ОБНАРУЖЕНИЕ ОБФУСЦИРОВАННЫХ ЭКСПЛОИТОВ В ФАЙЛАХ НЕИСПОЛНЯЕМЫХ ФОРМАТОВ

Архипов А. Н.[1], Кондаков С. Е.[2]

DOI: 10.21681/2311-3456-2024-6-65-75

Цель исследования: *разработка модели бинарной классификации файлов неисполняемых форматов, обеспечивающей повышение эффективности выявления обфусцированных эксплоитов относительно моделей, реализованных в существующих средствах антивирусной защиты.*

Методы исследования *базируются на положениях теории вероятности и математической статистики, теории множеств, методов проведения натурного эксперимента и обработки экспериментальных данных.*

Результат: *в ходе исследования на базе математической модели эксплоита сгенерировано множество потенциальных признаков, которые представлены численными значениями. Из сформированного признакового пространства осуществлен отбор информативных признаков и построение модели бинарной классификации, обладающей наилучшими показателями эффективности в выявлении обфусцированных эксплоитов. Разработана программа для ЭВМ, реализующая полученную модель. Эффективность разработанной модели подтверждена в рамках проведенных экспериментальных исследований по оценке эффективности выявления обфусцированных эксплоитов с использованием средств антивирусной защиты, включенных в реестр российского программного обеспечения, и средств антивирусной защиты иностранного производства, размещенных в свободном доступе, и авторской программы.*

Научная новизна *результатов определяется совокупностью авторских процедур, обеспечивающих выбор классификатора, его гиперпараметров, а также формирование информативного признакового пространства, включая признаки, разработанные авторами, и, позволяющих построить наиболее эффективную модель бинарной классификации, что обеспечивает обоснованность полученных результатов. Представлено подтверждение реализуемости и получения лучших значений показателей эффективности при выявлении обфусцированных эксплоитов относительно существующих средств антивирусной защиты.*

Практическая значимость: *представленная модель, в первую очередь, ориентирована на применение в системах антивирусной защиты, но может быть использована и для решения других задач обеспечения информационной безопасности.*

Ключевые слова: *кибербезопасность, компьютерные атаки, вредоносный код, защита информации, системы антивирусной защиты информации, система обнаружения вторжений.*

Введение

Проблема выявления угроз нарушения информационной безопасности, реализуемых посредством локальных эксплоитов, распространяемых в файлах неисполняемых форматов и позволяющих за счет эксплуатации уязвимостей исполнять произвольный код (далее – эксплоиты), является одной из актуальных и недостаточно решенных вопросов обеспечения информационной безопасности [1, 2].

Для решения указанной проблемы уже несколько десятилетий применяются средства антивирусной защиты информации.

Реализованные в данных средствах защиты информации модели и алгоритмы обнаружения вредоносного кода, главным образом, базируются на устойчивых статических и поведенческих паттернах, выделенных из уже выявленных и изученных образцов эксплоитов.

На практике указанный подход показывает хорошие результаты при обнаружении только известных эксплоитов и одновременно плохо справляется с выявлением эксплоитов, созданных на базе известных уязвимостей с применением технологий обфускации и уязвимостей нулевого дня.

В данной статье будет рассмотрено только обнаружение обфусцированных эксплоитов, так как данный инструмент проведения компьютерных атак значительно чаще встречается на практике [3, 4].

Общий алгоритм действий злоумышленника при распространении вредоносного программного обеспечения в форме обфусцированных эксплоитов представлен на рисунке (рис. 1).

Процедура обфускации позволяет значительно снизить эффективность обнаружения вредоносного кода применяемыми средствами антивирусной защиты в файлах неисполняемых форматов при проведении компьютерных атак, что подтверждают соответствующие экспериментальные исследования [5].

Одновременно файлы неисполняемого формата не вызывают подозрения у пользователей и в сочетании с методами социальной инженерии

1 Архипов Александр Николаевич, студент МГТУ им. Баумана, г. Москва, Россия. E-mail: diskpart111@mail.ru
2 Кондаков Сергей Евгеньевич, кандидат технических наук, доцент, МГТУ им. Баумана, г. Москва, Россия. E-mail: sergeikondakov@list.ru

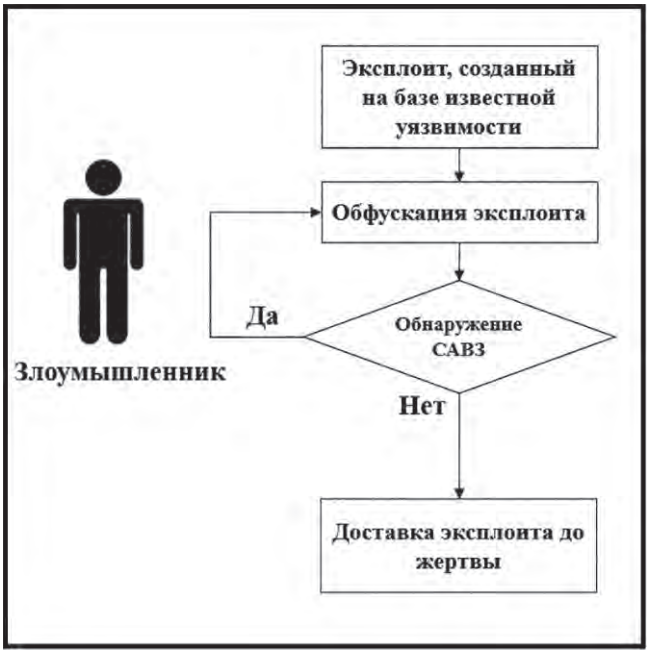

Рис. 1. Процесс подготовки эксплоита для обхода средств антивирусной защиты (САВЗ)

позволяют атакующему распространять вредоносное программное обеспечение под видом легитимного посредством электронной почты, социальных сетей и различных мессенджеров.

