper, Beer}
{Bread, Egg, Cookies}

→ match patterns → less suspicious ← low score

→ not match → more suspicious ← high score

[1] Kuchar, J., & Svátek, V. (2018, January). Spotlighting anomalies using frequent patterns. In KDD 2017 Workshop on Anomaly Detection in Finance (pp. 33-42)

Same Idea Applied in Security

What is a normal event and how much suspiciousness it is?

- Security Use Case
  - Given Windows logs, mining a user's normal authentication behavior

History Logs (Itemsets)
{Network, myLogonProcess, myDevice}
{Network, myLogonProcess, myApp, myDevice, myAccountName}
{myapp, myDevice, myAccountName}
{Network, myLogonProcess, anotherApp, myDevice, myAccountName}
{myLogonProcess, anotherApp, myAccountName}
...

Frequent Itemset Mining →

Frequent Patterns
{myLogonProcess}
{myDevice, myLogonProcess}
{myDevice, myApp}
{myAccountName, myDevice}
{myAccountName, myDevice, myApp}
{myAccountName, myDevice, anotherApp}
...

New Authentications (Itemsets)
{myAccountName, myDevice, myApp}
{firstObsDevice, rarelySeenApp}

→ match patterns → less suspicious ← low score

→ not match → more anomalous ← high score

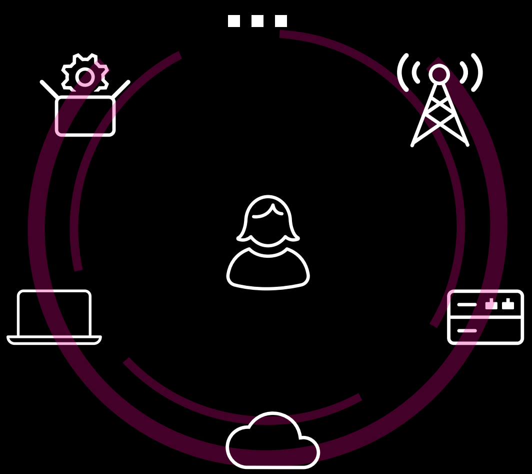
Model Overview

Pre-process the historical logs and learn the frequent patterns

- Historical events profiled as a combination of different field values extracted by proper Core Splunk Queries
- Frequent patterns generated automatically

Oct 10 23:12:00 1,2019/10/10 23:12:00,,TRAFFIC,end,1,2019/10/10 23:12:00, userDomainName, deviceName, 0.0.0.0,0.0.0.0,PAN-Agent- Access, NTLM(authentication type), logonProcess, sourceZoneName, Feature1Value1, Feature2Value1, Feature3Value1, xx.xxxx,, xxxx...
Oct 11 23:12:14 1,2019/10/11 23:12:14,,TRAFFIC,end,1,2019/10/11 23:12:14, userDomainName, deviceName, 0.0.0.0,0.0.0.0,PAN-Agent- Access,, duosecurity(application), ,GuestZone(zone name),... Feature1Value2, Feature2Value2, Feature3Value2...
Oct 11 23:12:50 1,2019/10/11 23:12:50,,TRAFFIC,end,1,2019/10/11 23:12:50, userName, device name,0.0.0.0,0.0.0.0,PAN-Agent- Access,,, Kerberos(authenticationPackage)... Feature1Value3, Feature2Value2,

Frequent Patterns



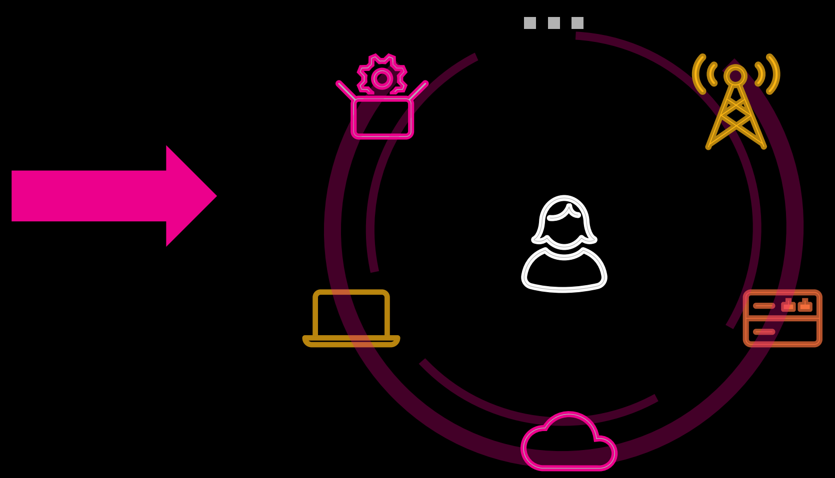
Model Overview

Pre-process the new events in a same procedure

- New events profiled in the same procedure like historical events
- Prepare for scoring

