Алгоритмы сбора и обработки данных в системах информационной безопасности

Д.К. Агапов

***Аннотация*—В данной статье представлено исследование системы сбора и корреляции событий (SIEM), сравнение разных методов сбора и обработки данных в системах информационной безопасности и методов сбора и обработки данных в системах информационной безопасности, а также определены основные проблемы современных SIEM. Статья состоит из четырёх разделов, первый из которых является введением, и заключения. Во введении раскрывается актуальность темы, обозначаются цель и задачи. Второй раздел посвящён анализу методов сбора и обработки данных в системах информационной безопасности. Даётся определение таким понятиям, как правила корреляции и модели в SIEM системах. В третьем разделе приведены результаты анализа современных решений в области сбора и обработки данных в системах информационной безопасности. В четвёртом разделе описаны проблемы и ограничения текущих SIEM, а также предложены варианты, как можно нейтрализовать их при создании будущих SIEM. Заключение содержит выводы по исследованию и предложения по использованию его результатов.**

***Ключевые слова*— Information security, SIEM, information security, event collection, event correlation.**

I. Введение

Риски информационной безопасности очень сильно выросли за последние годы. В основном это связано с колоссально и повсеместно возросшей активностью кибернетической преступности и национальных государств в данной сфере. Злоумышленники стали более квалифицированными, опасными и изощрёнными.

В качестве примеров инцидентов информационной безопасности можно привести [1]: инциденты компрометации корпоративной электронной почты (выдача злоумышленником себя за другое лицо или захват учётной записи), атаки программ-вымогателей, утечка и кража данных, социальная инженерия для сбора конфиденциальной информации от персонала предприятий, фишинговые атаки, направленные на руководителей, помощников руководителей, ИТ-администраторов или других пользователей с повышенными правами, вредоносное программные обеспечение, негативно влияющее на осуществление предприятием его основной деятельности, ведение бизнеса и операций.

Для эффективного реагирования на угрозы системам системы управления должны обеспечивать своевременное (в реальном времени) обнаружение различных аномалий, а также обеспечивать аналитическую визуализацию сетей и всех их взаимосвязей и узлов. Системы сбора и корреляции событий (SIEM) учитывают эти требования как встроенные функции.

Поэтому темой данного исследования была выбрана «Алгоритмы сбора и обработки данных в системах информационной безопасности».

Это исследование определяет основные направления развития современных систем сбора и обработки данных в системах информационной безопасности и содержит:

- Преимущества и недостатки обнаружения и реагирования на текущие сценарии атак;
- Описание инструментов (коммерческих и с открытым исходным кодом), направленных на определение основных характеристик системы сбора и корреляции событий;
- Анализ функций и возможностей современных SIEM;
- Анализ факторов, которые могут потенциально повлиять на будущие SIEM.

Также данное исследование позволяет определить возможность потенциального развития решений SIEM.

Цель исследования: исследовать системы сбора и корреляции событий (SIEM), сравнить разные методы сбора и обработки данных в системах информационной безопасности.

Задачами исследования в данной работе являются следующие:

Изучить: принципы построения систем сбора и корреляции событий (SIEM)

Ознакомиться: с современными решениями в области сбора и обработки данных в системах информационной безопасности.

II. Анализ методов сбора и обработки данных в системах информационной безопасности

Так как в современно виде SIEM системы представляют из себя готовые правила корреляции и сложные модели

Статья получена 9 апреля 2023.

Д. К. Агапов работает в Межрегиональном бухгалтерском управлении Федерального казначейства, Москва, Россия (e-mail: agapov.daniil@inbox.ru).

для выявления широкого спектра аномального поведения и событий, под методами сбора и обработки данных в системах информационной безопасности в данном исследовании понимаются эти правила и модели.

A. Правила корреляции в SIEM системах

В SIEM системах может быть очень большое количество различных правил корреляции, направленных на обнаружение всевозможных инцидентов информационной безопасности. Часть из них достаточно просты в реализации, часть – сложнее. Правила корреляции нужны для того, чтобы система информационной безопасности могла выполнить соответствующие действия по реакции на враждебные действия в её отношении. Часто при этом происходит приостановка её работы или даже полное её отключение, а также уведомление о этом событии пользователя.

Правило корреляции, также известное как правило фактов, — это логическое выражение, которое заставляет систему предпринимать определённые действия, если происходит определённое событие. Например, «Если на компьютере есть вирус, предупредите пользователя». Другими словами, правило корреляции — это условие (или набор условий), которое действует как триггер.

Правила корреляции неразумны — они не оценивают историю оцениваемых событий. Например, им все равно, был ли компьютер вчера заражён вирусом.

Правила корреляции могут быть простыми и работать сами по себе, или они могут быть составными правилами, обрабатывающими комбинации событий.