Таким образом актуализируется задача совершенствования научно-методического аппарата обнаружения обфусцированных эксплоитов особенно в условиях открытого распространения техник и технологий обфускации в сети Интернет, включая программное обеспечение для автоматизации данного процесса.

Обзор релевантных работ

С формальной точки зрения указанная проблема обнаружения эксплоитов сводится к решению научной задачи бинарной классификации файлов неисполняемых форматов на «безопасные» и «вредоносные».

Эффективное решение указанной задачи классификации зависит от качества используемого признакового пространства, описывающего объект классификации, и классификатора, которые в совокупности формируют модель бинарной классификации.

С учетом изложенного дальнейшее рассмотрение моделей бинарной классификации, предлагаемых в научной и практической литературе, будем осуществлять с точки зрения подходов к формированию признакового пространства и выбора классификатора.

Проведем анализ работ, в которых представлены модели бинарной классификации, разработанные для обнаружения вредоносного программного обеспечения, распространяемого в форме эксплоитов в файлах неисполняемых форматов.

В работах [6–8] признаковое пространство формируется по результатам поиска и анализа в исследуемых файлах модулей, которые потенциально могут содержать исполняемый код: макросы VBA, OLE-объекты, вставки кода JavaScript, которые являются легитимными и определены форматом. При этом предметом поиска являются устойчивые потенциально опасные конструкции, ранее выявленные в других вредоносных файлах.

В качестве классификаторов отобраны классические решения такого рода задач алгоритмы: деревья решений, метод опорных векторов, наивный байесовский классификатор, метод k-ближайших соседей, а также их комбинация: объединение, пересечение или выбор одного из перечисленных.

В публикациях [9, 10] признаковое пространство формируется посредством извлечения типовых последовательностей различных системных вызовов API, которые формируются на базе моделей обработки естественного языка (NLP).

В качестве классификаторов предложено использовать градиентный бустинг деревьев решений, реализованный в библиотеках CatBoost, XGBoost, а также метод случайного леса.

В работе [11] предлагается построение классификатора на основе полученных авторами 89 реляционных правил из моделей, основанных на правилах PART, OneR и JRip.

В статье [12] предложены методы, представляющие исследуемый файл в виде графического изображения с последующей его классификацией с использованием сети с долговременной и кратковременной памятью (LSTM).

Вышеуказанные модели бинарной классификации имеют ряд общих недостатков, которые на практике могут значительно снизить эффективность обнаружения эксплоитов:

1. Каждая из рассмотренных моделей предназначена для обнаружения эксплоитов только в конкретном формате и непригодна для анализа других файлов неисполняемых форматов.
2. При построении представленных моделей не учитываются возможности атакующего по применению технологий обфускации вредоносного кода.
3. Генерация признакового пространства в упомянутых работах осуществляется экспертом (группой экспертов) и не обосновывается математически, например, за счет математического моделирования эксплоита.
4. Разработанные модели не учитывают иные возможные места внедрения вредоносного кода кроме структурных элементов файла, предусмотренных спецификацией на формат для размещения исполняемого кода.

Вышеуказанные факторы актуализируют задачу разработки новых моделей бинарной классификации файлов неисполняемых форматов в задаче обнаружения эксплойтов, потому их разработка формирует результат, обладающий научной новизной.

В настоящей работе предлагается авторская модель бинарной классификации файлов неисполняемых форматов на основе созданного вектора информативных признаков, учитывающая указанные недостатки.

Цель (постановка задачи) исследования

Содержательная (вербальная) постановка научной задачи.

Разработать модель бинарной классификации файлов неисполняемых форматов, обеспечивающую повышение эффективности выявления обфусцированных эксплоитов, относительно моделей, реализованных в существующих средствах антивирусной защиты.

Формальная постановка научной задачи.

Дано:

X – множество файлов неисполняемых форматов, с внедренными обфусцированными эксплоитами (вредоносные файлы) и без таковых (безопасные файлы); PF – множество потенциальных признаков-кандидатов; O – множество рассматриваемых обфускаторов; SZI – множество средств антивирусной защиты; Z – множество рассматриваемых классификаторов с дискретным набором гиперпараметров.

Задача: разработать модель бинарной классификации R, обеспечивающую максимальный показатель эффективности обнаружения обфусцированных эксплоитов q, и его улучшение относительно рассматриваемых классификаторов и существующих средств антивирусной защиты:

$$R: argmax\ q[Z^*(O(X),PF^*)] > \\ > q[SZI(O(X))],|PF^*| \le |PF|, \quad (1)$$

где q – повышаемый показатель эффективности; Z^* – выбранный классификаторов с дискретным набором гиперпараметров; PF^* – множество отобранных информативных признаков.

В качестве показателя эффективности бинарной классификации определим следующий [13]:

$$q = \frac{Se + Sp}{2}, \quad (2)$$

где Se – чувствительность, под которой понимается доля истинно положительных случаев классификации, которая вычисляется по формуле:

$$Se = \frac{TP}{TP + FN}, \quad (3)$$

где TP – число верно классифицированных «вредоносных» объектов, FN – число объектов, классифицированных как отрицательные (ошибка I рода).

При этом специфичность Sp определяется как доля истинно отрицательных случаев классификации, которые были корректно идентифицированы классификатором:

$$Sp = \frac{TN}{TN + FP}, \quad (4)$$

где TN – число верно классифицированных «безопасных» объектов, FP – число объектов, классифицированных как положительные (ошибка II рода).

