Добрышин Михаил Михайлович
Академии ФСО России, сотрудник к.т.н., г. Орёл
Жиляева Валерия Алексеевна
Академии ФСО России, сотрудник, г. Орёл

ОБРАБОТКА РЕЗУЛЬТАТОВ МОДЕЛИРОВАНИЯ КОМПЬЮТЕРНЫХ АТАК НА ОБЪЕКТ ЗАЩИТЫ С ИСПОЛЬЗОВАНИЕМ МЕТОДОВ МАШИННОГО ОБУЧЕНИЯ

Задача снижения ущерба от инцидентов информационной безопасности на текущий момент времени является одной из актуальных, а новости об успешных реализациях компьютерных атак регулярно появляются у всех информационных агентств. В статье представлен вариант выявления факта реализации атаки основанный на анализе эксплуатационных характеристик защищаемых объектов с применение комплекса средств машинного обучения.

Ключевые слова: компьютерные атаки, машинное обучение, эксплуатационные характеристики, дерево состояния.

Mikhail Mikhailovich Dobryshin
Academy of the Federal Guard Service of Russia, PhD in Engineering, Oryol
Valeria Alekseevna Zhilyaeva
Academy of the Federal Guard Service of Russia, employee, Oryol

PROCESSING OF SIMULATION RESULTS OF COMPUTER ATTACKS ON THE OBJECT OF PROTECTION USING MACHINE LEARNING METHODS

The task of reducing damage from information security incidents is currently one of the most urgent, and news about successful computer attacks regularly appears in all news agencies. The article presents an option for detecting the fact of an attack based on an analysis of the operational characteristics of protected objects using a set of machine learning tools.

Keywords: computer attacks, machine learning, performance characteristics, state tree

Интеграция инфо-телекоммуникационных систем во все сферы деятельности современного общества позволило в значительной мере ускорить

процессы управления, принятия решений, «действий» и «жизни» отдельного человека и общества. Однако совместно со множеством выгод внедрения технологичных систем возникают актуальная задача защиты этих систем от различных компьютерных атак (КА). Выход систем отвечающих за управлением деятельностью городской или транспортной инфраструктуры, финансового сектора, средств массовой коммуникации способно привести как значительным финансовым убыткам, так и к тяжелым последствиям для населения.

В связи с этим, задача по обеспечению информационной безопасности (ИБ) инфо-телекоммуникационных систем не только не теряет актуальности. С этой целью совершенствуются и разрабатываются новые средства, методы и способы обеспечения ИБ, однако существенная часть из них применяет традиционные подходы выявления фактов реализации КА. Под традиционными подходами понимается не только средства реализующие сигнатурный анализ, но и средства, анализирующие параметры, характеризующие атаки. Данный подход показывает свою эффективность в отношении известных тактик, техник, способов и средств реализации атак, и не в полной мере эффективен при анализе новых, ранее не применяемых атак.

В качестве возможного направления изменения парадигмы разработки средств обеспечения ИБ является применения подходов, направленных не на выявление в условиях неопределенности фактов (поиск конкретных значений параметров) реализации КА, а обработка и анализ эксплуатационных характеристик защищаемых объектов с целью выявления факта реализации КА [1]. Данный подход не исключает применение традиционных решений, а дополняет их.

Суть предложения заключается в выявлении схожести влияния конкретных видов воздействия на контролируемые эксплуатационные характеристики исследуемых объектов. В качестве примера, возможно, рассмотреть факт заражения вредоносным программным обеспечением локальной сети. В качестве ущерба наносимого такой атакой следует рассматривать не нарушение работоспособности одного компьютера (ПК) входящего в состав локальной сети, а нарушение работоспособности всей сети (предполагая, что если один ПК заражен вирусом, то применяемое в сети

антивирусное средство не способно выявить и локализовать этот вирус на других устройствах).

С целью реализации указанного замысла выявления вируса в сети и минимизации ущерба, задача декомпозирована на следующие этапы [2]:

– измерение, сбор, обработка и нормализация измеренных значений параметров эксплуатационных характеристик каждого ПК;

– сопоставление текущих значений с базой возможных состояний;

– принятие решения по реагированию, при выявлении факта реализации атаки.

Формирование базы возможных состояний возможно на основе проведения натурных экспериментов и кластеризации результатов.

