

# Правила сбора и анализа данных (Data Mining)

В отличие от потоковой корреляции, работающей в режиме реального времени, Data Mining правила позволяют с помощью языка SQL и функций ClickHouse ([примеры](#) запросов, почти все возможно использовать) распознавать и анализировать события, сохраненных в хранилище KUMA (можно указать и конкретный спейс хранилища).

Для работы необходимо указать, рассмотрим на примере:

- В **Ресурсах - Правила сбора** и анализа данных Создать правило

Ресурсы и сервисы / Правила сбора и анализа данных

## Правила сбора и анализа данных

Все Мои

☆ Избранное

🔍 Coverage Test

+ Добавить 📄 Дублировать 🗑 Удалить 🏷 Теги 📄 Показа

<input type="checkbox"/>	Название	Путь до ресурса	Последн...
<input type="checkbox"/>	Demo_Data_Mining	Shared	10.01.2025 11:2...

Всего 1 / Выбрано 0

- В правиле указать:
  - **Интервал (частота) выполнения SQL-запроса** можно указать в минутах, часах и днях (минимум 1 минута)
  - **SQL-запрос** должен содержать функцию агрегации ([примеры](#)) и/или группировку (GROUP BY) данных с обязательным указанием ограничения LIMIT (от 1 до 10 000)

Каждое выполнение такого правила происходит в виде запроса в Хранилище, а это значит неосторожным движением в виде частого или тяжелого правила можно нагрузить базу больше чем хотелось бы

В примере рассматривается запрос на основе событий Windows по пользователям (DestinationUserName) событиям входа (EventID 4624) и выхода (EventID 4634) с расчетом среднего времени сессии пользователя за последние 24 часа.

## Посмотреть SQL запрос (пример)

```
SELECT
    login_events.DestinationUserName AS destination_user_name,
    round(AVG(logout_events.logout_time - login_events.login_time)/1000) AS avg_time_diff_s,
    COUNT(DISTINCT login_events.login_time) AS total_logins,
    COUNT(DISTINCT logout_events.logout_time) AS total_logouts,
    concat(
        toString(floor(avg_time_diff_s / 86400)), ' days, ',
        toString(floor((avg_time_diff_s % 86400) / 3600)), ' hours, ',
        toString(floor((avg_time_diff_s % 3600) / 60)), ' minutes, ',
        toString(avg_time_diff_s % 60), ' seconds'
    ) AS human_readable_diff
FROM
    (SELECT
        DestinationUserName,
        toUnixTimestamp(EndTime) AS login_time,
        FlexString1 AS logon_id
    FROM `events`
    WHERE DeviceEventClassID = '4624'
    AND EndTime >= now() - INTERVAL 24 HOUR
    AND DestinationUserName NOT LIKE '%$%') AS login_events
INNER JOIN
    (SELECT
        DestinationUserName,
        toUnixTimestamp(EndTime) AS logout_time,
        FlexString1 AS logon_id
    FROM `events`
    WHERE DeviceEventClassID = '4634'
    AND EndTime >= now() - INTERVAL 24 HOUR
    AND DestinationUserName NOT LIKE '%$%') AS logout_events

    ON login_events.DestinationUserName = logout_events.DestinationUserName
    AND logout_events.logon_id = login_events.logon_id

WHERE logout_events.logout_time >= login_events.login_time
GROUP BY login_events.DestinationUserName
ORDER BY avg_time_diff_s DESC
LIMIT 100
```

- **Добавить маппинг** (сопоставление) по полям запроса и модели KUMA

+ Добавить сопоставление   Удалить		
Исходное поле	Поле события	Подпись
<input type="checkbox"/> destination_user_name	DestinationUserName	
<input type="checkbox"/> avg_time_diff_s	DeviceCustomNumber1	avg_time_diff_s
<input type="checkbox"/> total_logins	DeviceCustomString1	total_logins
<input type="checkbox"/> total_logouts	DeviceCustomString2	total_logouts
<input type="checkbox"/> human_readable_diff	DeviceCustomString3	human_readable_diff

- **Привязать хранилище** по которому будет осуществляться поиск
- **Привязать коррелятор** с соответствующим правилом корреляции для сработки
- В **Ресурсах - Сбор и анализ данных** добавить ранее созданное правило
- Для ручного запуска нажмите кнопку **Запустить**