Структура исследования

Процедура разработки модели бинарной классификации для обнаружения обфусцированных эксплоитов

Рис. 2. Схема подготовки исходных данных

в файлах неисполняемого формата декомпозирована на следующие этапы:

1. Подготовка исходных данных;
2. Определение минимального размера выборки;
3. Формирование множества потенциальных признаков-кандидатов;
4. Формированные и отбор моделей бинарной классификации;
5. Выбор наилучшей модели бинарной классификации;

Подготовка исходных данных

Подготовка необходимых для проведения исследования наборов «вредоносных» и «безопасных» файлов осуществлялось в порядке, представленном на схеме (рис. 2).

Формат файлов Office Open XML (расширения: .docx, .xlsx, .pptx, .vsdx) был выбран ввиду его наибольшей распространенности при проведении компьютерных атак и соответственно наличия значительно большего числа образцов с внедренными эксплоитами в открытом доступе.

Рис. 3. Процедура получения эмпирического закона вероятности обнаружения обфусцированных эксплоитов

С учетом изложенного на данном этапе получены три выборки, включающие в себя на момент проведения исследования доступные образцы эксплоитов, содержащие известные разновидности уязвимостей.

Определение минимального размера выборки

Определение минимального размера выборки осуществлялась в следующем порядке:

1. Получение эмпирического закона распределения вероятности обнаружения обфусцированных эксплоитов.
2. Проверка статистической гипотезы о нормальности эмпирического закона распределения.
3. Расчет минимального размера выборки с уровнем значимости равным 0,01.

Учитывая, что на момент проведения настоящего исследования опубликовано 223 подтвержденных уязвимости, позволяющих выполнить произвольный код в файлах неисполняемого формата Office Open XML, процедура получения эмпирического закона распределения вероятности обнаружения обфусцированных эксплоитов реализована следующим образом (рис. 3).

При этом отбор из исходной выборки образцов, содержащих отличные уязвимости, осуществлялся методом бутстрэпа[3] [14].

По результатам выполнения первого этапа получено следующее распределение (рис. 4), характеризующее процесс выявления угроз нарушения информационной безопасности, представленных эксплоитами, подвергшихся процедуре обфускации.

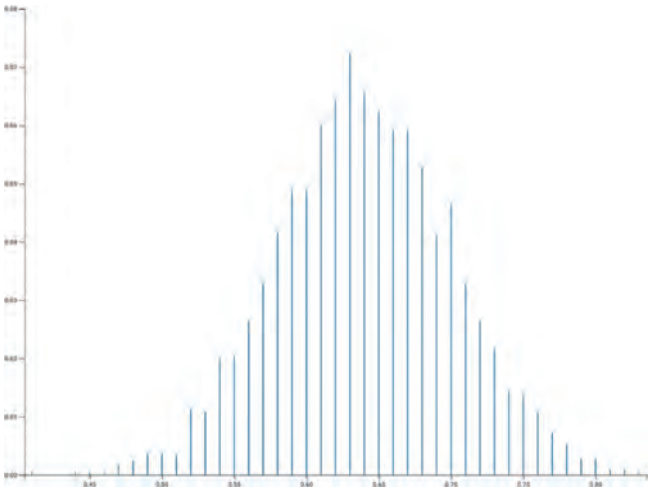

Рис. 4. Распределение, характеризующее процесс обнаружения обфусцированных эксплоитов

3 Бутстрэп (англ. bootstrap) в статистике — практический компьютерный метод исследования распределения статистик вероятностных распределений, основанный на многократной генерации выборок методом Монте-Карло на базе имеющейся выборки.

На следующем этапе для проверки гипотезы о нормальности полученного распределения использовались специализированные статистические тесты: Колмогорова–Смирнова[4] и Шапиро–Уилка[5].

По результатам выполнения указанных статистических тестов получены следующие значения контрольных критериев: D = 0.012, W = 0,99, на основании которых можно сделать вывод о том, что полученное распределение являются нормальным с уровнем значимости равным 0.05.

С учетом подтверждения гипотезы о нормальности эмпирического закона вероятности обнаружения обфусцированных эксплоитов расчет минимального размера выборки n осуществлен по следующей формуле с учетом неизвестной численности генеральной совокупности [15]:

$$n = \frac{Z^2\,\sigma^2}{\delta^2} = \frac{2^2 0.12^2}{0.01^2} = 576, \qquad (5)$$

где Z – критическое значение нормального распределения, σ – стандартное отклонение, δ – уровень значимости.

Формирование множества потенциальных признаков-кандидатов

Формирование множества потенциальных признаков-кандидатов осуществлено на базе разработанной авторами математической модели эксплоита внедренного в файл неисполняемого формата [16].

С учетом изложенного сформировано два множества потенциальных признаков-кандидатов, предназначенных для обнаружения: модуля эксплуатации уязвимости и полезной нагрузки эксплоита.

Получение численных значений признаков осуществлялась путем применения различных математических методов.

Указанные методы применялись не ко всему содержимому исследуемых файлов неисполняемого формата, а только к содержимому, входящему в состав сегментов, полученных по результатам авторского алгоритма сегментации [17].

В качестве математических преобразований применялись методы, используемые в математической статистике и теории информации, а также ряд авторских процедур, которые приведены в таблицах (табл. 1, табл. 2).

Формирование и отбор моделей бинарной классификации

Формирование и отбор моделей бинарной классификации осуществлялся по следующей схеме (рис. 5) на основе обучающей выборки и показателя эффективности бинарной классификации q (2).