Oct 12 23:52:53 1,2019/10/12 23:52:53,,TRAFFIC,end,1,2019/10/12 23:52:53, username, deviceName, 0.0.0.0,0.0.0.0,PAN-Agent- Access,,,active-directory(applicationName),, sourceZoneName,destinationZoneName,xxx.xxx,xx.xxxx,,2019/10/12 23:52:53,x.x.x.x,tcp,allow, 2019/ 10/12 23:52:20,31,any,0, Feature1Value1, Feature2Value2, Feature3Value1 xxxx...
...

New Event



Model Overview

Equipped with frequent patterns, let's do scoring by matching

- Itemset

  {a server that has never been accessed by the user before. a blacklist application is used for access
  a rarely seen device is used, access comes from a commonly seen source zone ...}

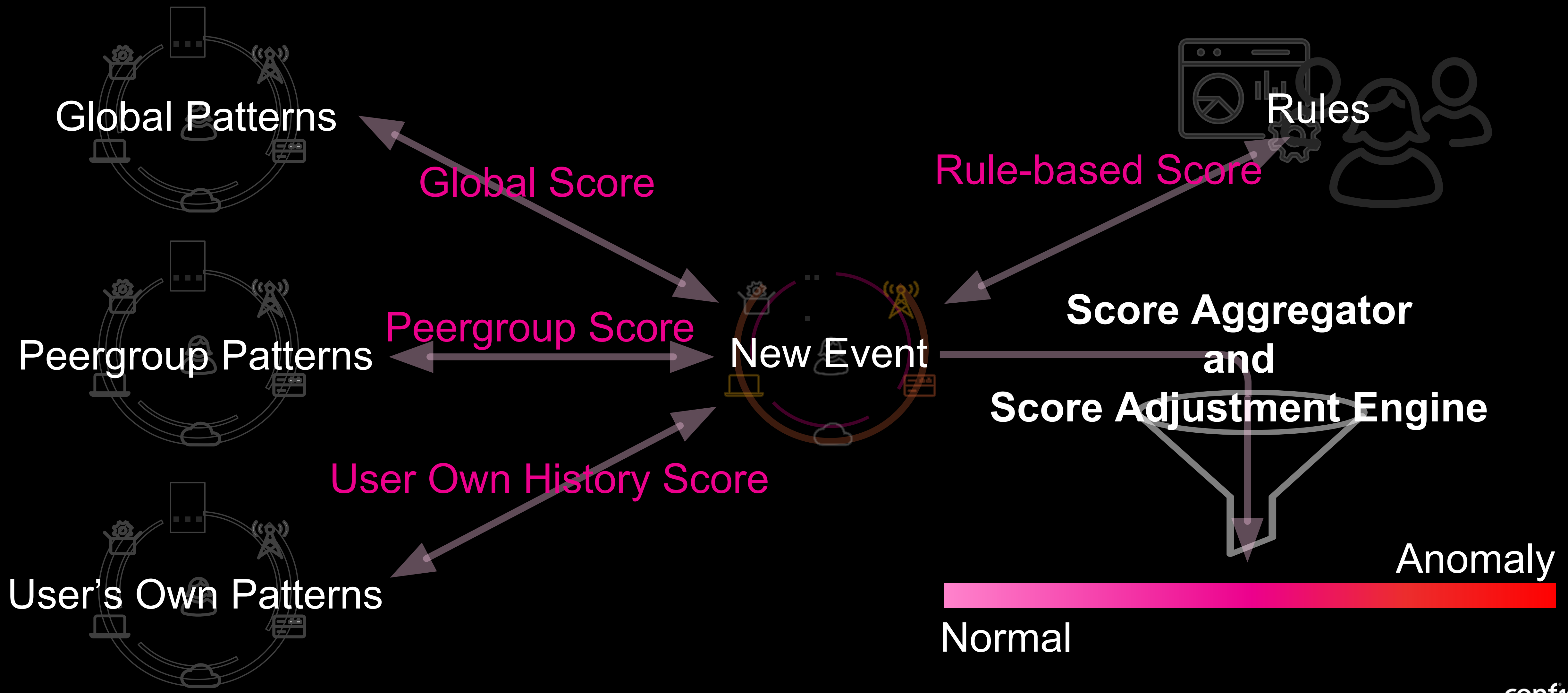
- Match itemset to the frequent patterns, score can be decomposed to

Fields share values with frequent patterns	+	Fields contain values rarely observed	+	Fields contain values completely new	+	Fields marked as anomalous by rules
Low score gain		Medium score gain		Relative High score gain		Extremely High score gain

ONE Combined Score for the New Event

Model Overview

Aggregate and adjust scores generated from various scopes



Case Study

How does the model perform on a real data example?