Указанные частные задачи обработки данных, возможно, обработать на основе применения традиционных статистических методов (наивный байесовский классификатор, дискриминантный анализ, логистическая регрессия), однако при наличии значительного объема выборки (в настоящий момент процесс функционирования ПК возможно описать на основе изменения 12 эксплуатационных характеристик) их использование потребует значительных вычислительных ресурсов (для выявления факта деструктивного действия вируса на ПК и сопоставление его конкретному типу необходимо сопоставить изменение не одного параметра, а всего картежа параметров, причем это необходимо сделать не точечно, а интервально – в течение некоторого времени наблюдения) и привлечения в организацию дополнительных специалистов.

Вместе с этим, в настоящее время активно развиваются инструменты на основе применения методов машинного обучения [3-7]:

Метод k-средних – один из наиболее распространённых алгоритмов кластеризации, относящийся к методам обучения без учителя (unsupervised learning). Алгоритм разбивает множество объектов (наблюдений) на заранее заданное число k непересекающихся кластеров таким образом, чтобы каждый объект принадлежал кластеру с ближайшим центром (центроидом), а суммарное внутрикластерное расстояние было минимальным. Данный метод получил

развитие и был усовершенствован в ряде решений, например, k-ближайших соседях.

Линейный дискриминантный анализ (LDA – Linear Discriminant Analysis)

Проецирует данные на подпространство, максимизирующее межклассовое расстояние и минимизирующее внутриклассовое. Одновременно снижает размерность и классифицирует. Предполагает нормальное распределение признаков и равные ковариационные матрицы классов

Метод опорных векторов (SVM – Support Vector Machine) – позволяет формировать оптимальную разделяющую гиперплоскость с максимальным разделением (margin) между классами. Ядровое преобразование (kernel trick) позволяет строить нелинейные границы: полиномиальное ядро, RBF (радиальная базисная функция), сигмоидное. Метод эффективен в пространствах высокой размерности и устойчив к переобучению при правильном выборе параметров.

Наивный байесовский классификатор (Naive Bayes) – Основан на теореме Байеса с допущением о условной независимости признаков. Варианты: гауссовский, мультиномиальный, бернуллиев. Применение данного инструмента позволяет быстро обучать модели для классификации. Эффективен для текстовых данных (спам-фильтрация, классификация документов). Допущение о независимости редко выполняется, но метод всё равно работает на практике.

Байесовские сети (Bayesian Networks) – графические вероятностные модели, описывающие условные зависимости между переменными. Допускают частичную зависимость признаков. Интерпретируемость через структуру графа.

Случайный лес (Random Forest) – ансамблевый метод, основанный на множестве деревьев решений, каждый из которых обучается на случайных подмножествах данных и признаков. Позволяет строить модель классификации, которая на входе получает вектор признаков (многомерный кортеж) телеметрии, а на выходе выдает метку класса, например, normal, xmrig, wannacry и т.д. Случайный лес позволяет оценить, какие признаки важны для классификации. Для каждого признака вычисляется среднее снижение индекса Джини по всем деревьям и узлам.

Градиентный бустинг (Gradient Boosting) – последовательное построение ансамбля слабых моделей (обычно деревьев), каждая из которых корректирует ошибки предыдущих. Высочайшая точность на табличных данных. Обработка пропусков, категориальных признаков (CatBoost). Риск переобучения при неправильной настройке.

Анализ практических решений показывает, что на текущем этапе развития программных средств, выделить наиболее эффективное средство достаточно сложно, а учитывая то, что для решения задач применяют комплекс моделей, рационально оценивать не эффективность моделей ИИ, а расходуемые вычислительные ресурсы, предоставляемый функционал.

В качестве варианта решения рассмотренной задачи, возможно применять программное средство объединяющее следующие модели обработки данных: *k-ближайших соседей, случайный лес.*

Разработанная программа [8] позволяет выявить кортежи значений контролируемых параметров СОИ при реализации в отношении него атак. Данное средство, фиксирует и выявляет аномалии изменения контролируемых параметров СОИ, вызванные, в том числе реализацией КА. Применение ПО позволяет определить наиболее информативные параметры и выявить сочетание параметров, характеризующих нормальные состояния СОИ и условия, характеризующие реализацию различных видов КА, в том числе ранее не известных.