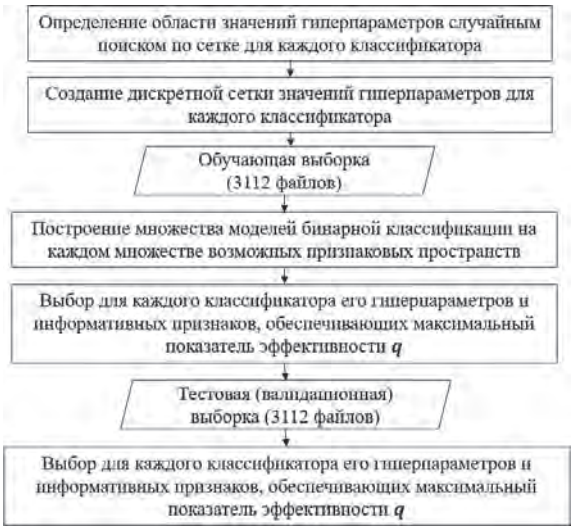

Рис. 5. Схема формирования и отбора моделей бинарной классификации

Таблица 1.
Множество признаков-кандидатов для обнаружения модуля эксплуатации уязвимости эксплоита

Методы получения числовых значений признака	Принадлежность
Форматный	Авторский
Математическое ожидание	Мат. статистика
Дисперсия	Мат. статистика
Мода	Мат. статистика
Медиана	Мат. статистика
Среднее линейное отклонение	Мат. статистика
Среднеквадратичное отклонение	Мат. статистика
Коэффициент асимметрии	Мат. статистика
Коэффициент эксцесса	Мат. статистика
Максимальное значение	Мат. статистика
Минимальное значение	Мат. статистика
Размах вариации	Мат. статистика
Базисный абсолютный прирост	Мат. статистика
Цепной абсолютный прирост	Мат. статистика
Коэффициент осцилляции	Мат. статистика
Относительное линейное отклонение	Мат. статистика
Относительный показатель квартильной вариации	Мат. статистика
Коэффициент вариации	Мат. статистика
Эмпирический коэффициент детерминации	Мат. статистика
Децильный коэффициент дифференциации	Мат. статистика
Тест Колмогорова-Смирнова	Мат. статистика
Тест Шапиро-Уилка	Мат. статистика
Энтропия	Теория информации

4 Леман Э. Проверка статистических гипотез; [пер. с англ.]. М.: Наука, 1979. 408 с.

5 Shapiro S. S., Wilk M. B. An analysis of variance test for normality (complete samples). Biometrika, 1965, vol. 52, no. 3-4, pp. 591–611. DOI: 10.1093/BIOMET/52.3-4.591.

Таблица 2.

Множество признаков-кандидатов для обнаружения полезной нагрузки эксплоита

Методы получения числовых значений признака	Принадлежность
Форматный	Авторский
Признак наличия машинного кода	Авторский
Признак наличия программного кода интерпретируемых языков программирования	Авторский
Признак наличия программного кода встроенных средств командного процессора	Авторский
Признак наличия программного кода встроенных средств программирования	Авторский
Показатель числа управляющих символов в строке	Авторский
Показатель числа специальных символов в строке	Авторский
Математическое ожидание	Мат. статистика
Дисперсия	Мат. статистика
Среднее линейное отклонение	Мат. статистика
Среднеквадратичное отклонение	Мат. статистика
Максимальное значение	Мат. статистика
Минимальное значение	Мат. статистика
Размах вариации	Мат. статистика
Коэффициент вариации	Мат. статистика
Энтропия	Теория информации
2-граммная энтропия	Теория информации
3-граммная энтропия	Теория информации
Эмпирический коэффициент детерминации	Мат. статистика
Децильный коэффициент дифференциации	Мат. статистика
Тест Колмогорова-Смирнова	Мат. статистика
Тест Шапиро-Уилка	Мат. статистика
Энтропия	Теория информации

По результатам выполнения вышеуказанных процедур отобраны сформированные модели бинарной классификации (табл.3), имеющие наилучшие значения показателя эффективности (2) для каждого из классификаторов, с учетом его гиперпараметров и признакового пространства.

Таблица 3.

Отобранные модели бинарной классификации

Метод классификации	Число признаков	Значения критерия q
Алгоритмы на основе деревьев решений		
Decision Tree	31	0.77
Extra Tree	29	0.73
Ансамблевые алгоритмы		
AdaBoost	22	0.90
Extra Trees	28	0.89
Random Forest	28	0.87
Bagging	23	0.85
Gradient Boosting	21	0.95
Hist Gradient Boosting	21	0.94
CatBoost	21	0.96
LightGBM	23	0.95
XGBoost	21	0.99
Линейные классификаторы		
Gaussian Naive Bayes	32	0.66
Multinomial Naive Bayes	29	0.65
Complement Naive Bayes	30	0.63
Bernoulli Naive Bayes	27	0.59
Ridge Classification	29	0.71
Stochastic Gradient Descent	26	0.73
Logistic Regression	28	0.69
Passive Aggressive Classifier	31	0.59
Метрические классификаторы		
K-nearest neighbors	25	0.88
Nearest Centroid	25	0.83
Метод опорных векторов		
SVM	28	0.84
Искусственные нейронные сети		
Perceptron	24	0.87
Multi-layer Perceptron	28	0.89

Выбор наилучшей модели бинарной классификации

Выбор наилучшей модели бинарной классификации осуществлялся по следующей схеме (рис. 6) на основе валидационной выборки и показателя эффективности бинарной классификации q (2).