Простые правила SIEM определяют тип события и вызывают ответ. Например, если к электронному письму прикреплен ZIP-файл, они запускают предупреждение.

Составные правила объединяют два или более правил для достижения более сложного поведения. Например, если семь попыток аутентификации на одном компьютере завершились неудачно с одного и того же IP-адреса в течение десяти минут и используются разные имена пользователей, и если успешный вход в систему происходит на любом компьютере в сети и происходит с того же IP-адреса, они запускают предупреждение.

Правила корреляции зависят от знаний и опыта по обнаружению ошибочного поведения пользователя SIEM, а не от каких-либо атрибутов интеллектуальной системы. Чтобы написать правило корреляции, необходимо знать, какое нежелательное поведение нужно обнаружить.

Примеры правил корреляции:

- Если пользователю не удаётся более трёх попыток входа на один и тот же компьютер в течение часа, активируется предупреждение.
- Если после большого количества неудачных попыток входа в систему следует одна успешная, активируется предупреждение.

Бывают ситуации, когда правила корреляции — лучший и самый простой вариант. Вот несколько примеров, когда они наиболее эффективны:

- Мониторинг хорошо известных угроз — правила корреляции могут легко обнаруживать распространённые угрозы, которые хакеры неоднократно используют, чтобы попытаться получить доступ к вашим ресурсам. Многие решения SIEM заранее содержат правила для обработки этих типов угроз.
- Нарушение нормативных требований — организации в каждой отрасли должны продемонстрировать, что они соблюдают определённые законы, правила и нормы, например GDPR, HITECH и PCI DSS. У каждого есть требования, выполнение которых можно проверить с помощью правил корреляции. Например, «Оповещать, если антивирусное программное обеспечение отключено на любом компьютере, подключённом к сети».
- Обнаружение угроз на основе сигнатур. Системы обнаружения вредоносных программ имеют постоянно расширяющиеся репозитории, содержащие сотни миллионов известных сигнатур, идентифицирующих угрозы. Правила — лучший способ их обнаружить.

B. Модели в SIEM системах

Помимо правил корреляции, SIEM также может иметь модели. Модели несколько отличаются от правил корреляции, но при правильном применении могут быть столь же полезными. Вместо использования однозначной корреляции модель требует выполнения ряда шагов, чтобы вызвать предупреждение. Обычно это означает первое правило, за которым следует аномальное поведение. В качестве примера можно привести сценарий, при котором пользователь входит в систему из другого места, чем обычно, а затем выполняет передачу большого файла.

Это может быть чрезвычайно полезно, поскольку единичное событие не обязательно означает компрометацию серверов или сети организации, это может быть просто сотрудник, работающий из кафе для смены обстановки.

Модель профилирует поведение пользователя или актива, вызывая предупреждение, когда поведение отклоняется от нормального. После того, как модель выявляет ненормальное поведение, она использует правила для его оценки и предупреждения. Как правило, вы можете определять правила в моделях, которые классифицируют различные типы поведения, чтобы они могли создавать разные профили предупреждений:

- Правило первого раза — правило для нового события, выходящего за рамки того, что когда-либо происходило.
- Аномальное поведение — правило для события, которое фиксировалось, но не часто.

Особенность моделей, которая отличает их от правил корреляции, заключается в том, что они не всегда оцениваются или запускаются, даже если событие

происходит, тогда как правила корреляции всегда оцениваются. Выражение классификации в модели определяет, запускается ли это правило. Например, можно оценивать модель только в том случае, если пользователь удалён и пытается в третий раз создать новую системную учётную запись.

По сравнению с правилами SIEM, модели обычно имеют гораздо более простое выражение правила, которое запускает предупреждение — то есть, если поведение наблюдается более определённого числа раз, а коэффициент достоверности превышает заданное значение. Интеллект модели заключается в её выражении классификации и типах событий, которые она отслеживает.

Модели зависят от способности пользователя SIEM определять результаты необычного поведения, а также от способности системы отслеживать и выявлять такое поведение. Они не требуют от него глубокого понимания отдельных угроз информационной безопасности

Как и в случае с правилами корреляции, оценка отдельной модели обычно не вызывает предупреждения. Вместо этого каждая модель, которую применяет система, добавляет баллы к данному сеансу. Если точки сеанса превышают заданное значение, система выдает предупреждение.

Примеры моделей:

- Пользователь переключается со своей обычной учетной записи на привилегированную, а затем выполняет ненормальную передачу данных во внешнюю службу или из нее.
- Пользовательский VPN впервые подключается к сети из нового места, а затем обращается к общей файловой системе.
- Пользователь входит в систему удалённо в 3 часа ночи (обычно это делается локально в обычные рабочие часы), а затем предпринимает неоднократные попытки подключиться к производственной базе данных в качестве администратора.