Графическое представление применения метода k-ближайших соседей в разработанной программы представлено на рисунках 1-3. Применение k-ближайших соседей позволяет выявить важность контролируемых параметров и сократить количество обрабатываемых параметров.

```
Результаты обнаружения аномалий:
Всего записей: 1000
Обнаружено аномалий: 50 (5.00%)

Важность признаков для предсказания скорости записи:
CPU (%): 0.6388
RAM (%): 0.3612

Средняя абсолютная ошибка (MAE) предсказания скорости записи: 2.7058 MB/s

Статистика по подозрительным записям:
        RAM (%)     CPU (%)  ...  Anomaly  Predicted_Disk_Write
count  50.000000   50.000000  ...     50.0             50.000000
mean   58.696000   66.998000  ...      1.0             36.274764
std     7.236048   42.163378  ...      0.0             83.330955
min    49.700000    0.000000  ...      1.0              0.000000
25%    52.750000   15.400000  ...      1.0              0.021440
50%    56.500000  100.000000  ...      1.0              0.096858
75%    63.700000  100.000000  ...      1.0              6.285938
max    74.900000  100.000000  ...      1.0            374.576074

[8 rows x 5 columns]

Примеры подозрительных записей:
    RAM (%)  CPU (%)  Disk Write Speed (MB/s)  Anomaly  Predicted_Disk_Write
15     51.0     17.4                 0.000000        1             59.690625
19     63.0    100.0                 0.836426        1              0.270271
22     56.5    100.0                 0.000000        1              0.006726
23     56.5    100.0                 0.000000        1              0.006726
24     56.5    100.0                 0.000000        1              0.006726
```

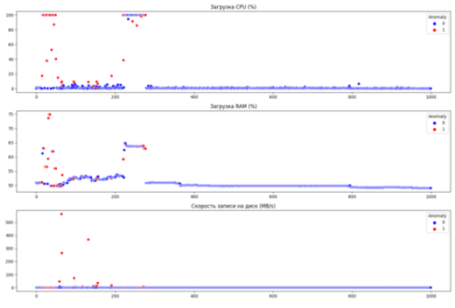


Рис. 1. Графическое представление измеренных параметров СОИ при его заражении ВПО

Рис. 2. Графическое представление определения важности контролируемых параметров

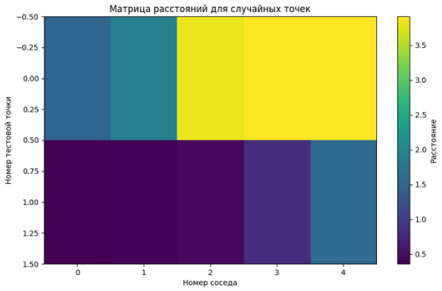
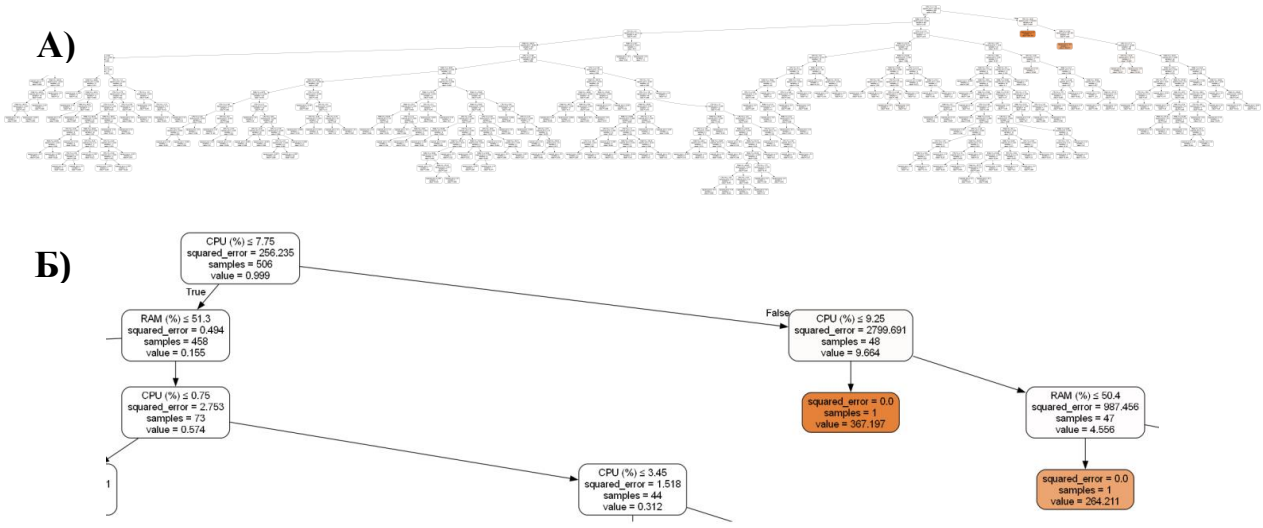