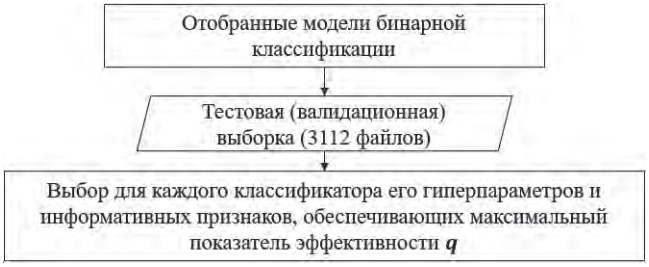

Рис. 6. Схема выбора наилучшей модели бинарной классификации

По результатам выполнения вышеуказанных процедур получены следующие значения показателя эффективности бинарной классификации q (2) для каждой из сформированных моделей бинарной классификации (табл. 4).

Таблица 4.
Множество признаков-кандидатов для обнаружения модуля эксплуатации уязвимости эксплоита

Метод классификации	Значения критерия q
Decision Tree	0.73
Extra Tree	0.72
AdaBoost	0.87
Extra Trees	0.86
Random Forest	0.83
Bagging	0.76
Gradient Boosting	0.91
Hist Gradient Boosting	0.92
CatBoost	0.93
LightGBM	0.93
XGBoost	0.98
Gaussian Naive Bayes	0.65
Multinomial Naive Bayes	0.65
Complement Naive Bayes	0.58
Bernoulli Naive Bayes	0.57
Ridge Classification	0.69
Stochastic Gradient Descent	0.72
Logistic Regression	0.67
Passive Aggressive Classifier	0.58
K-nearest neighbors	0.84
Nearest Centroid	0.80
SVM	0.83
Perceptron	0.83
Multi-layer Perceptron	0.87

С учетом полученных результатов в качестве наилучшей модели бинарной классификации выбрана модель, построенная на базе классификатора XGBoost[6] и двух конечных множеств информативных признаков $P'_M = \{F'_1,F'_2,\ldots,F'_{11}\}$ и $P'_S = \{G'_1,G'_2,\ldots,G'_{10}\}$, где:

1. P'_M – признаковое пространство для обнаружения модуля эксплуатации уязвимости эксплоита, P'_S – признаковое пространство для обнаружения модуля полезной нагрузки эксплоита.
2. F'_1 – математические ожидание, которое вычисляется по формуле:
$$F'_1 = \overline{x} = \sum_{i=1}^{256} b_i\, p_i, \quad (6)$$
где b_i – значение байта, p_i – вероятность появления b_i.
3. F'_2 – дисперсия, которая вычисляется по формуле:
$$F'_2 = D = \sum_{i=1}^{256} (b_i - \overline{x})\, p_i, \quad (7)$$
4. F'_3 – среднеквадратичное отклонение, которое вычисляется по формуле:
$$F'_3 = \sigma = \sqrt{D} \quad (8)$$
5. F'_4 – коэффициент асимметрии, который вычисляется по формуле:
$$F'_4 = A_3 = \frac{m_3}{\sigma_3}, \quad (9)$$
где m_3 – центральный эмпирический момент третьего порядка.
6. F'_5 – коэффициент эксцесса, который вычисляется по формуле:
$$F'_5 = E = \frac{m_4}{\sigma_4}, \quad (10)$$
где m_4 – центральный эмпирический момент четвертого порядка.
7. F'_6 – максимальное значение, которое вычисляется по формуле:
$$F'_6 = max(p_i) \quad (11)$$
8. F'_7 – минимальное значение, которое вычисляется по формуле:
$$F'_7 = min(p_i) \quad (12)$$
9. F'_8 – размах вариации, который вычисляется по формуле:
$$F'_8 = max(p_i) - min(p_i) \quad (13)$$
10. F'_9 – коэффициент вариации, который вычисляется по формуле:
$$F'_9 = \frac{\sigma * 100}{\overline{x}} \quad (14)$$
11. F'_{10} – энтропия, которая вычисляется по формуле:
$$F'_{10} = -\sum_{i=1}^{32640} p_i \log_2 p_i \quad (15)$$

6 XGBoost. – URL: https://github.com/dmlc/xgboost/ (дата обращения: 01.07.2024).

12. F'_{11} – форматный признак, который вычисляется по формуле:

$$F'_{11} = \begin{cases} 1, \text{ если все } B \text{ соотвествуют формату} \\ 0, \text{ если хотя бы один } B \text{ не соотвествует формату} \end{cases}, \quad (16)$$

где B – сегмент, полученный по результатам алгоритма сегментации, для обнаружения модуля эксплуатации уязвимости эксплоита.

13. G'_1 – форматный признак, который вычисляется по формуле:

$$G'_1 = \begin{cases} 1, \text{ если все } B' \text{ соотвествуют формату} \\ 0, \text{ если хотя бы один } B' \text{ не соотвествует формату} \end{cases}, \quad (17)$$

где B' – сегмент, полученный по результатам алгоритма сегментации для обнаружения модуля полезной нагрузки эксплоита.

14. G'_2 – признак наличия машинного кода, G'_3 – признак наличия программного кода интерпретируемых языков программирования, G'_4 – признак наличия кода встроенных средств командного процессора, G'_5 – признак наличия кода встроенных средств программирования, которые вычисляется по общей формуле:

$$G'_i(B') = \frac{G''_i}{|B'|}, \begin{cases} G''_i = G'_i + 1, \text{ если } d_{G'_i} \in B' \\ G''_i = G'_i + 0, \text{ если } d_{G'_i} \notin B' \end{cases}, \quad (18)$$

где $d_{G'_i}$ – команда (инструкция) множества программного кода, $i = 2...5$, $|B'|$ – размер сегмента в байтах, G''_i – общее число найденных $d_{G'_i}$ в сегменте B'.