Следует предпочесть модели правилам, когда:

- Нет возможности точно определить событие, которое указывает на нежелательное поведение.
- Динамические условия делают правила слишком сложными или заставляют их возвращать слишком много ложных срабатываний.

Примеры использования моделей:

- Угрозы на основе поведения. Динамические среды, имеющие несколько точек входа (например, организации, где сотрудники могут использовать свои собственные устройства или где корпоративные данные хранятся в облаке) предъявляют более высокие требования к вашим системам обнаружения угроз. Отслеживая как аномальное, так и нормальное поведение пользователя, модели могут выявлять угрозы, которые в противном случае могли бы остаться незамеченными.
- Обнаружение кражи данных. Ни одна организация не хочет, чтобы её ценные данные копировались в неизвестные внешние системы. Кража данных часто принимает форму хакеров, которые получают доступ, а затем перемещаются по сети в разные стороны, чтобы обнаружить активы, которые могут быть украдены. А иногда сами сотрудники компании или бывшие сотрудники будут пытаться забрать все, к чему у них есть доступ.
- Угрозы нулевого дня. По определению, угрозы нулевого дня ещё не встречались, и поэтому не могут быть точно идентифицированы правилом корреляции. Возможно использовать модели для обнаружения результатов этих угроз путём определения их нежелательного поведения, например, аномальных или удалённых входов в систему, доступа к файлам и аномальных загрузок данных.

III. Сравнение современных решений в области сбора и обработки данных в системах информационной безопасности

Системы управления информацией и событиями безопасности (SIEM) были разработаны, чтобы помочь администраторам разрабатывать политики безопасности и управлять событиями из различных источников. Как правило, простой SIEM состоит из отдельных блоков (например, исходное устройство, инструмент для сбора журналов, нормализация анализа, механизм правил, хранилище журналов, мониторинг событий), которые могут работать независимо друг от друга, но без их совместной работы SIEM не будет функционируют должным образом [2].

Платформы SIEM обеспечивают анализ в реальном времени событий безопасности, генерируемых сетевыми устройствами и приложениями. Кроме того, несмотря на то, что новое поколение SIEM обеспечивает возможности реагирования для автоматизации процесса выбора и развёртывания контрмер, текущие системы реагирования выбирают и развёртывают меры безопасности без выполнения всестороннего анализа воздействия атак и сценариев реагирования.

Помимо этих общих черт, текущие SIEM имеют различия, которые классифицируют их как лидеров, претендентов, нишевых игроков или визионров, согласно ежегодному отчету Gartner по SIEM Magic Quadrant. В этом исследовании представлены основные решения SIEM, доступные на рынке на сегодняшний день, и представлены основные преимущества и недостатки каждого из них на основе последнего отчета Gartner и исследований, связанных с технологиями SIEM [3,4,5,6,7,8,9,10,11].

A. Классификация SIEM

Анализ и оценка систем безопасности широко представлены в литературе. В то время как одни исследования сосредоточены на коммерческих аспектах, другие концентрируются на технических характеристиках, которые могут быть улучшены в текущих решениях SIEM. Хорошо известные

организации, такие как Gartner [12], например, предлагают коммерческий анализ систем SIEM на основе рынка и основных поставщиков, для чего ежегодно публикуется отчет, в котором поставщики SIEM позиционируются как лидеры рынка, претенденты, нишевые игроки, или дальновидныйы.

Другие институты безопасности (например, Techtarget и Info-Tech Research Group) широко сообщалось о возможностях решений SIEM и о том, как можно сравнивать и оценивать поставщиков SIEM. Techtarget, с одной стороны, выпускает периодические электронные руководства по обеспечению безопасности SIEM-систем и по определению SIEM-стратегии, управления и успеха на предприятии [13]. Info-Tech, с другой стороны, предоставляет технические отчеты о среде поставщиков SIEM [14], фокусируясь на преимуществах и недостатках основных коммерческих SIEM. Обе организации используют «Магический квадрант» Gartner в качестве основы для своего анализа.

В течение последнего десятилетия Gartner классифицировал SIEM-решения как лидеров (организации, которые хорошо справляются со своим текущим видением и имеют хорошие возможности для завтрашнего дня), дальновидных (организации, которые понимают, куда идет рынок или имеют видение изменения рыночных правил), нишевые игроки (организации, которые успешно фокусируются на небольшом сегменте или не сфокусированы и не внедряют инновации и не превосходят других) и претенденты (организации, которые сегодня хорошо работают или могут доминировать в большом сегменте, но не демонстрируют понимание направления рынка).

