DOI 10.37882/2223–2966.2022.03.17

МЕТОД ОБНАРУЖЕНИЯ СЕТЕВОЙ СТЕГАНОГРАФИИ НА ОСНОВЕ МАШИННОГО ОБУЧЕНИЯ[1]

A METHOD FOR DETECTING NETWORK STEGANOGRAPHY BASED ON MACHINE LEARNING

A. Krasov

Summary. Classification of steganographic methods of information transformation is described in the basic model of threats to the security of personal data during their processing in personal data information systems, approved by the FSTEC of the Russian Federation in 2008. Due to the growing need for confidentiality in data transmission over modern communication channels, the popularity of various methods of steganography has increased. However, the development of methods of covert transmission of information has led to the need to identify participants in the illegal dissemination of information. This article discusses the possibility of analyzing network traffic using machine learning technologies. The article presents the results of work on the Grant-I B5/2020 project, proposals for improving the basic model.

Keywords: machine learning, steganography, TCP, ICMP, Bayesian networks.

Красов Андрей Владимирович
К.т.н., доцент, Санкт-Петербургский государственный университет телекоммуникаций им. проф. М.А. Бонч-Бруевича
krasov@inbox.ru

Аннотация. Классификация стеганографических методов преобразования информации описано в базовой модели угроз безопасности персональных данных при их обработке в информационных системах персональных данных, утверждённой ФСТЭК РФ в 2008 году. В связи с ростом потребности в конфиденциальности при передаче данных по современным каналам связи возросла популярность различных методов стеганографии. Однако развитие методов скрытой передачи информации привело к необходимости выявления участников нелегального распространения информации. В данной статье рассматривается возможность анализа сетевого трафика с применением технологий машинного обучения. В статье приводятся результаты работы по проекту Грант-ИБ 5/2020, предложения по совершенствованию базовой модели.

Ключевые слова: машинное обучение, стеганография, TCP, ICMP, байесовские сети.

Сетевая стеганография — технология позволяющая скрытно осуществлять передачу информации по общедоступным каналам связи, причём данные скрыты не только от человека, но и от промежуточных устройств. Применение данного метода привлекает не только законопослушных граждан, желающих добиться конфиденциальности в интернете, но также злоумышленников, пытающихся скрыть свои намерения от правоохранительных органов, в связи с чем растёт потребность в обнаружении фактов использования сетевой стеганографии. В [1–2] описывается процесс передачи данных по каналам связи и возможность использования механизмов стеганографии.

Сокрытие информации может достигаться за счёт инъекции полезной нагрузки в поля заголовков пакетов, модификации основных данных, передающихся по каналу связи или изменения структуры очерёдности пересылки пакетов. Также могут использоваться различные комбинации данных методов, например намеренная потеря аудио пакетов с целью последующего изменения и ретрансляции.

Существующие средства детектирования стеганографии различаются в зависимости от входных данных. Большинство технологий основаны на использовании статистических данных об имеющемся канале или сигнатур.

Метод с использованием сигнатур сопоставляет пакет с уже имеющимися наборами правил и определяет является ли он стеганоконтейнером. Несовершенство данного метода заключается в невозможности распознавания неизвестных технологий сокрытия данных.

Обнаружение с использованием статистических данных заключается в обнаружении стеганографии на основании отклонений от имеющихся параметров

[1] Исследование выполнено при финансовой поддержке Минобрнауки России (грант ИБ) в рамках научного проекта № 5/2020.