15. G'_6 – показатель числа управляющих символов в строке, который вычисляется по формуле:

$$G'_6 = \frac{N_s}{|B'|}, \quad (19)$$

где N_s – количество специальных символов в сегменте B'.

16. G'_7 – показатель числа специальных символов в строке, который вычисляется по формуле:

$$G'_7 = \frac{N'_s}{|B'|}, \quad (20)$$

где N'_s – количество управляющих символов в сегменте B'.

17. G'_8 – энтропия, которая вычисляется по формуле:

$$G'_8 = -\sum_{i=1}^{256} p_i \log_2 p_i \quad (21)$$

18. G'_9 – 2-граммная энтропия, которая вычисляется по формуле:

$$G'_9 = -\sum_{j=1}^{32640} p_j \log_2 p_j \quad (22)$$

19. G'_{10} – 3-граммная энтропия, которая вычисляется по формуле:

$$G'_9 = -\sum_{k=1}^{2763520} p_k \log_2 p_k \quad (23)$$

Оценка эффективности обнаружения обфусцированных эксплоитов за счет применения разработанной модели

В целях оценки эффективности обнаружения обфусцированных эксплоитов за счет применения предложенной модели бинарной классификации проведены натурные экспериментальные исследования.

Предложенная модель реализована в программе для ЭВМ AntigenExploits (Свидетельство о госрегистрации № 2023687464).

Исследования эффективности проведены с использованием указанной программы.

Экспериментальные исследования проводились в следующем порядке:

1. Подготовлена тестовая выборка из 3200 файлов неисполняемых форматов, не используемых при построении модели бинарной классификации, содержащей файлы с внедренными эксплоитами (вредоносные файлы – 1600 образцов), и без таковых (безопасные файлы – 1600 образцов), сгенерированных в автоматическом режиме с использованием штатных шаблонов приложений Microsoft Office.

Для подтверждения наличия/отсутствия вредоносного кода в сформированной выборке применялись средства (технологии) антивирусной защиты информации, размещенные в открытом доступе [https://virustotal.com/].

Вредоносные файлы, входящие в выборку, прошли процедуру обфускации с использованием свободно доступных программных средств [https://github.com/], реализующих обфускацию программных кодов в автоматизированном режиме.

Таблица 5.

Результаты экспериментального исследования

Средства антивирусной защиты	Значения критерия q
AntigenExploits (СГР № 2023687464)	**0,99**
Kaspersky Endpoint Security	0,88
Dr.Web Enterprise Security Suite	0,74
NANO Антивирус Pro	0,67
Антивирус «VR Protect» для Linux	0,69
Avast	0,72
ClamAV	0,54
AVG	0,63
Symantec	0,84
Microsoft	0,69
McAfee Scanner	0,85
Panda	0,75

2. Проведен анализ тестовой выборки с использованием средств антивирусной защиты, включенных в реестр российского программного обеспечения, и средств антивирусной защиты иностранного производства, размещенных в свободном доступе, и авторской программы.
3. Проведен анализ результатов экспериментального исследования.

Результаты исследования представлены в таблице (таб. 5).

Результаты экспериментального исследования показали, что предложенная модель бинарной классификации в конкретном исследовании позволила повысить эффективность обнаружения обфусцированных эксплоитов относительно существующих средств антивирусной защиты в среднем на 26 %.

Заключение

В статье представлена модель бинарной классификации файлов неисполняемых форматов, применимая в задаче обнаружения эксплоитов, определяющая используемый классификатор, его гиперпараметры, а также множество информативных признаков.

Модель является универсальной и позволяет производить классификацию файлов неисполняемых форматов на «вредоносные» и «безопасные», обеспечивающую высокие значения выбранного показателя эффективности обнаружения относительно существующих средств антивирусной защиты. При этом предложенная модель бинарной классификации файлов неисполняемых форматов ориентирована, в первую очередь, на обнаружение эксплоитов, подвергшихся процедуре обфускации.

Указанные аспекты отличают представленную модель от релевантных аналогов.

В качестве направления для дальнейших исследований в данной области целесообразно рассмотреть возможность решения задачи оптимизации подбора гиперепараметров для классификатора, используемого в предложенной модели.