Недавнее исследование [14] рассматривает 22 игрока в карте поставщиков SIEM 2020 года на основе трех основных возможностей: обнаружение информации об угрозах, соответствие требованиям и управление журналами. Помимо анализа угроз, соответствия нормативным требованиям и управления журналами, разработчики SIEM рассматривают возможности UEBA и интеллектуальные информационные панели как инновации, которые следует добавить в свои решения. В результате новые системы SIEM помогут администраторам безопасности с помощью предварительно созданных информационных панелей, отчетов, рабочих процессов реагирования на инциденты, расширенной аналитики, корреляционного поиска и индикаторов безопасности. Кроме того, был проведен углубленный анализ расширяемости SIEM, который показал, что текущие решения SIEM нуждаются в улучшении таких функций, как поведенческий анализ, анализ и развертывание рисков, визуализация, хранение данных и возможности реагирования, чтобы идти в ногу с рынком.

B. Функции и возможности SIEM

По сути, все SIEM обладают способностью собирать, хранить и коррелировать события, генерируемые управляемой инфраструктурой. Помимо этих ключевых возможностей, существует множество различий между существующими системами, которые обычно отражают различные позиции SIEM на рынке. В этом разделе представлен список функций, которые следует учитывать при анализе решений SIEM.

Правила корреляции: Успех обнаружения события SIEM зависит от силы правил корреляции. В то время как большинство SIEM обладают базовыми правилами корреляции, некоторые из них обладают надежными возможностями поиска и поддерживают языки обработки поиска для написания сложных запросов, которые можно использовать для данных SIEM.

Источники данных: Одна из ключевых особенностей SIEM-системы — это возможность сбора событий из множества и разнообразных источников данных в управляемой инфраструктуре. Большинство SIEM изначально поддерживают несколько типов источников данных, включая как поддерживаемые датчики, так и поддерживаемые типы данных (например, анализ угроз). Для других решений (например, QRadar, USM) такая функция может поддерживаться дополнительными компонентами, интегрированными в SIEM. Эта функция оценивает изначально поддерживаемые источники данных и возможность SIEM автоматически настраивать их.

Обработка в реальном времени: эта функция учитывает способность SIEM обрабатывать данные в реальном времени при постоянном изменении. Он оценивает возможности управления, мониторинга и конвейерной обработки в реальном времени, развернутые инструментом для предотвращения или реагирования на инциденты кибербезопасности, а также возможности вычисления производительности, которые SIEM имеют для анализа миллионов событий в реальном времени. Все изученные SIEM обладают расширенными возможностями обработки в реальном времени.

Объем данных: анализ больших объемов данных, поступающих из разных источников, важен для получения более глубокого понимания собранных событий и для лучшего мониторинга. Однако хранение больших объемов собранных данных в действующей SIEM-системе часто бывает дорогостоящим и непрактичным. Эта функция оценивает возможность текущих систем поддерживать большие объемы данных для операций корреляции, индексации и хранения.

Визуализация: одним из ключевых факторов, препятствующих анализу событий безопасности, является отсутствие поддержки надлежащих методов визуализации данных и слабая поддержка интерактивного исследования собранных данных. Поэтому важно понимать возможности анализируемых систем с точки зрения создания новых методов визуализации данных и настраиваемых информационных панелей.

Аналитика данных: более свежие версии ведущих SIEM поддерживают обширную интеграцию с

детекторами аномалий на основе приложений и пользователей. Эти возможности включают анализ поведения сотрудников, сторонних подрядчиков и других сотрудников организации. Для этого SIEM должен включать в себя управление профилями пользователей/приложений и использование методов машинного обучения для обнаружения ненадлежащего поведения.

Производительность: эта функция оценивает производительность решения SIEM с точки зрения вычислительной мощности, возможностей работы с данными (например, чтение/запись), обработки корреляции правил (например, высокопроизводительный механизм корреляции), а также поиска, индексации и мониторинга данных.

Сложность: SIEM известны своей сложностью в развертывании и управлении. Однако важно понимать, можно ли установить анализируемую систему для тестирования с небольшими или умеренными усилиями. Из восьми изученных SIEM ArcSight представляет собой инструмент с высочайшей сложностью для развертывания и управления, тогда как LogRhythm и Splunk рассматриваются как простые и удобные инструменты для установки, развертывания и использования.

Масштабируемость: эта функция рассматривает возможность роста развертывания SIEM не только с точки зрения оборудования, но и с точки зрения количества событий безопасности, собранных на границе инфраструктуры SIEM. Новое цифровое преобразование приводит к тому, что больше датчиков и устройств (например, серверов, агентов, узлов) подключаются к одной и той же сети.

Анализ рисков. Последние версии ведущих систем SIEM (например, QRadar, LogRhythm, Splunk) включают функции для анализа рисков для активов управляемой инфраструктуры. Эта функция оценивает, поддерживает ли SIEM изначально анализ рисков или его можно интегрировать с внешними устройствами для этой цели.

Хранение: учитывая, что SIEM обычно хранят информацию не более 90 дней, эта функция оценивает продолжительность, с которой текущие технологии SIEM хранят данные, хранящиеся в их системах, для дальнейшей обработки и криминалистических операций.

