УДК 004.056.53:004.89

Федін Вадим Віталійович,

здобувач другого (магістерського) рівня вищої освіти
Харківського національного університету радіоелектроніки

Сєвєрінов Олександр Васильович,

к.н.н., професор професор кафедри безпеки інформаційних технологій
Харківського національного університету радіоелектроніки

АНАЛІЗ ПОВЕДІНКОВИХ АНОМАЛІЙ ПРАЦІВНИКІВ ДЛЯ РАННЬОГО ВИЯВЛЕННЯ ІНСАЙДЕРСЬКИХ ЗАГРОЗ

Анотація. Розглянуто проблему своєчасного виявлення потенційно небезпечної активності працівників, які використовують легітимні облікові записи корпоративних інформаційних систем. Обґрунтовано структуру гібридного аналізу поведінки, що поєднує Isolation Forest, Autoencoder та LSTM. Описано джерела подій, формування поведінкових ознак, процедуру оцінювання аномальності та інтерпретацію сповіщень. Для пояснення критеріїв якості наведено контрольний розрахунок за матрицею помилок LSTM, поданою у вихідних матеріалах дослідження; розмежовано розрахункові показники окремого алгоритму й ефективність гібридної системи, яка потребує незалежної експериментальної перевірки.

Ключові слова: інсайдерські загрози, поведінкова аналітика, UEBA, машинне навчання, Isolation Forest, Autoencoder, LSTM, оцінювання ризику.

У корпоративних системах користувач із чинними повноваженнями може виконувати дії, що формально не порушують правил доступу, проте в сукупності свідчать про підготовку до витоку даних або зловживання привілеями. Наприклад, нетиповий час входу, масове читання документів і незвичний обсяг передавання файлів окремо можуть мати законні пояснення. Їх поєднання, зіставлене зі звичайним режимом роботи конкретного працівника, потребує перевірки фахівцем із безпеки. Тому метою є не автоматичне встановлення вини, а формування обґрунтованого сигналу для раннього аналізу ризику. Виявлення таких відхилень пов’язане з технологією UEBA (User and Entity Behavior Analytics), що ґрунтується на порівнянні поточної активності з характерними моделями поведінки [1, 2].

Запропонований підхід використовує події автентифікації, відомості про доступ до інформаційних ресурсів, операції з файлами та мережеву

активність. Після нормалізації записи пов’язуються з користувачем і часовим інтервалом. Формуються ознаки, зокрема кількість входів, частка дій поза робочим часом, інтенсивність звернень до файлів, обсяг переданих даних і відмінність поточної активності від попередніх періодів. Водночас необхідно враховувати роль працівника, зміну завдань і обґрунтованість його доступу.

Аналітичний модуль складається з трьох компонентів. Isolation Forest визначає статистично нетипові вектори поведінкових ознак, Autoencoder оцінює відхилення складних комбінацій ознак за похибкою відновлення, а LSTM опрацьовує впорядковані часові послідовності дій. Оцінки приводяться до узгодженої шкали та можуть об’єднуватися з контекстом доступу для формування інтегрального рівня ризику. Перевищення встановленого порога слугує підставою для сповіщення й подальшого аналізу, а не доказом інсайдерського порушення.

Для оцінювання придатності моделі передбачено використання CERT Insider Threat Dataset [3], який містить синтетичні сценарії поведінки користувачів. Відтворювана перевірка потребує зазначення версії набору, способу формування часових спостережень, розмежування навчальних і тестових періодів, конфігурації алгоритмів і порогів сповіщення. Дані тестового періоду не повинні використовуватися для навчання або налаштування порогів.

Отже, запропонована архітектура поєднує статистичні, багатовимірні й часові характеристики активності працівників і визначає послідовність переходу від подій журналювання до пріоритизованого сповіщення про потенційний ризик. Її прикладна цінність полягає в можливості доповнення засобів SIEM контекстною поведінковою аналітикою.

Список використаних джерел

1. Пасєка І., Сєвєрінов О. В. Проблеми впровадження системи аналітики

поведінки користувачів та сутностей. Global Cyber Security Forum: матеріали Першого міжнародного науково-практичного форуму, 14–16 листопада 2019 р. Харків: ХНУРЕ, 2019. С. 82–83.
2. Сєвєрінов О.В., Чубенко О.В. Аналіз проблем використання SIEM систем. Проблеми інформатизації: тези доповідей VIII Міжнародної науково-технічної конференції, 26–27 листопада 2020 р. 2020. Т. 1. С. 111.
3. Carnegie Mellon University, Software Engineering Institute. Insider Threat Test Dataset. 2016. DOI: https://doi.org/10.1184/R1/12841247.v1.