Литература

1. *Середкин С. П. Особенности кибератак на объекты критической информационной инфраструктуры в современных условиях // Информационные технологии и математическое моделирование в управлении сложными системами. – 2022. – № 4 (16). – С. 56–66. DOI: 10.26731/2658-3704.2022.4(16).56-66.*
2. *Ланецкая А. Ю., Александрова Е. Н. Современные угрозы информационной безопасности // Международный журнал гуманитарных и естественных наук. – 2022. – Том 7–2. – № 20. – С. 192–195. DOI:10.24412/2500-1000-2022-7-2-192-195.*
3. *Павлычев А. В., Стародубов М. И., Галимов А. Д. Использование алгоритма машинного обучения Random Forest для выявления сложных компьютерных инцидентов // Вопросы кибербезопасности. – 2022. – Том 51. – № 5. – С. 74–81. DOI:10.21681/2311-3456-2022-5-74-81.*
4. *Таловерова Д. В. Сравнительный анализ сценариев реализации угроз безопасности информации методики ФСТЭК РФ и Mitre Att&ck и их применение на практике // Фундаментальные и прикладные аспекты компьютерных технологий и информационной безопасности. Сборник статей Всероссийской научно-технической конференции. Таганрог, 2023. – С. 34–37.*
5. *Архипов А. Н., В. А. Пиков, В. В. Кабаков Порядок и результаты экспериментальных исследований влияния обфускации на качество выявления угроз информационной безопасности, реализуемых посредством эксплоитов, в файлах неисполняемых форматов // Научно-практический журнал «Вопросы защиты информации» (Доверенная среда). – 2023. – С. 32–37.*
6. *Kamran Saeed, M. Fatih Adak Detection of Unknown Malicious Microsoft Office Documents based on Hidden Feature Extraction by using Machine Learning // Authorea. – 2024.– P. 1–16. DOI: 10.22541/au.170664344.41804021/v1.*
7. *Salman Abdul Jabbaar Wiharja, Deden Pradeka Wirmanto Suteddy, Designing A Pdf Malware Detection System Using Machine Learning // Jurnal Poli-Teknologi. – 2024. – Vol. 23, No. 1. – P. 40–54. DOI:10.32722/pt.v23i1.6540.*
8. *Fran Casino, Nikolaos Totosis, Theodoros Apostolopoulos, Nikolaos Lykousas Analysis and Correlation of Visual Evidence in Campaigns of Malicious Office Documents // Digital Threats Research and Practice. – 2022. – Vol. 4, No. 2. – P. 1–19. DOI:10.1145/3513025.*
9. *Candra Ahmadi, Jiann Chen, Yi-Cheng Lai Enhancing Detection of Malicious VBA Macros in Office Documents: An Integrated Approach Employing P-Code Analysis and XGBoost-based Machine Learning Model // IEEE Access. – 2024. – Vol. 12. – P. 71746–71760. DOI: 10.1109/ACCESS.2024.3402956.*
10. *V Ravi, S. P. Gururaj, H. K. Vedamurthy, M. B. Nirmala Analysing corpus of office documents for macro-based attacks using Machine Learning // Siddaganga Institute of Technology. – 2022. – Vol. 8, No. 3. – P. 20–24. DOI:10.1016/j.gltp.2022.04.004.*
11. *Geet C. Salame, Nirlepa T. Shinde, Prajakta P. Baad, Deepak D. Kshirsagar A. relational rule-based system for PDF malware detection // Journal of Information and Optimization Sciences. – 2024. – Vol. 45, No. 4. – P. 925–934. DOI:10.47974/JIOS-1616.*
12. *Maheshwaran T., Manideep M., Sai Chaitanya K., Karthik A. Securing pdfs: an innovative lstm algorithm for image-based malware detection // Interantional journal of scientific research in Engineering and Management. – 2024. – Vol. 8, No. 5. – P. 1–5. DOI:10.55041/IJSREM34090.*
13. *Старовойтов В. В., Голуб Ю. И. Сравнительный анализ оценок качества бинарной классификации // Информатика. – 2020. – Т. 17, № 1. – С. 87–101. DOI:10.37661/1816-0301-2020-17-1-87-101.*
14. *Bradley Efron. Bootstrap Methods: Another Look at the Jackknife // Annals of Statistics. – 1979. – Vol. 7, no. 1. – P. 1–26.*
15. *Donna L. M., William J. W., Rudolf J. F. Statistical Methods // University of North Florida. – 2021. – Vol. 4. – P. 123-167. DOI:10.1016/C2019-0-02521-6.*
16. *Кондаков С. Е., Архипов А. Н. Математическая модель эксплоита, внедренного в файл неисполняемого формата // Изв. ИИФ. 2023. Т. 69. № 3. С. 93–96.*
17. *Архипов А. Н., Кондаков С. Е. Сегментация файлов неисполняемых форматов для выявления угроз нарушения информационной безопасности, реализуемых в форме эксплоитов // Программные продукты и системы. 2024. Т. 37. № 2. С. 186–192. DOI: 10.15827/0236-235X.142.186-192.*

IDETECTING OBFUSCATED EXPLOITS IN NON-EXECUTABLE FORMAT FILES

Arkhipov A. N.[7], Kondakov S. E.[8]

The purpose of the research: *is the development of a model of binary classification of non-executable file formats, which provides increased efficiency of detection of obfuscated exploits, relative to the models implemented in existing anti-virus protection tools.*

Research methods *are based on the provisions of probability theory and mathematical statistics, set theory, methods of conducting field experiments and processing experimental data.*

Result: *in the course of the research, on the basis of the mathematical model of the exploit, a set of potential features, which are represented by numerical values, was generated. Informative features were selected from the generated feature space and a binary classification model with the best performance in detecting obfuscated exploits was built. A computer program implementing the obtained model was developed. The effectiveness of the developed model is confirmed in the framework of experimental studies to assess the effectiveness of detecting obfuscated exploits using anti-virus protection tools included in the register of Russian software, and foreign anti-virus protection tools placed in free access, and the author's program.*

The scientific novelty *of the results is determined by a set of author's procedures providing the choice of classifier, its hyperparameters, as well as the formation of an informative feature space, including features developed by the authors, and, allowing to build the most effective model of binary classification, which ensures the validity of the obtained results. The author presents the confirmation of realizability and obtaining the best values of efficiency indicators in detecting obfuscated exploits in relation to the existing means of antivirus protection.*

Practical significance: *the presented model, first of all, is oriented on application in antivirus protection systems, but it can be used for solving other tasks of information security.*

Keywords: *cybersecurity, computer attacks, local exploits, malicious code, information protection, anti-virus information protection systems, intrusion detection system.*

