Недекларированные возможности SIEM

Иван Лопатин, *технический эксперт АО "ДиалогНаука"*

Очень многие заказчики используют SIEM-системы для выявления инцидентов информационной безопасности разной степени сложности и зачастую даже не представляют себе все возможности собственной системы мониторинга. Это связано с тем, что из-за довольно большой загрузки отдела информационной безопасности его сотрудникам не приходится задумываться о дополнительных функциях и возможностях, хотя эти знания могли бы существенно помочь в автоматизации целого ряда процессов в отделе ИБ или всей компании в целом.

В данной статье мы постараемся подтолкнуть специалистов к исследованию дополнительных свойств SIEM-систем, которые могут быть использованы без необходимости закупки и расширения существующих лицензий. К основным таким недекларированным возможностям относятся:

- запуск внешних скриптов и программ;
- использование скоринговой модели;
- применение нейросетей для выявления инцидентов ИБ.

Одним из кейсов, которые можно реализовать на базе недекларированной возможности, является запуск сторонних или внешних скриптов.

Одним из инструментов, которые можно использовать для централизованного запуска скриптов, является Ansible. Можно заранее настроить сценарии выполнения команд и действий для их более качественного и бесперебойного исполнения.

Запуск внешних скриптов и программ

Одним из кейсов, которые можно реализовать на базе недекларированной возможности, является запуск сторонних или внешних скриптов, например для:

- сканирования хоста, на котором произошел инцидент информационной безопасности;
- записи в сторонние средства и системы информации об инциденте;
- исполнения команд на удаленной системе для произведения действий над пользователем или узлом, фигурирующим в инциденте ИБ.

Остановимся несколько подробнее на первом пункте (сканирование хоста) и рассмотрим ситуацию, при которой происходит инцидент, связанный с попытками нелегитимных действий в части очистки журнала аудита и/или создание административной группы на Linux-системе. При выявлении правилом данного события запускается скрипт сбора с узла дополнительных сведений, связанных с данным пользователем или группой. Результаты работы скрипта направляются обратно в SIEM-систему для фиксации всей полученной информации о пользователе или группе.

При такой реализации работа правила становится максимально продуктивной. Но стоит учитывать тот факт, что все правила, которые производят запуски внешних команд, должны пройти этапы тестирования и отладки для минимизации ложных срабатываний и оптимизации загрузки системы. Для этого в SIEM-системе создаются механизмы автоматической проверки корректности работы правил и частоты их срабатываний.

Вернемся к кейсу сканирования хоста. Если данный инцидент приводит к его заведению в системе управления инцидентами (Case Manager), то оператору и аналитику будет гораздо проще принять решение о критичности данного инцидента/кейса и необходимости его эскалации, если в кейсе будут дополнительные признаки инцидента.

Кейс: сканирование хоста

Для принятия решения о критичности данного инцидента оператору или аналитику значительно поможет следующая информация в кейсе:

1. ID пользователя в Linux-системе, который выполняет злонамеренные действия;

2. Список неудачных входов (команда #lastb), IP-адрес, время события. Данная информация полезна в случае недостижения порога "неуспешных входов" для срабатывания инцидента Brute Force (попытка несанкционированного доступа методом перебора паролей):

support ssh:notty 103.138.109.76 Wed Dec 4 12:20 – 12:20 (00:00)
tom ssh:notty 140.207.46.136 Wed Dec 4 12:51 – 12:51 (00:00)
ubnt ssh:notty 185.43.209.8 Wed Dec 4 11:17 – 11:17 (00:00)
user ssh:notty 185.43.209.8 Wed Dec 4 11:17 – 11:17 (00:00)
admin ssh:notty 185.43.209.8 Wed Dec 4 11:17 – 11:17 (00:00)
ftpuser ssh:notty 5.238.245.53 Thu Dec 5 06:00 – 06:00 (00:00)

3. Список запущенных процессов (команда: #ps -eo pid, lstart, cmd)

4547 15:48:23 /bin/bash /tmp/<name>.py

4. Другие команды, необходимые для более эффективного расследования инцидентов ИБ в Linux-системах, такие как запуск сканеров, антивирусов, блокировка пользователей и процессов (в случае необходимости).

Использование скоринговой модели

В обиходе словосочетание "скоринговая модель" ассоциируется с антифрод-системами и борьбой с финансовыми мошенниками. Но реализация скоринговой модели в информационной безопасности и в SIEM-системе уже является той необходимостью, без которой полнота картины выявления инцидентов ИБ не будет достаточно исчерпывающей.

Для использования более сложных корреляционных логик и выявления более сложных цепочек инцидентов необходимо реализовывать скоринговую модель для подсчета баллов по

каждому пользователю и узлу, участвующему в инцидентах ИБ.

Так, например, при срабатывании правила Brute Force производится занесение в список подсчета скоринга как пользователя, так и хоста, на котором происходит попытка подбора пароля.