Рис. 3. Графическое представление матрицы расстояний для результатов эксперимента

Обработка собранного набора (кортежей) данных реализовано с применением случайного леса (рис. 4), что позволило определить условия, характеризующие состояние ПК, которое характеризует действие вируса.



А) Полный вид дерева состояний ПК в различных условиях функционирования
Б) Фрагмент дерева состояний ПК в условиях реализации атаки

Рис. 4. Графическое представление дерева состояний СОИ при реализации атаки на него

Собранные наборы параметров подтверждают гипотезу о том, что анализ значений параметров, характеризующих эксплуатационные характеристики, потенциально способен выявить факт реализации неизвестных КА, а также являются основой для формирования базы данных для выработки стратегии защиты для случаев изменения значений контролируемых параметров.

Основываясь на том, что выявление аномалий параметров, описывающих состояние исследуемого СОИ производится в рамках контролируемых экспериментов (известно время начало и окончания воздействия, понятны цели воздействия, в том числе непосредственно изменяемые контролируемые параметров), результаты отражают реализуемые и протекающие процессы и не несут искажений (ошибок).

Применение такого подхода в отличие от традиционных методов выявления признаков реализации КА основанных на знаниях о возможных признаках КА, позволяет обнаружить факт реализации неизвестной КА в отношении ПК, что в дальнейшем позволит за счет своевременного реагирования и выбора стратегии защиты повысить уровень безопасности элемента сети (*например, при обнаружении факта заражения ПК вирусом, локализовать его, тем самым защитить локальную сеть от нанесения ущерба*).

Список литературы:

1. Добрышин М.М. Порядок, формирование и подтверждение гипотез, и их влияние на парадигму теории информационной безопасности / Международный научно-практический электронный журнал «Экономика и качество систем связи» : – 2025. – № 3 (37). – С. 148-156.

2. Белов А. С., Добрышин М. М., Душкин А. В. Системы обеспечения информационной безопасности: системный анализ, синтез, управление обработкой информации / Учебное пособие для вузов. под науч. ред. А. В. Душкина // М. : Горячая линия – Телеком, 2023. – 232 с.

3. Намиот Д. Е., Ильюшин Е. А., Чижов И. В. Искусственный интеллект и кибербезопасность / International Journal of Open Information Technologies ISSN : 2307-8162 vol. 10, no 9, 2022. С. 135-144.

4. Рассел, Стюарт, Норвиг, Питер. Искусственный интеллект: современный подход, 4-е изд., том 1. Решение проблем: знания и рассуждения. : Пер. с англ. - СПб. : ООО "Диалектика", – 2021. – 704 с.

5. Рассел, Стюарт, Норвиг, Питер. Искусственный интеллект: современный подход, 4-е издание, том 3. Обучение, восприятие и действие: Пер. с англ. – СПб.: ООО "Диалектика", – 2022. – 640 с. - Парал. тит. англ.

6. Исаков А. А. Искусственный интеллект и расследование киберпреступлений / Вестник науки – № 5 (62) Т.3 – С. 597-602.

7. Добрышин М. М. К вопросу применения в средствах обеспечения информационной безопасности элементов доверенного искусственного интеллекта / Международный научно-практический электронный журнал «Экономика и качество систем связи» : – 2025. – № 4 (38). – С. 118-126.

8. Добрышин М. М. Программный модуль выявления вредоносного программного обеспечения, на основе дерева состояния, защищаемого ЭВМ / Свидетельство о государственной регистрации программы для ЭВМ № 2025666 120 от 23.06.2025 Бюл. № 7.