References

1. *Seredkin S.P. Osobennosti kiberatak na ob"ekty' kriticheskoj informacionnoj infrastruktury' v sovremenny'x usloviyax // Informacionny'e texnologii i matematicheskoe modelirovanie v upravlenii slozhny'mi sistemami. – 2022. – № 4 (16). – S. 56–66. DOI: 10.26731/2658-3704.2022.4(16).56-66.*
2. *Laneczkaya A. Yu., Aleksandrova E.N. Sovremenny'e ugrozy' informacionnoj bezopasnosti // Mezhdunarodny'j zhurnal gumanitarny'x i estestvenny'x nauk. – 2022. – Tom 7-2. – № 20. – S. 192–195. DOI:10.24412/2500-1000-2022-7-2-192-195.*
3. *Pavly'chev A. V., Starodubov M. I., Galimov A. D. Ispol'zovanie algoritma mashinnogo obucheniya Random Forest dlya vy'yavleniya slozhny'x komp'yuterny'x incidentov // Voprosy' kiberbezopasnosti. – 2022. – Tom 51. – № 5. – S. 74–81. DOI:10.21681/2311-3456-2022-5-74-81.*
4. *Taloverova D. V. Sravnitel'ny'j analiz scenariev realizacii ugroz bezopasnosti informacii metodiki FSTE'K RF i Mitre Att&ck i ix primenenie na praktike // Fundamental'ny'e i prikladny'e aspekty' komp'yuterny'x texnologij i informacionnoj bezopasnosti. Sbornik statej Vserossijskoj nauchno-texnicheskoj konferencii. Taganrog, 2023. – S. 34–37.*
5. *Arxipov A. N., V. A. Pikov, V. V. Kabakov Poryadok i rezul'taty' e'ksperimental'ny'x issledovanij vliyaniya obfuskacii na kachestvo vy'yavleniya ugroz informacionnoj bezopasnosti, realizuemy'x posredstvom e'ksploitov, v fajlax neispolnyaemy'x formatov // Nauchno-prakticheskij zhurnal «Voprosy' zashhity' informacii» (Doverennaya sreda). – 2023. – S. 32-37.*
6. *Kamran Saeed, M. Fatih Adak Detection of Unknown Malicious Microsoft Office Documents based on Hidden Feature Extraction by using Machine Learning // Authorea. – 2024.– P. 1–16. DOI: 10.22541/au.170664344.41804021/v1.*
7. *Salman Abdul Jabbaar Wiharja, Deden Pradeka Wirmanto Suteddy, Designing A Pdf Malware Detection System Using Machine Learning // Jurnal Poli-Teknologi. – 2024. – Vol. 23, No. 1. – P. 40-54. DOI:10.32722/pt.v23i1.6540.*
8. *Fran Casino, Nikolaos Totosis, Theodoros Apostolopoulos, Nikolaos Lykousas Analysis and Correlation of Visual Evidence in Campaigns of Malicious Office Documents // Digital Threats Research and Practice. – 2022. – Vol. 4, No. 2. – P. 1-19. DOI:10.1145/3513025.*
9. *Candra Ahmadi, Jiann Chen, Yi-Cheng Lai Enhancing Detection of Malicious VBA Macros in Office Documents: An Integrated Approach Employing P-Code Analysis and XGBoost-based Machine Learning Model // IEEE Access. – 2024. – Vol. 12. – P. 71746–71760. DOI: 10.1109/ACCESS.2024.3402956.*
10. *V Ravi, S. P. Gururaj, H. K. Vedamurthy, M. B. Nirmala. Analysing corpus of office documents for macro-based attacks using Machine Learning // Siddaganga Institute of Technology. – 2022. – Vol. 8, No. 3. – P. 20–24. DOI:10.1016/j.gltp.2022.04.004.*
11. *Geet C. Salame, Nirlepa T. Shinde, Prajakta P. Baad, Deepak D. Kshirsagar A. relational rule-based system for PDF malware detection // Journal of Information and Optimization Sciences. – 2024. – Vol. 45, No. 4. – P. 925–934. DOI:10.47974/JIOS-1616.*

7 Alexander N. Arkhipov, student Bauman Moscow State Technical University, Moscow, Russia. E-mail: diskpart111@mail.ru
8 Sergey E. Kondakov, Ph.D. (tech.), associate Professor, Bauman Moscow State Technical University, Moscow, Russia. E-mail: sergeikondakov@list.ru

12. *Maheshwaran T., Manideep M., Sai Chaitanya K., Karthik A. Securing pdfs: an innovative lstm algorithm for image-based malware detection // Interantional journal of scientific research in Engineering and Management. – 2024. – Vol. 8, No. 5. – P. 1–5. DOI:10.55041/IJSREM34090.*
13. *Starovojtov V. V., Golub Yu. I. Sravnitel'ny'j analiz ocenok kachestva binarnoj klassifikacii // Informatika. – 2020. – T. 17, № 1. – S. 87–101. DOI:10.37661/1816-0301-2020-17-1-87-101.*
14. *Bradley Efron. Bootstrap Methods: Another Look at the Jackknife // Annals of Statistics. – 1979. – Vol. 7, no. 1. – P. 1–26.*
15. *Donna L. M., William J. W., Rudolf J. F. Statistical Methods // University of North Florida. – 2021. – Vol. 4. – P. 123–167. DOI:10.1016/C2019-0-02521-6.*
16. *Kondakov S. E., Arxipov A. N. Matematicheskaya model' e'ksploita, vnedrennogo v fajl neispolnyaemogo formata // Izv. IIF. 2023. T. 69. № 3. S. 93–96.*
17. *Arxipov A. N., Kondakov S. E. Segmentaciya fajlov neispolnyaemy'x formatov dlya vy'yavleniya ugroz narusheniya informacionnoj bezopasnosti, realizuemy'x v forme e'ksploitov // Programmny'e produkty' i sistemy'. 2024. T. 37. № 2. S. 186–192. DOI: 10.15827/0236-235X.142.186-192.*