УДК 004.056

МОДЕЛИ ПРЕДСКАЗАНИЯ КИБЕРАТАК С ИСПОЛЬЗОВАНИЕМ МАШИННОГО ОБУЧЕНИЯ

Киселев Р.Г.
магистр, Южный федеральный университет
(Ростов-на-Дону, Россия)

CYBERATTACK PREDICTION MODELS USING MACHINE LEARNING

Kiselev R.
master's degree, Southern Federal University
(Rostov-on-Don, Russia)

Аннотация

В статье рассматриваются методы машинного обучения (МО) для предсказания кибератак и повышения уровня кибербезопасности. Основное внимание уделено алгоритмам классификации и методам обнаружения аномалий, которые позволяют выявлять как известные, так и новые типы атак. Приведены примеры применения классификационных моделей, таких как случайные леса, и методов обнаружения аномалий, таких как изоляционный лес, для анализа сетевой активности. Также обсуждаются перспективы использования глубокого обучения и методов МО с подкреплением для анализа временных рядов и динамических данных. В заключении подчеркивается значимость данных подходов для обеспечения устойчивости информационных систем в условиях постоянно меняющейся киберугрозы.

Ключевые слова: кибератаки, машинное обучение, обнаружение аномалий, классификация, кибербезопасность.

Abstract

The article examines machine learning (ML) methods for predicting cyberattacks and enhancing cybersecurity. The focus is on classification algorithms and anomaly detection techniques that identify both known and unknown types of attacks. Examples of applying classification models, such as random forests, and anomaly detection methods, such as isolation forests, for network activity analysis are provided. The prospects for using deep learning and reinforcement learning methods to analyze time series and dynamic data are also discussed. The conclusion emphasizes the importance of these approaches for maintaining the resilience of information systems amid constantly evolving cyber threats.

Keywords: cyberattacks, machine learning, anomaly detection, classification, cybersecurity.

Введение

С развитием технологий и расширением использования цифровых инфраструктур возрастает и число кибератак, что представляет серьезную угрозу для безопасности данных и информационных систем. Традиционные методы защиты, такие как системы предотвращения вторжений и антивирусные программы, зачастую не способны эффективно противостоять новым видам атак. В условиях, когда злоумышленники используют продвинутые методы и инструменты, необходимы более гибкие и эффективные подходы для защиты данных. Одним из таких методов является использование алгоритмов машинного обучения (МО), которые способны анализировать большие объемы данных, выявлять аномалии и предсказывать потенциальные атаки.

Модели предсказания кибератак, основанные на МО, позволяют обнаруживать подозрительную активность, анализируя сетевые пакеты, логи и события системы безопасности. Алгоритмы, такие как деревья решений, случайные леса и нейронные сети, могут обучаться на исторических данных, чтобы определять паттерны, характерные для кибератак. На основе этих моделей можно прогнозировать вероятность атаки, что позволяет предпринять меры по защите системы до момента реального вторжения. Преимущество МО заключается в способности обнаруживать новые типы атак, что значительно расширяет возможности традиционных методов безопасности.

Цель данной статьи состоит в рассмотрении различных моделей МО, применяемых для предсказания кибератак, а также анализа их эффективности и применимости в различных условиях. В статье будут описаны основные методы, включая алгоритмы классификации и обнаружения аномалий, а также представлены примеры их практического применения для повышения уровня кибербезопасности.

Основная часть

Алгоритмы классификации, такие как логистическая регрессия, деревья решений и случайные леса, часто применяются для определения, принадлежит ли сетевое событие к категории нормальной активности или кибератаки.

На рисунке 1 представлено распределение типов кибератак, с которыми чаще всего сталкиваются информационные системы. Данные подчеркивают разнообразие угроз и необходимость применения алгоритмов машинного обучения для их обнаружения.

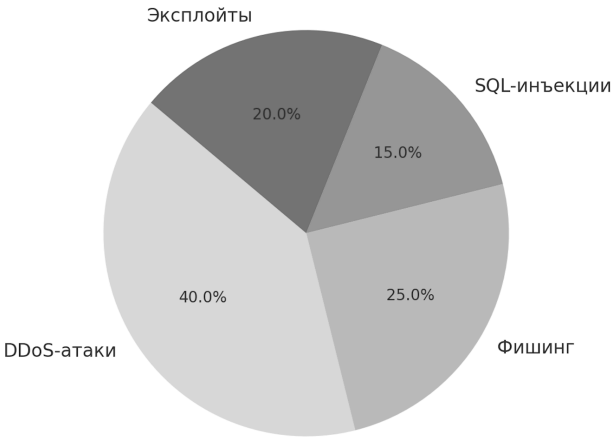

Рисунок 1. Распределение типов кибератак

Как видно из рисунка 1, наиболее частыми типами атак являются DDoS-атаки (40%) и фишинг (25%), что связано с их относительной простотой и высокой эффективностью для злоумышленников. Реже встречаются сложные атаки, такие как SQL-инъекции (15%) и кража данных через эксплойты (20%). Эти данные подтверждают важность разработки универсальных алгоритмов, способных распознавать как стандартные, так и более сложные типы угроз.

