Контекстное выявление аномалий сервисных учётных записей методами машинного обучения

Березкин Ф.Е.

ФГБОУ ВО "Брянский государственный технический университет",

г. Брянск, Российская Федерация

Сервисная учётная запись для ночного резервного копирования ежедневно обращается к базе данных с одного узла. Если в два часа ночи она входит с другого сервера и открывает административный интерфейс, время активности остаётся привычным, а действие – нет. Правило, учитывающее только часы работы или число входов, не покажет разницы между такой сессией и штатным копированием. Для оценки служебной активности важен контекст: какая запись, с какого узла и к какому ресурсу обратилась, а также что произошло после входа.

Сведения для такой оценки уже присутствуют в журналах аутентификации и аудита: время, идентификатор учётной записи, исходный узел, целевой ресурс, результат входа и последующие операции. Исследование И. С. Лошак и Е. Ю. Костюченко показывает, что при анализе журналов полезны временные, параметрические и количественные признаки [1]. Для сервисной записи их следует связывать в сеанс по одному запуску задания. Отдельное успешное подключение может быть обычным, тогда как цепочка из нескольких отказов и обращения к новому ресурсу требует внимания.

Общий профиль всех сервисных записей здесь малоинформативен. Резервное копирование обычно выполняется ночью, интеграция с внешней системой – по иному расписанию, а обновление программного обеспечения порождает редкие, но разрешённые сеансы. Шелухин О.И. и Рябинин В.С. отмечают, что журнальное событие без контекста не всегда достаточно для поиска причины аномалии [2]. Поэтому обычное поведение описывается отдельно для каждой записи: время запуска, исходные узлы, пары «узел – ресурс», число операций и порядок событий.

Рассмотрим условный фрагмент журнала за 30 дней. Из 40 сеансов одной записи 34 прошли с узла A к базе данных, три – с узла B к той же базе, ещё три – с узла A к административному интерфейсу. Новый сеанс с узла B к административному интерфейсу начался в привычные два часа ночи. Время, узел и ресурс по отдельности уже встречались; новым является их сочетание. Оценка частот отдельных значений может не выделить сеанс, а признак редкости пары «узел – ресурс» направит его на проверку. Эти числа иллюстрируют логику признака, а не результат испытания системы.

Сеансы можно ранжировать алгоритмом поиска выбросов, например изолирующим лесом, по числовому представлению признаков. Время запуска кодируется циклически, а для узла и ресурса учитываются частота их появления у данной записи и частота совместного появления. Дополнительные признаки – число отказов перед успешным входом и отклонение объёма операций от обычного. В обучающую выборку включаются периоды, подтверждённые владельцем сервиса как штатные. Подходы к поиску необычных действий субъектов доступа по журналам аудита с помощью машинного обучения систематизированы К. А. Гайдуком и А. Ю. Исхаковым [3].

Пока новая сервисная запись не накопила историю, её профиль дополняется данными записей с тем же назначением. После накопления подтверждённых сеансов вес индивидуальной истории повышается. Перенос задания на новый сервер сначала даст сигнал об аномалии; владелец системы может подтвердить изменение. Обновлять профиль только по факту повторения событий нельзя: серия несанкционированных обращений постепенно стала бы для него нормой.

Оценка одного сеанса не заменяет анализ последовательности. Вход с нового узла, серия отказов и обращение к нескольким ресурсам вместе дают более веское основание для проверки. Методы машинного обучения для анализа связанных системных событий рассмотрены И. В. Котенко и Д. А. Левшуном [4]. Сигнал должен пояснять, какие сочетания не встречались ранее, как сеанс расходится с расписанием и какие действия последовали за входом. Так

специалист сможет отличить подозрительную активность от согласованного изменения задания.

Аномальный балл определяет очередь проверки, но не доказывает компрометацию записи. Качество модели оценивается на размеченных сеансах с разделением по времени. При одинаковом числе сигналов для просмотра её следует сравнить с правилами по часам и частоте входов: по доле подтверждённых инцидентов среди первых сигналов, числу пропусков и нагрузке на аналитика. Пока такой проверки нет, преимущество контекстного анализа показано лишь на разборе типовых сценариев.

Список использованной литературы

1. Лошак И.С., Костюченко Е.Ю. *Продлённая аутентификация на основе анализа журналов пользователя в ОС* // Доклады Томского государственного университета систем управления и радиоэлектроники. 2025. Т. 28, № 4. С. 39–49. URL: https://journal.tusur.ru/ru/arhiv/4-2025/prodlyonnaya-autentifikatsiya-na-osnove-analiza-zhurnalov-polzovatelya-v-os (дата обращения: 27.09.2026).

2. Шелухин О.И., Рябинин В.С. *Обнаружение аномалий больших данных неструктурированных системных журналов* // Вопросы кибербезопасности. 2019. № 2(30). С. 36–41. URL: https://cyberrus.info/wp-content/uploads/2019/07/36-41-230-19_5.-Sheluhin.pdf (дата обращения: 27.09.2026).

3. Гайдук К.А., Исхаков А.Ю. *К вопросу о реализации алгоритмов выявления внутренних угроз с применением машинного обучения* // Вестник СибГУТИ. 2022. № 4. С. 80–95. URL: https://www.ipu.ru/node/72074 (дата обращения: 27.09.2026).

4. Котенко И.В., Левшун Д.А. *Методы интеллектуального анализа системных событий для обнаружения многошаговых кибератак: использование методов машинного обучения* // Искусственный интеллект и принятие решений. 2023. № 3. С. 3–15. URL: https://www.mathnet.ru/rus/iipr32 (дата обращения: 27.09.2026).