При фиксации ряда инцидентов с пользователем или узлом, находящимся в листе скоринга, суммарный балл повышается в зависимости от критичности инцидента. Таким образом, суммируя баллы за каждый инцидент, мы наглядно видим картину происходящего, а также в случае реализации дополнительных механизмов можно контролировать происходящие инциденты и накладывать на жизненный цикл кибератак Kill Chain с последующей классификацией по MITRE.

Приведем еще один абстрактный пример. Создаем два списка (листа) с названиями User List и Host List:

- поля в User List – пользователь, скоринг, дата;
- поля в Host List – узел, скоринг, дата.

Добавляем в правила, выявляющие инциденты ИБ, действия по занесению пользователя или хоста в соответствующий лист, а также скоринговое значение.

На выходе получаем, что при выявлении инцидентов в листы добавляется информация о пользователях и хостах, фигурирующих в инцидентах, а также подсчитывается скоринг по каждому пользователю и хосту.

Такими действиями мы создаем дополнительный механизм анализа и возможности для изменения критичности выявления инцидентов, связанных с тем или иным хостом или пользователем. В случае грамотно построенной модели можно и активно реагировать на инциденты (например, осуществлять действия, описанные в начале статьи).

Применение нейросетей

Сейчас все больший оборот набирает использование математических моделей в области информационной безопасности, одним из примеров которых являются нейросети. Внедрение нейросетей в ИБ пока носит скорее желательный характер, чем обязательный, и они дополняют функционал уже существующих СЗИ. Перед крупными игроками, использующими нейросети для предсказания тех или иных параметров и ситуаций, встает проблема создания качественных наборов данных (dataset), по которым могла бы обучаться и переобучаться модель.

Подготовка набора данных для модели – это кропотливый труд специалистов в области сбора и анализа больших массивов данных (Data Science и Data Mining). Обычно в небольших организациях со своим штатом сотрудников на сбор, очистку и подготовку данных для обучения математических моделей уходит порядка 70% времени.

Использование заранее размеченного набора данных с принудительным обучением (значение – реакция) является эффективным способом обучения нейросети по предсказанию аномалий, но в то же время трудозатратным.

Предлагается использовать мощности SIEM-систем для генерации части наборов данных для обучения нейросети и минимизации трудозатрат специалистов по сбору и анализу данных.

SIEM-система собирает огромное количество событий ИБ и производит их нормализацию, агрегацию и корреляцию – данная информация и будет использоваться в дальнейшем. Архитектура построения интеграции заключается в настройке SIEM-системы в части правил корреляции, минимизации их ложных срабатываний и дальнейшей выгрузке корреляционных событий в набор данных (dataset) через шину данных.

Для наглядности приведем следующий пример:

1. Мы хотим обучить математическую модель для анализа доменов третьего уровня на предмет наличия ряда артефактов, которые могут быть в запросе. Например, проверка длины доменного имени:
   - длина доменного имени третьего уровня свыше 20 символов является подозрительной;
   - не человеко-читаемые слова в доменном имени представляют собой дополнительный признак аномалии;
   - фиксация доменного имени в репутационных базах – признак инцидента ИБ.
2. Производим подключение DNS к SIEM и настройку правил корреляции на выявление длины доменов третьего уровня, в случае превышения длины запроса список доменов помещается в лист.
3. Лист отправляется на API лингвистического анализа для получения семантического коэффициента распознания имени домена.
4. Как дополнительная мера производится отправка запроса в платформу Threat Intelligence для анализа доменного имени.

Таким образом, мы получаем многоуровневую систему анализа доменных имен и составления списков "нормальных" и "ненормальных" имен для дальнейшей загрузки в набор данных (dataset).

Основная идея данного примера заключается в демонстрации необходимости применения автоматических интеграций с сервисами, которые могут уменьшить время подготовки набора данных для обучения модели и сэкономить ресурсы специалистов DS/DM.

Хотелось бы сделать пометку, что приведенные выше примеры являются абстрактными и недостаточными механизмами для решения каких-либо задач. Выбор решений и путей всегда остается на стороне компании и/или партнера, и компания "ДиалогНаука" со своей стороны готова предоставить услуги настройки интеграций с SIEM системой, создания интеграционных шин-данных, а также настройку и отладку контента любой сложности.

Рост эффективности

В данной статье на конкретных кейсах были продемонстрированы недекларированные возможности, используемые в SIEM-системах, и то, как они позволяют повысить эффективность ситуационного центра ИБ в целом. Необходимость автоматизации процессов обработки инцидентов ИБ является одним из краеугольных камней в работе любого подразделения защиты информации. Чтобы эффективно управлять системами мониторинга событий ИБ, необходимо перципировать (воспринимать) и уметь задействовать весь спектр возможностей SIEM-систем. ●

Ваше мнение и вопросы присылайте по адресу is@groteck.ru

Для использования более сложных корреляционных логик и выявления более сложных цепочек инцидентов необходимо реализовывать скоринговую модель для подсчета баллов по каждому пользователю и узлу, участвующему в инцидентах ИБ.

SIEM-система собирает огромное количество событий ИБ и производит их нормализацию, агрегацию и корреляцию – данная информация и будет использоваться в дальнейшем для обучения нейросети и минимизации трудозатрат специалистов по сбору и анализу данных.