Real Data Experiment

Windows-based authentication events from Los Alamos National Laboratory [1] Dataset

- An approximate one month data related with user U66's authentication events (~1.5M)
- Labeled compromised events of this user from their red team as ground truth (118)
- First 17-days for training, following ~18k events for testing (normal + anomalies)

Objective: detect the compromise events

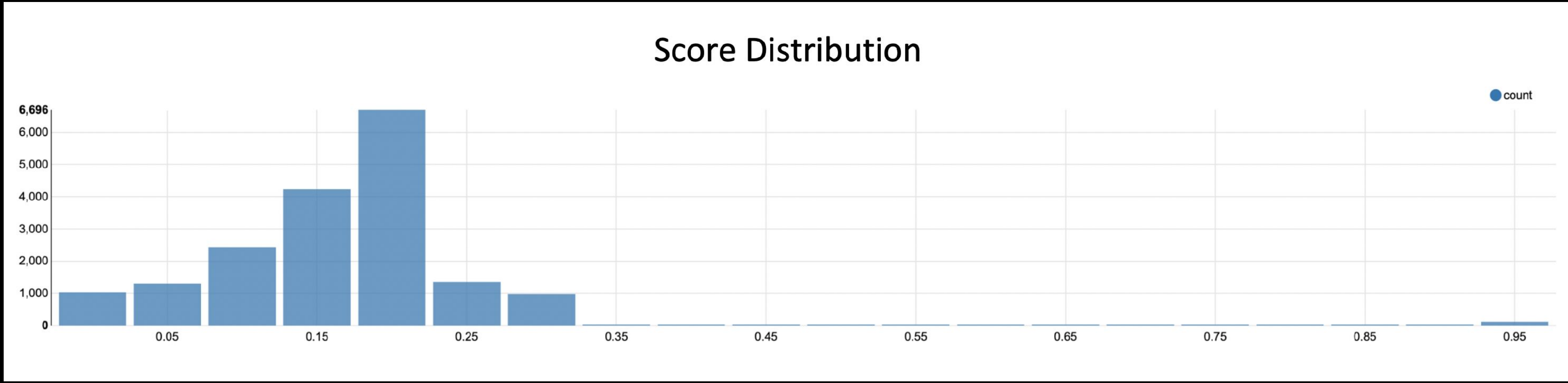
Label	Time	Source User	Source Domain	Source Computer	Destination User	Destination Domain	Destination Computer	Authentication Type	Logon Type	Authentication Orientation	Success OrFailure
Normal	766689	U66	DOM1	C1747	U66	DOM1	C1747	Kerberos	Network	LogOn	Success
Normal	769019	C3873$	DOM1	C3873	U66	DOM1	C3873	Kerberos	Network	LogOn	Success
Normal	774414	U53	DOM1	C1710	U66	DOM1	C1710	Negotiate	Interactive	LogOn	Fail
...											
Anomaly	2372551	U66	DOM1	C17693	U66	DOM1	C626	NTLM	Network	LogOn	Success
Anomaly	2370126	U66	DOM1	C17693	U66	DOM1	C5653	NTLM	Network	LogOn	Success
...											

[1] Kent, A. D. (2015). Comprehensive, multi-source cyber-security events data set (No. LA-UR-15-23810). Los Alamos National Lab.(LANL), Los Alamos, NM (United States).