Цена: эта функция оценивает метод лицензирования, связанный с решением SIEM (например, корпоративное, бесплатное, бета-версию, премиум), и ограничения на количество пользователей, запросов, томов индекса, предупреждений, корреляций, отчетов, панелей мониторинга и автоматических корректирующих действий. Большинство изученных решений очень дороги, за исключением LogRhythm, USM и SolarWinds, с более разумной стоимостью и возможностью использовать решения с открытым кодом с более ограниченными возможностями.

Устойчивость: устойчивость или отказоустойчивость — важная характеристика любой критически важной системы мониторинга. Важно понимать, каковы возможности отказоустойчивости существующих SIEM, например, поддерживает ли механизм корреляции отказоустойчивость. Также важно понимать, присутствует ли в SIEM способ поддержки аварийного восстановления и репликации в хранилище событий; если соединители поддерживают функции высокой доступности.

Возможности реагирования и отчетности: эта функция изучает действия, которые изначально поддерживаются SIEM для реагирования на инциденты безопасности (включая возможности совместного использования и отчетности), а также то, как такие действия выражаются механизму корреляции.

UEBA: эта функция оценивает, предоставляет ли решение SIEM встроенную функцию анализа поведения пользователей и объектов (UEBA) или обеспечивает интеграцию со сторонними решениями UEBA.

Безопасность: эта функция оценивает возможность реализации автоматизации безопасности, а также встроенные возможности шифрования, присутствующие в SIEM, во время мониторинга, обнаружения, корреляции, анализа и представления результатов.

IV. Ограничения текущих SIEM

Несмотря на то, что новое поколение SIEM обеспечивает мощные функции с точки зрения корреляции, хранения, визуализации и производительности, а также возможность автоматизировать процесс реакции путем выбора и развертывания контрмер, существующие системы реагирования очень ограничены и меры противодействия выбираются и развертываются без проведения всестороннего анализа воздействия атак и сценариев реагирования.

Кроме того, большинство SIEM поддерживают интеграцию новых коннекторов или парсеров для сбора событий или данных и предоставляют API или интерфейсы RESTful для сбора событий позже. Эти механизмы позволяют создавать дополнения и расширения к существующим системам. Будущие SIEM должны использовать эту функцию, чтобы повысить качество событий, подаваемых в систему (например, используя новые системы мониторинга или сбор внешних данных из разведки с открытым исходным кодом) через настраиваемые соединители, а также предоставлять новые инструменты визуализации путем сбора данных из SIEM. хранилище данных.

В этом разделе подробно описаны основные ограничения, имеющиеся в текущих решениях SIEM, и даны некоторые перспективы возможных улучшений.

A. Неполные данные

Хотя современные SIEM имеют дело с тоннами данных, ни у одной из них нет всех данных, необходимых для обработки и обнаружения всех инцидентов безопасности. Причина в том, что сбор и обработка всех

необходимых данных нерентабельны. Как правило, все SIEM сопоставляют журналы из сетей VPN, брандмауэров, элементов управления доменом, неудачных подключений и т. д. Большинство SIEM могут сопоставлять журналы входа в систему, вредоносное ПО и веб-журналы, но лишь немногие из текущих SIEM коррелируют трафик DNS, журналы данных конечных точек и т. д., а также журналы электронной почты. В результате невозможно узнать, кто все в системе.

Идентификационные данные фрагментированы, у людей есть общие учетные записи, и разные роли связаны с одним и тем же пользователем, но по закону мы не можем раскрыть личность данного человека, поскольку это создает проблемы с конфиденциальностью, как это представлено в Общем регламенте защиты данных (GDPR). Если SIEM не может собрать все данные о пользователях и ценных активах, корреляция никогда не будет работать должным образом, что приведет к большому количеству ложных срабатываний и отрицательных результатов. Таким образом, следующее поколение SIEM должно отвечать требованиям конфиденциальности GDPR, в то же время предоставляя аналитикам достаточно информации для выявления инцидентов безопасности.

Обзор существующих SIEM позволил подтвердить, что эти системы не предоставляют высокоуровневые метрики рисков безопасности. Важным достижением по сравнению с текущими SIEM будет разработка полезных операционных показателей, которые позволят SOC принимать решения, подкрепленные количественными доказательствами, в которых явно указывается неопределенность в показателях, а также с улучшенной поддержкой визуализации, позволяющей лучше информировать об этих решениях соответствующие заинтересованные стороны в организации. Такие измерения должны поддерживаться на нескольких уровнях защиты (например, межсетевые экраны, IDS, антивирусные продукты, операционные системы, приложения) и различных продуктах каждого типа.

Хотя чувствительные к стоимости метрики трудно вычислить из-за сложности оценки затрат на безопасность организаций, новые SIEM-системы должны приближаться к этой категории метрик с использованием высокодетализированной оценки затрат.