Эти модели обучаются на большом объеме данных, содержащих метки классов (например, «атака» и «норма»), что позволяет алгоритму обнаруживать паттерны, характерные для атакующих действий. Примером такого подхода может служить случайный лес, который строит несколько деревьев решений и принимает итоговое решение на основе их совокупного результата [1]. Пример кода на Python для классификации сетевых событий с использованием случайного леса:

```
from sklearn.ensemble import RandomForestClassifier
from sklearn.model_selection import train_test_split
from sklearn.metrics import accuracy_score

# Пример данных, где X - признаки, а y - метки классов (атака или норма)
# X может включать такие параметры, как количество запросов в секунду, IP-адрес, тип пакета и т.д.
```

```
X = [[200, 0.5, 10], [100, 0.1, 5], [300, 0.7, 8], [50, 0.2, 3]]
y = ["норма", "атака", "атака", "норма"]

# Разделение данных на обучающую и тестовую выборки
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.25, random_state=42)

# Инициализация и обучение модели случайного леса
clf = RandomForestClassifier(n_estimators=10, random_state=42)
clf.fit(X_train, y_train)

# Предсказание на тестовой выборке
y_pred = clf.predict(X_test)

# Вывод точности модели
accuracy = accuracy_score(y_test, y_pred)
print(f"Точность классификации: {accuracy * 100:.2f}%")
```

В данном примере используется модель случайного леса для классификации событий на «атаку» и «норму». Модель обучается на небольшом наборе данных, включающем параметры сетевой активности, и затем применяется для предсказания на тестовой выборке. Функция accuracy_score позволяет оценить, насколько точно модель определяет атаки, что является важным показателем ее эффективности [2]. Использование таких алгоритмов классификации дает возможность системам безопасности автоматически реагировать на потенциальные угрозы, снижая нагрузку на специалистов по кибербезопасности. Случайные леса и другие алгоритмы классификации могут эффективно выявлять кибератаки в условиях высокой динамичности сетевой активности, обеспечивая дополнительный уровень защиты для информационных систем. Методы обнаружения аномалий представляют собой еще один важный инструмент в предсказании кибератак. В отличие от классификационных алгоритмов, которые требуют размеченных данных для обучения, методы обнаружения аномалий работают с неразмеченными данными и позволяют выявлять неизвестные типы атак. Это особенно важно для ситуаций, когда системы безопасности сталкиваются с новыми, ранее неизвестными угрозами [3]. Примером подобного подхода является использование изоляционного леса, который строит несколько деревьев для выявления аномальных записей в данных, что позволяет обнаруживать возможные признаки кибератак. Также важную роль играют ассоциативные правила, которые помогают выявлять последовательности событий, предшествующих кибератакам. Эти алгоритмы позволяют найти связи между различными действиями пользователей или сетевой активностью, что может сигнализировать о подготовке атаки. Например, частое повторение определенных запросов или команд может указывать на подозрительную активность, что позволяет системе безопасности предпринять соответствующие меры до начала атаки.

Глубокое обучение также является перспективным направлением в области предсказания кибератак, особенно когда требуется анализировать большие объемы сложных данных. Нейронные сети способны работать с неструктурированными данными, такими как сетевые журналы и текстовые записи, позволяя системе безопасности находить аномальные паттерны. Например, сверточные нейронные сети можно использовать для анализа сетевых пакетов и поиска подозрительных последовательностей, указывающих на угрозу. Рекуррентные нейронные сети (РНС) особенно полезны для анализа временных рядов сетевой активности. Это дает возможность выявлять последовательные аномалии и связывать их с потенциальной атакой, особенно если атака разворачивается постепенно. Использование РНС позволяет моделировать сложные зависимости между событиями, что делает ее незаменимой в анализе данных с временной составляющей, таких как журналы активности системы [4].

Методы МО с подкреплением также находят применение в предсказании кибератак, позволяя системе обучаться через взаимодействие с реальной средой. Такие модели способны

адаптироваться к быстро меняющимся условиям, что особенно актуально для кибербезопасности, где характер угроз может значительно меняться. Алгоритмы с подкреплением помогают системе оптимизировать свои действия, повышая вероятность обнаружения угроз в режиме реального времени и снижая риски, связанные с возможными атаками.