Ресурсы и сервисы / Сбор и анализ данных

## Сбор и анализ данных

+ Добавить | Удалить | Запустить сейчас | Включить по расписанию | Отменить

Имя правила	Имя клиента
Demo_Data_Mining	Shared
lidzhiev-test	Main

### Редактирование правила сбора и анализа данных

Общие | Привязанные хранилища | Привязанные корреляторы

#### Планировщик

Состояние: Включено

Планировщик завершил работу в: 10.01.2025 12:30:41

Статус выполнения правила: OK

#### Настройки правила

Название\*: Demo\_Data\_Mining

Тенант\*: Shared

Теги:

Частота запуска запроса\*: 1 Час

Глубина: now-24h

Sql:

```

1 SELECT
2 ....login_events.DestinationUserName AS
3 destination_user_name,
4 ....round(AVG(logout_events.logout_time--
5 login_events.login_time)/1000) AS avg_time_diff_s,
6 ....COUNT(DISTINCT login_events.login_time) AS
7 total_logins,
8 ....COUNT(DISTINCT logout_events.logout_time) AS
9 total_logouts,
10 ....concat(

```

☐ Добавить сопоставление из SQL

Описание:

Сохранить | Сохранить с комментарием | **Запустить** | Отмена

- По результатам запроса на выходе будут какие-то данные, которые не будут нигде сохраняться, но на них можно настроить правило корреляции. В нашем случае правило ловит события, где время сессии меньше 5 секунд:

# Редактирование правила корреляции



Общие **Селекторы** Действия Корреляторы Исключения

Параметры

Локальные переменные

Параметры фильтра

Фильтр\*

Создать

☐ Сохранить фильтр

Конструктор

</> Код

И

+ Добавить условие

+ Добавить группу

Если

е: DeviceCustomNumber1Label

=

avg\_time\_diff\_s

Если не

е: Type

=

3

Если

е: DeviceCustomNumber1

<

5

Корреляционное событие выглядит следующим образом:

Алерты >

DataMiningBoris- Win AVG Session less 5 sec 

Новый

Уровень важности: 

Низкий

Назначить: 

Не назначено

Закрывать алерт

Создать инцидент

Привязать

Информация об алерте

Уровень важности правила корреляции

Первое появление

Тенант

Низкий

10.01.2025 11:30:43

Main

Наивысшая важность категории активов

Последнее появление

Правило корреляции

Нет значения

10.01.2025 11:30:43

[DataMiningBoris- Win AVG Session less 5 sec](#)

Идентификатор алерта

e7a32232-c6ae-4f6a-bf83-1d66e9636915

Связанные события

Время ↓

Информация о событии

ТЕНАНТ

10.01.2025 11:30:43

DestinationUserName: ksmg-idap, DeviceCustomString1: 193, DeviceCustomString2: 193, DeviceCustomString3: 0 days, 0 hours, 0 minutes, 0 seconds

Main

10.01.2025 11:30:43

DestinationUserName: myznikov, DeviceCustomString1: 1111, DeviceCustomString2: 1109, DeviceCustomString3: 0 days, 0 hours, 0 minutes, 3 seconds, DeviceCustomNumber1: 3

Main

10.01.2025 11:30:40

EndTime: 10.01.2025 11:30:40, DeviceTimeZone: +03:00, DestinationUserName: myznikov, DeviceCustomNumber1: 3, DeviceCustomNumber1Label: avg\_time\_diff\_s, DeviceCustomString1: 1111, DeviceCustomString1Label: total\_logins, DeviceCustomString2: 1109, DeviceCustomString2Label: total\_logouts, DeviceCustomString3: 0 days, 0 hours, 0 minutes, 3 seconds

Main

Найти в событиях 1

Main

Информация о корреляционном событии

Копировать

Подробные сведения

TenantName	Main
Timestamp	10.01.2025 11:30:43:274
Name	DataMiningBoris- Win- AVG- Session- less- 5- sec
StartTime	10.01.2025 11:30:40:180
EndTime	10.01.2025 11:30:40:180
DeviceProduct	KUMA
DeviceTimeZone	+03:00
DeviceVendor	Kaspersky
DestinationUserName	myznikov
DeviceCustomNumber1	3
DeviceCustomString1	1111
DeviceCustomString2	1109
DeviceCustomString3	0 days, 0 hours, 0 minutes, 3 seconds
CorrelationRule	<a href="#">DataMiningBoris- Win AVG Session less 5 sec</a>
Service	<a href="#">[QOTB] Correlator</a>
BaseEventCount	1
Priority	Низкий
Type	Correlated
Поля расширенной схемы событий	
N.KL_CorrelationRulePriority	1