Будущие SIEM должны исследовать и внедрять новые неконтролируемые методы, которые сочетают статистический и многокритериальный анализ решений для автоматического моделирования приложений и поведения пользователей, а затем выявляют аномалии и отклонения от известных хороших моделей поведения, которые являются статистически значимыми. Это приведет к развертыванию усовершенствованных датчиков мониторинга приложений, которые будут снабжать SIEM-системы различными типами событий, которые могут быть коррелированы с более традиционными событиями безопасности, собираемыми с хоста и сетевых устройств.

Комбинируя события, связанные с аномалиями, с событиями, предоставляемыми более традиционными эвристическими и сигнатурными инструментами, SIEM уменьшит количество ложных срабатываний этих компонентов, которые традиционно были основным камнем преткновения при их широком применении в реальных операциях.

B. Основные правила корреляции

Платформы SIEM обеспечивают анализ в реальном времени событий безопасности, генерируемых сетевыми устройствами и приложениями. Эти системы получают большие объемы информации из разнородных источников и обрабатывают их на лету. Таким образом, их развертывание сосредоточено, во-первых, на написании специальных сборщиков и переводчиков для сбора информации и ее нормализации, а во-вторых, на написании правил корреляции для агрегирования информации и уменьшения объема данных. Эта операционная направленность заставляет разработчиков SIEM отдавать предпочтение синтаксису, а не семантике, и использовать языки корреляции, которые бедны с точки зрения функций. Однако по мере увеличения количества атак и, следовательно, разнообразия предупреждений, получаемых SIEM, потребность в надлежащей обработке этих предупреждений становится важной.

Текущие правила корреляции SIEM слабые. Большинство из них используют базовую логическую цепочку событий, которые проверяют конкретный путь атаки (один из многих тысяч возможных). Очень немногие решения SIEM имеют встроенный усовершенствованный механизм корреляции, способный выполнять корреляцию отклонений и историческую корреляцию, полезную, например, для проверки после обнаружения атак нулевого дня.

C. Основные возможности хранения

Для большинства существующих решений SIEM данные, которые заархивированы и удалены из действующей системы, больше не используются. Более того, то, как обрабатываются архивные данные, где они хранятся или передаются, зависит от пользователя и обычно выполняется вручную. Поскольку существует множество вариантов для хранения архивных данных, некоторые пользователи SIEM выбирают присоединённое хранилище, другие будут использовать внутреннюю распределенную файловую систему, например, Hadoop Distributed File System (HDFS) или коммерческое облачное хранилище, такое как Amazon S3, Amazon Glacier.

Независимо от используемого решения для архивирования, фактический процесс архивирования состоит из выполнения сценариев, которые часто создаются специально для конкретной ИТ-среды.

Следовательно, сценарий, используемый одним клиентом, может оказаться бесполезным для нужд другого клиента, а изменение параметра архивации требует переписывания сценария архивирования.

Кроме того, архивирование устаревших данных из SIEM может быть дорогостоящим и может создать проблемы с безопасностью и надежностью, если с архивными данными не обращаться правильно.

Текущие инфраструктуры обычно хранят необработанные события в течение ограниченного периода времени, чтобы ограничить пространство для хранения, используемое для такого архивирования. Учитывая, что некоторые сложные постоянные угрозы обнаруживаются через много месяцев после их появления в системе, таких возможностей хранения может быть недостаточно, чтобы помочь в определенных инцидентах.

Хотя это многообещающе, большинство компаний избегают использования облака из-за опасений, связанных с конфиденциальностью событий (которые содержат конфиденциальную информацию), и опасений, связанных с доверием таких важных данных третьим сторонам.

Цель будущих SIEM должна быть сосредоточена на предоставлении безопасного и гибкого решения для архивирования данных независимо от требований к хранению данных с возможностью настройки политик в соответствии с требованиями к хранению.

V. Заключение

На сегодняшний день внедрение SIEM систем является неотъемлемой частью процесса обеспечения информационной безопасности практически любого предприятия, так как для реакции на инциденты в этой сфере нужно вовремя собирать информацию о них и обрабатывать её. Основой данных систем являются правила корреляции и модели. В ходе реализации и использования SIEM систем необходимо чётко понимать, какие именно методы сбора и обработки информации нужно применять в конкретной ситуации, а также как минимизировать количество ложных срабатываний.

С точки зрения поведенческого анализа, анализа и развертывания рисков, методы и инструменты для анализа, оценки и руководства оптимальным развертыванием различных механизмов безопасности в управляемой инфраструктуре (включая многоуровневые метрики на основе рисков) должны быть разработаны вместе со структурой. для развертывания разнообразных и дублирующих датчиков.