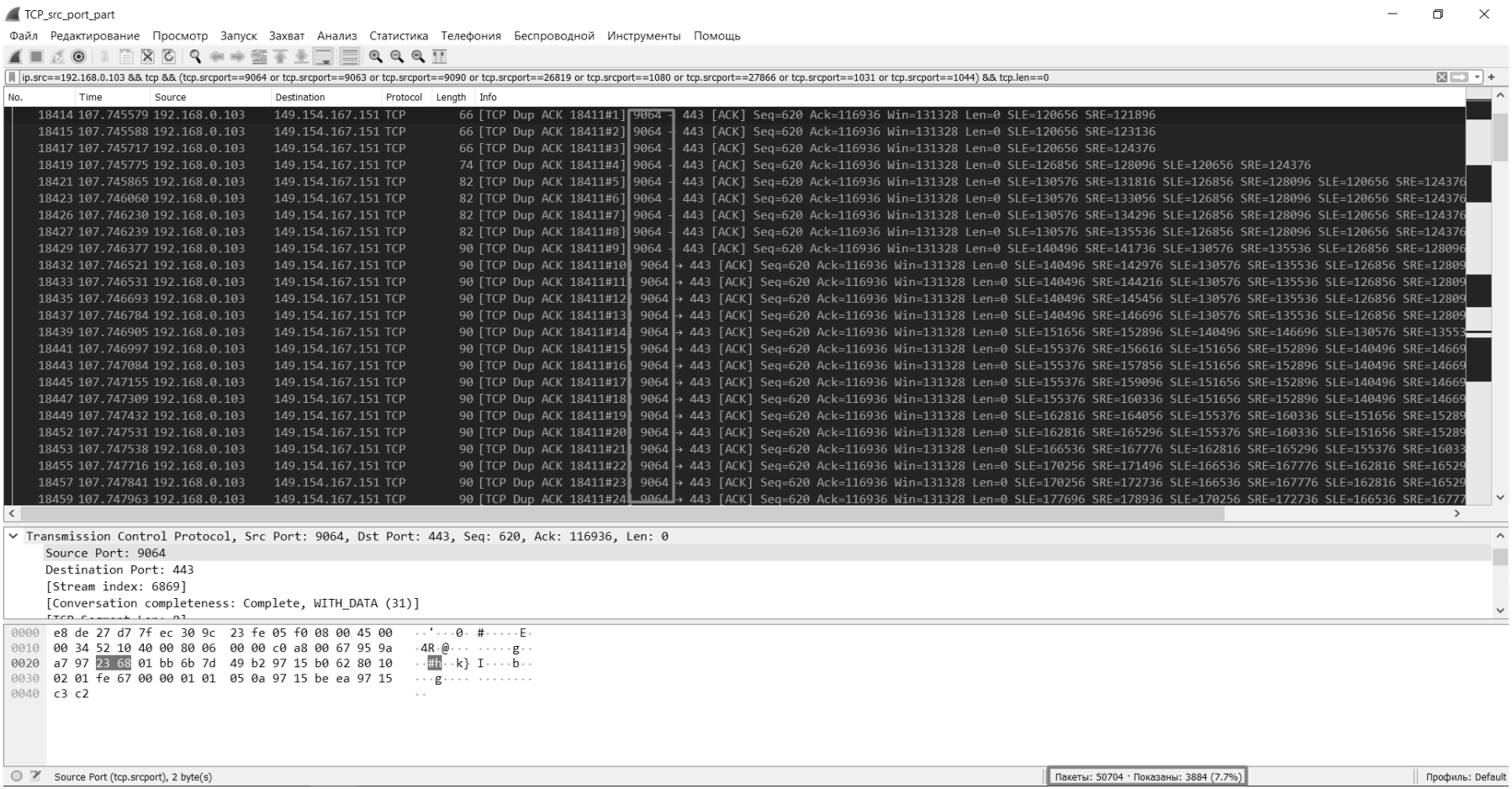

Рис. 1. Дамп пакетов с вложениями в номер порта источника

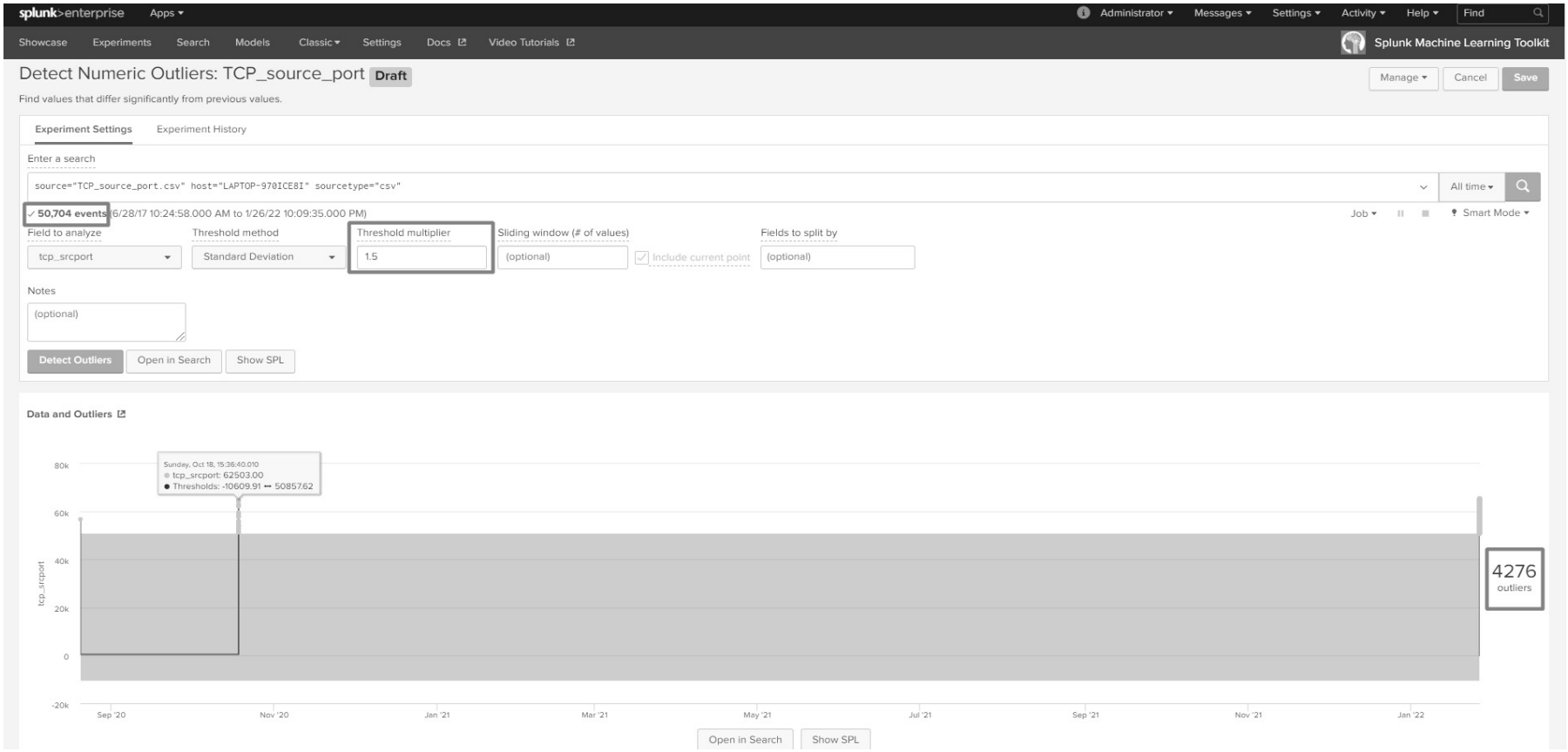