А событие на основе которого произошла сработка:

Алерты >

DataMiningBoris- Win AVG Session less 5 sec Новый

Уровень важности: Низкий

Назначить: Не назначено

Заккрыть алерт

Создать инцидент

Привязать

Информация об алерте

Уровень важности правила корреляции  
Низкий

Первое появление  
10.01.2025 11:30:43

Тенант  
Main

Наивысшая важность категории активов  
Нет значения

Последнее появление  
10.01.2025 11:30:43

Правило корреляции  
DataMiningBoris- Win AVG Session less 5 sec

Идентификатор алерта  
e7a32232-c6ae-4f6a-bf83-1d66e9636915

Связанные события

Время ↓

Информация о событии

Тенант

10.01.2025 11:30:43

DestinationUserName: ksmg-ldap , DeviceCustomString1: 193 , DeviceCustomString2: 193 , DeviceCustomString3: 0 days, 0 hours, 0 minutes, 0 seconds

Main

10.01.2025 11:30:43

DestinationUserName: myznikov , DeviceCustomString1: 1111 , DeviceCustomString2: 1109 , DeviceCustomString3: 0 days, 0 hours, 0 minutes, 3 seconds , DeviceCustomNumber1: 3

Main

10.01.2025 11:30:40

EndTime: 10.01.2025 11:30:40 , DeviceTimeZone: +03:00 , DestinationUserName: myznikov , DeviceCustomNumber1: 3 , DeviceCustomNumber1Label: avg\_time\_diff\_s , DeviceCustomString1: 1111 , DeviceCustomString1Label: total\_logins , DeviceCustomString2: 1109 , DeviceCustomString2Label: total\_logouts , DeviceCustomString3: 0 days, 0 hours, 0 minutes, 3 seconds

Информация о событии

Копировать

Timestamp

10.01.2025 11:30:40:180

EndTime

10.01.2025 11:30:40:180

DeviceTimeZone

+03:00

DestinationUserName

myznikov

DeviceCustomNumber1

3

DeviceCustomNumber1Label

avg\_time\_diff\_s

DeviceCustomString1

1111

DeviceCustomString1Label

total\_logins

DeviceCustomString2

1109

DeviceCustomString2Label

total\_logouts

DeviceCustomString3

0 - days, - 0 - hours, - 0 - minutes, - 3 - seconds

DeviceCustomString3Label

human\_readable\_diff

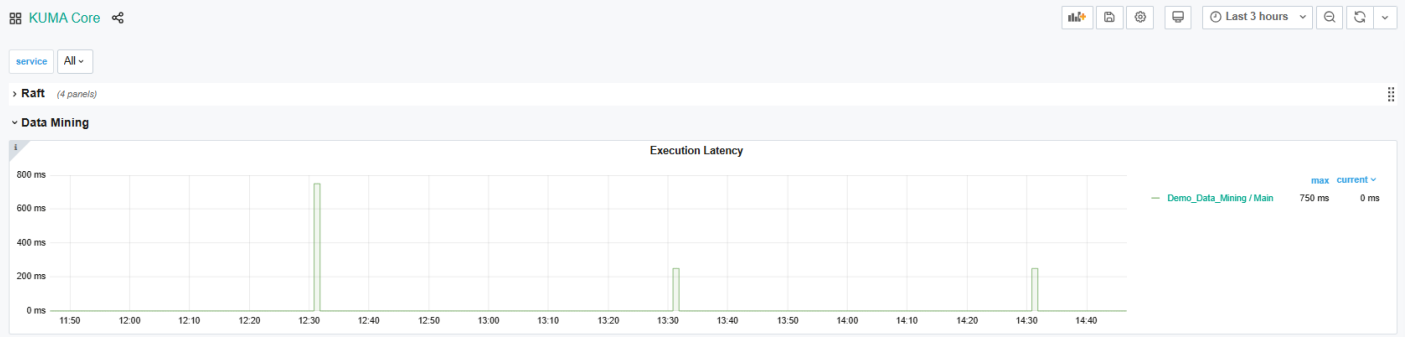
Service

core

Type

Base

Работу правил можно отслеживать с помощью метрик в разделе KUMA Core:



Revision #4  
Created 10 January 2025 08:44:23 by Boris RZR  
Updated 13 January 2025 08:14:42 by Boris RZR