Хотя большинство проанализированных решений предоставляют удобный графический интерфейс, возможности визуализации и реагирования ограничены для работы с огромным количеством собранных событий. Поэтому важно разработать расширения для визуализации и анализа, которые помогут дать пользователям высокий уровень понимания ситуации и более эффективные возможности принятия решений и реагирования.

Данная статья может помочь:

Специалистам в области информационной безопасности сориентироваться среди методов, используемых сегодня в SIEM, принять решение, какие из них лучше использовать в данной области, что поможет более верно настроить SIEM и тем самым увеличить его эффективность, минимизировав количество ложных срабатываний, проанализировать доступные сейчас решения SIEM, основываясь на уровне реализации в них тех или иных функций, что позволит выбрать максимально подходящее решение. Всё это поможет своевременно реагировать на угрозы информационной безопасности и максимально оптимизировать использование ресурсов в этой области;

Разработчикам SIEM сделать выводы о том, в каком направлении целесообразно развивать решения в этой области, представляемые ими, и на какие недостатки в них стоит обратить внимание. Благодаря этому они могут повысить рентабельность этих решений.

Библиография

[1] WaterISAC. 15 Основы кибербезопасности для предприятий водоснабжения и канализации. Лучшие практики для уменьшения уязвимых мест и атак. Доступно в Интернете: https://www.waterisac.org/system/files/articles/15%20Cybersecurity%20Fundamentals%20%28WaterISAC%29.pdf (доступ 10 октября 2022 г.).

[2] Miller, D.; Harris, S.; Harper, A.; Van Dyke, S.; Blask, C. Внедрение управления информацией и событиями безопасности (SIEM) ; Мак Гроу Хилл: Нью-Йорк, Нью-Йорк, США, 2010. [Google Scholar]

[3] Nicolett, M.; Kavanagh, K.M. Magic Quadrant для информации о безопасности и управления событиями, технический отчет Gartner. Доступно в Интернете: http://docplayer.net/2407833-Magic-quadrant-for-security-information-and-event-management.html (доступ 10 ноября 2022 г.).

[4] Nicolett, M.; Kavanagh, K.M. Magic Quadrant для информации о безопасности и управления событиями, технический отчет Gartner. Доступно в Интернете: https://www.bankinfosecurity.com/whitepapers/2012-gartner-magic-quadrant-for-siem-w-602 (доступ 12 ноября 2022 г.).

[5] Nicolett, M.; Kavanagh, K.M.; Rochford, O. Магический квадрант для информации о безопасности и управления событиями, Технический отчет Gartner. Доступно в Интернете: https://www.bwdigitronik.ch/application/files/5814/5450/7565/www.gartner.com.com.pdf (доступ 12 ноября 2022 г.).

[6] Kavanagh, K.M.; Rochford, O.; Бусса, Т. Магический квадрант для информации о безопасности и управления событиями, Технический отчет Gartner. Доступно в Интернете: https://securelink.net/wp-

content/uploads/sites/7/2016-Magic-Quadrant-for-SIEM.pdf (доступ 10 ноября 2022 г.).
[7] Kavanagh, K.M.; Sadowski, T.B.G. Magic Quadrant for Security Information and Event Management, Gartner Technical Report. Доступно в Интернете: https://virtualizationandstorage.files.wordpress.com/2018/03/magic-quadrant-for-security-information-and-event-3-dec-2018.pdf (доступ 10 ноября 2022 г.).
[8] Kavanagh, K.M.; Sadowski, T.B.G. Magic Quadrant for Security Information and Event Management, Gartner Technical Report. Доступно в Интернете: https://www.gartner.com/en/documents/3981040/magic-quadrant-for-security-information-and-event-manage (доступ 25 ноября 2022 г.).
[9] Скарфоне, К. Сравнение лучших SIEM-систем на рынке. Интернет-исследования. Доступно в Интернете: http://searchsecurity.techtarget.com/feature/Comparing-the-best-SIEM-systems-on-the-market (доступ 27 ноября 2022 г.).
[10] Nirvana, I. Сравнение продуктов SIEM-2016. Доступно в Интернете: http://infosecnirvana.com/siem-product-comparison-201/ (доступ 21 ноября 2022 г.).
[11] Rochford, O .; Kavanagh, K.M.; Бусса, Т. Критические возможности для управления информацией и событиями безопасности; Технический отчет; Gartner: Стэмфорд, Коннектикут, США, 2016 г. [Google Scholar]
[12] McAfee. Информация о безопасности и управление событиями (SIEM). Официальный сайт. Доступно в Интернете: https://www.mcafee.com/enterprise/en-us/products/siem-products.html (доступ 29 ноября 2022 г.).
[13] Trustwave. SIEM Enterprise. Краткое описание продукта. Доступно в Интернете: https://trustwave.azureedge.net/media/13581/tw-siem-enterprise.pdf?rnd=131659475410000000 (доступ 29 ноября 2022 г.).
[14] LogRhythm. Информация о безопасности и управление событиями (SIEM). Доступно в Интернете: https://logrhythm.com/solutions/security/siem/ (доступ 31 мая 2022 г.).