Применение моделей обнаружения аномалий в предсказании кибератак

Модели обнаружения аномалий играют критически важную роль в выявлении кибератак, так как они позволяют обнаруживать отклонения от нормального поведения системы, указывающие на возможную угрозу. В отличие от классификационных моделей, которые требуют размеченные данные для обучения, методы обнаружения аномалий способны работать с неразмеченными данными, что делает их более универсальными в условиях, когда полный набор данных о кибератаках отсутствует или содержит неизвестные типы угроз. Эти модели подходят для случаев, когда атакующие используют новые техники, еще не встречавшиеся в исторических данных.

Основной принцип работы моделей обнаружения аномалий заключается в создании профиля «нормального» поведения системы на основе сетевых пакетов, активности пользователей или логов системы, после чего любые значительные отклонения от этого профиля классифицируются как аномалии [5]. Например, если в обычной сетевой активности среднее количество запросов к серверу составляет 50 в минуту, а вдруг происходит всплеск до 200 запросов, это может свидетельствовать о начале DDoS-атаки.

Изоляционный лес (Isolation Forest) – это один из алгоритмов, эффективных для обнаружения аномалий в кибербезопасности. Алгоритм строит множество деревьев решений, каждое из которых изолирует случайные точки данных. При этом аномалии (выбросы) изолируются на более низких уровнях дерева, так как для их выделения требуется меньше шагов из-за их редкости или необычности. Такой подход позволяет алгоритму выделять подозрительные события на основе отклонений от нормы без необходимости разметки данных.

Методы обнаружения аномалий обладают способностью выявлять ранее неизвестные типы атак, так как они не ограничены размеченными данными. Однако данные модели чувствительны к качеству исходных данных и профилю «нормального» поведения, на основе которого строятся прогнозы [6]. Например, если модель обучена на данных, содержащих аномалии, они могут быть приняты за норму, что снизит точность предсказаний. Применение моделей обнаружения аномалий для предсказания кибератак позволяет улучшить проактивный подход к безопасности, реагируя на угрозы до их непосредственного проявления [7].

Заключение

Использование алгоритмов МО для предсказания кибератак позволяет повысить эффективность и проактивность систем безопасности. Классификационные алгоритмы, такие как случайные леса, показывают высокую точность в определении известных атак на основе исторических данных, что позволяет системам безопасности оперативно реагировать на угрозы. Эти модели дают возможность автоматизировать обработку и анализ данных, снижая нагрузку на специалистов по безопасности и позволяя оперативно предотвращать потенциальные атаки. Методы обнаружения аномалий также занимают важное место в предсказании кибератак, так как они способны выявлять ранее неизвестные угрозы, что особенно важно в условиях постоянно меняющейся киберугрозы. Такие алгоритмы, как изоляционный лес, помогают выделить отклонения от нормального поведения и предсказать потенциальные угрозы, даже если их паттерны ранее не были зафиксированы в данных. Это позволяет поддерживать высокий уровень безопасности в условиях, когда новые техники атак появляются регулярно.

Список литературы

1. Мишин А.Е., Былевский П.Г. Применение машинного обучения для прогнозирования кибератак // Hi-Hume Journal. 2023. Т. 3. С. 80.

2. Асяев Г.Д., Соколов А.Н. Модели предиктивной защиты информации автоматизированной системы управления водоснабжением на основе временных рядов с использованием технологий машинного обучения // Вестник УрФО. Безопасность в информационной сфере. 2021. №4 (42). С. 39-45.
3. Кечеджиев А.С., Цветкова О.Л., Дубровина А.И. Методика выявления аномалий в данных оценки кибератак с использованием Random Forest и градиентного бустинга в машинном обучении // Вестник Дагестанского государственного технического университета. Технические науки. 2024. Т. 51. №3. С. 72-85.
4. Власенко А.В., Дзьобан П.И., Жук Р.В. Обзор инструментов машинного обучения и их применения в области кибербезопасности // Прикаспийский журнал: управление и высокие технологии. 2020. №1 (49). С. 144-155.
5. Боброва М.В., Мастилин А.Е. Машинное обучение в кибербезопасности // Научные междисциплинарные исследования. 2021. №2. С. 24-29.
6. Котенко И.В., Саенко И.Б., Лаута О.С., Крибель А.М. Методика обнаружения аномалий и кибератак на основе интеграции методов фрактального анализа и машинного обучения // Информатика и автоматизация. 2022. Т. 21. №6. С. 1328-1358.
7. Ibor A.E., Oladeji F.A., Okunoye O.B., Ekabua O.O. Conceptualisation of Cyberattack prediction with deep learning // Cybersecurity. 2020. Vol. 3. P. 1-14.