Real Data Experiment

Windows-based authentication events from Los Alamos National Laboratory

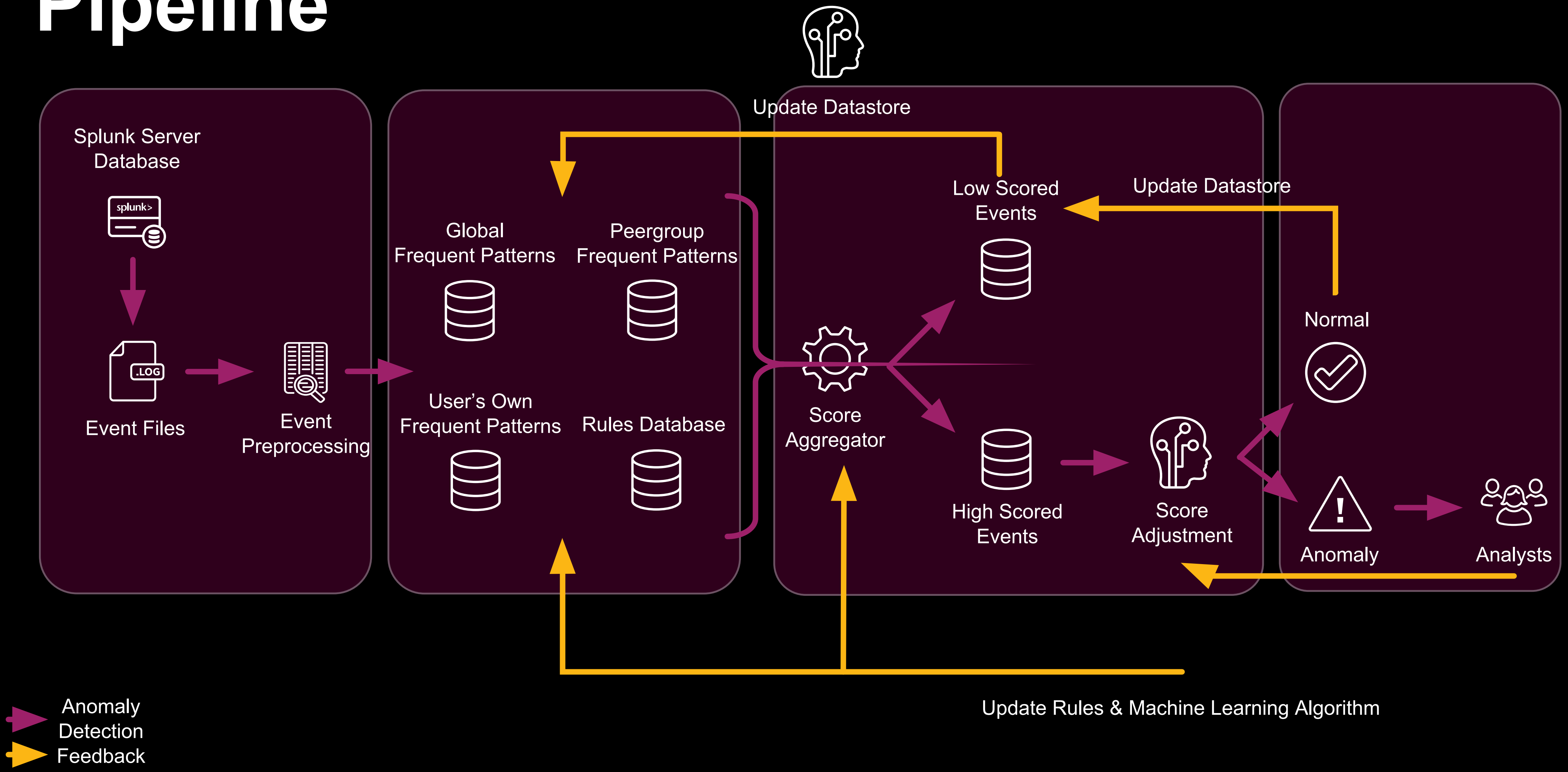
Results



- Precision = 1.00
- Recall = 0.91

Implementation

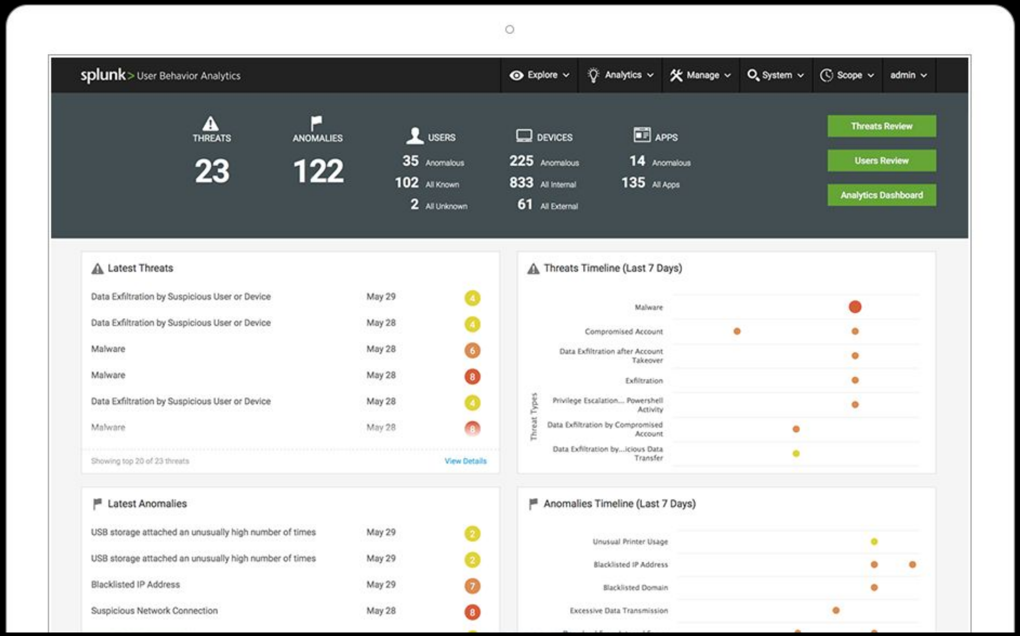
Pipeline


Update Datastore
Splunk Server
Database
splunk>
.LOG
Event Files
Event
Preprocessing
Global
Frequent Patterns
Peergroup
Frequent Patterns
User's Own
Frequent Patterns
Rules Database
Score
Aggregator
Low Scored
Events
Update Datastore
High Scored
Events
Score
Adjustment
Normal
Anomaly
Analysts
Update Rules & Machine Learning Algorithm