Д.К. Агапов родился в городе Москва 29 декабря 1999 года. Окончил обычную общеобразовательную школу, а затем РТУ МИРЭА по направлению Информационная безопасность телекоммуникационных систем.

Algorithms for data collection and processing in information security systems

D. K. Agapov

***Abstract*—This article presents a study of the event collection and correlation system (SIEM), a comparison of different methods of collecting and processing data in information security systems and methods of collecting and processing data in information security systems, and also identifies the main problems of modern SIEM. The article consists of four sections, the first of which is an introduction, and a conclusion. The introduction reveals the relevance of the topic, identifies the goal and objectives. The second section is devoted to the analysis of data collection and processing methods in information security systems. The definition of such concepts as correlation rules and models in SIEM systems is given. The third section presents the results of the analysis of modern solutions in the field of data collection and processing in information security systems. The fourth section describes the problems and limitations of current SIEMs, and also suggests options for how to neutralize them when creating future SIEMs. The conclusion contains conclusions on the study and suggestions for using its results.**

***Keywords*— Information security, SIEM, information security, event collection, event correlation.**

REFERENCES

[1] WaterISAC. 15 Fundamentals of cybersecurity for water supply and sewerage enterprises. Best practices for reducing vulnerabilities and attacks. Available online: https://www.waterisac.org/system/files/articles/15%20Cybersecurity%20Fundamentals%20%28WaterISAC%29.pdf (accessed October 10, 2022).

[2] Miller, D.; Harris, S.; Harper, A.; Van Dyke, S.; Blask, C. Implementing Security Information and Event Management (SIEM); McGraw Hill: New York, NY, USA, 2010. [Google Scholar]

[3] Nicolett, M.; Kavanagh, K.M. Magic Quadrant for Security Information and Event Management, Gartner Technical Report. Available online: http://docplayer.net/2407833-Magic-quadrant-for-security-information-and-event-management.html (accessed November 10, 2022).

[4] Nicolett, M.; Kavanagh, K.M. Magic Quadrant for Security Information and Event Management, Gartner Technical Report. Available online: https://www.bankinfosecurity.com/whitepapers/2012-gartner-magic-quadrant-for-siem-w-602 (accessed November 12, 2022).

[5] Nicolett, M.; Kavanagh, K.M.; Rochford, O. Magic Quadrant for Security Information and Event Management, Gartner Technical Report. Available online: https://www.bwdigitronik.ch/application/files/5814/5450/7565/www.gartner.com.com.pdf (accessed November 12, 2022).

[6] Kavanagh, K.M.; Rochford, O.; Bussa, T. The Magic Quadrant for Security Information and Event Management, Gartner Technical Report. Available online: https://securelink.net/wp-content/uploads/sites/7/2016-Magic-Quadrant-for-SIEM.pdf (accessed November 10, 2022).

[7] Kavanagh, K.M.; Sadowski, T.B.G. Magic Quadrant for Security Information and Event Management, Gartner Technical Report. Available online: https://virtualizationandstorage .files.wordpress.com/2018/03/magic-quadrant-for-security-information-and-event-3-dec-2018.pdf (accessed November 10, 2022).

[8] Kavanagh, K.M.; Sadowski, T.B.G. Magic Quadrant for Security Information and Event Management, Gartner Technical Report. Available online: https://www.gartner.com/en/documents/3981040/magic-quadrant-for-security-information-and-event-manage (accessed November 25, 2022).

[9] Scarfone, K. Comparison of the best SIEM systems on the market. Internet research. Available online: http://searchsecurity.techtarget.com/feature/Comparing-the-best-SIEM-systems-on-the-market (accessed November 27, 2022).

[10] Nirvana, I. Comparison of SIEM-2016 products. Available online: http://infosecnirvana.com/siem-product-comparison-201 / (accessed November 21, 2022).

[11] Rochford, O.; Kavanagh, K.M.; Bussa, T. Critical Capabilities for Security Information and Event Management; Technical Report; Gartner: Stamford, Connecticut, USA, 2016 [Google Scholar]

[12] McAfee. Security Information and Event Management (SIEM). Official website. Available online: https://www.mcafee.com/enterprise/en-us/products/siem-products.html (accessed November 29, 2022).

[13] Trustwave. SIEM Enterprise. A brief description of the product. Available online: https://trustwave.azureedge.net/media/13581/tw-siem-enterprise.pdf?rnd=131659475410000000 (accessed November 29, 2022).

[14] LogRhythm. Security Information and Event Management (SIEM). Available online: https://logrhythm.com/solutions/security/siem / (accessed May 31, 2022).