Anomaly
Detection
Feedback

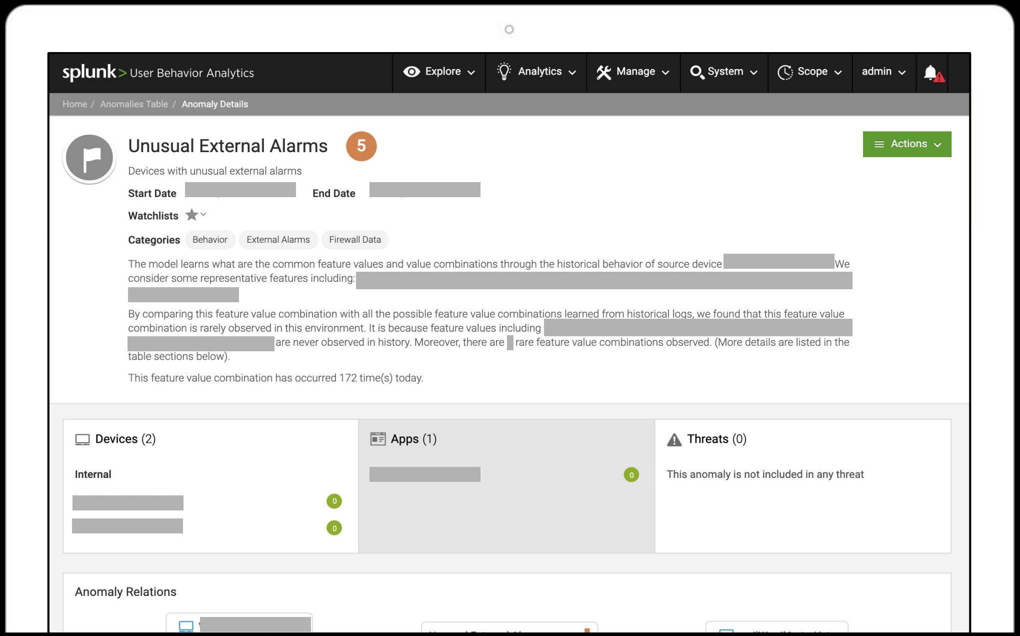
Splunk UBA

Available in Splunk UBA 4.3.2

Splunk® User Behavior Analytics



Deploy with Various Use Cases



Final Words

Last but not least, here’s some key takeaways!

Key Takeaways

Flexible

Apply to multiple use cases

Intelligent

Combine rules and domain knowledge

Interpretable

Easy to explain

Scalable

Implement at scale

Q & A

Contact Us

Product Manager: Koulick Ghosh (kghosh@splunk.com)

Data Scientist: Zhuxuan (Nancy) Jin (njin@splunk.com)

Thank You!

Go to the .conf19 mobile app to

RATE THIS SESSION