Рис. 2. Результат работы модели Detect Numeric Outliners

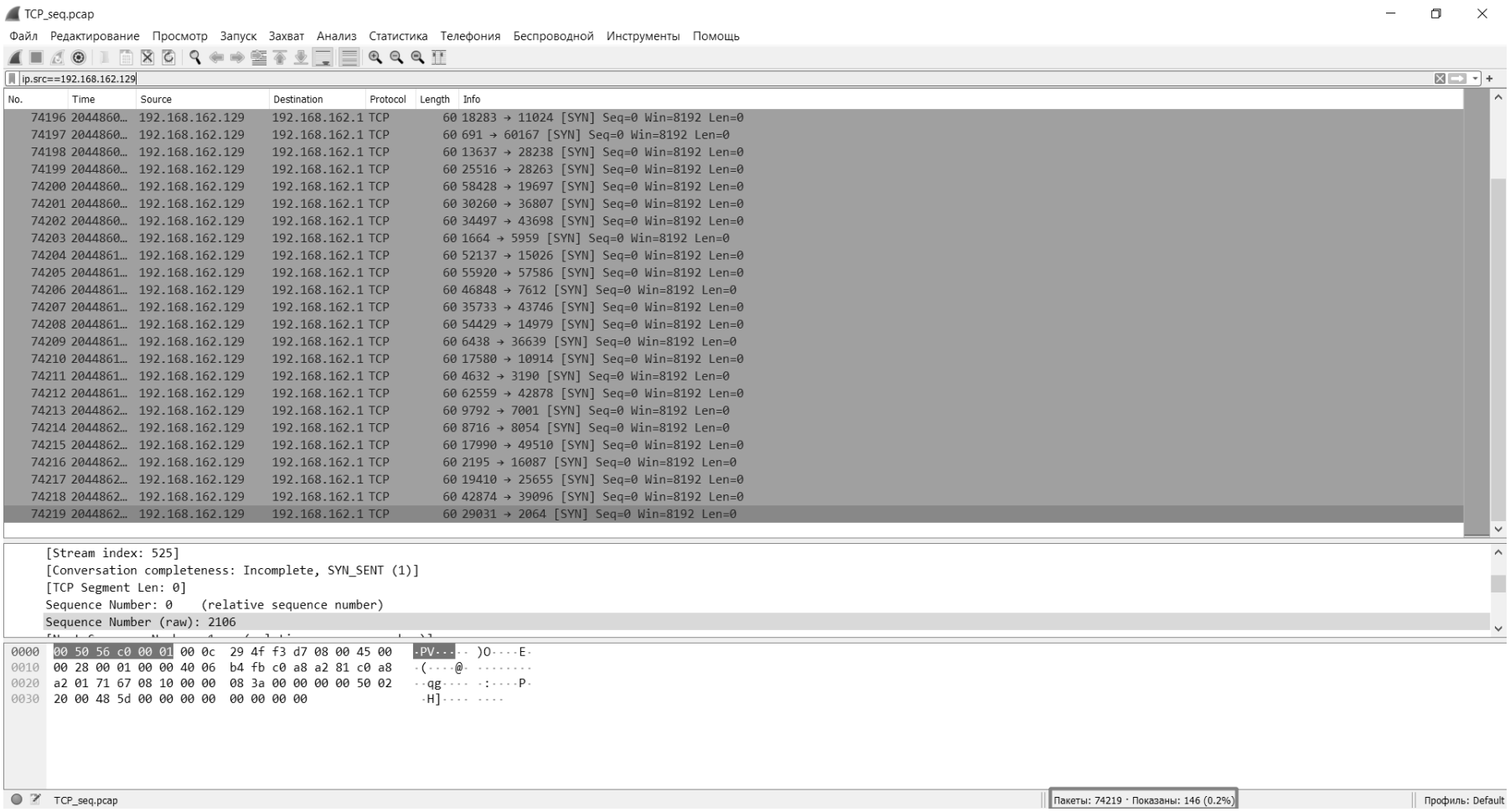

Рис. 3. Дамп с вложениями в поле identification заголовка ICMP пакетов

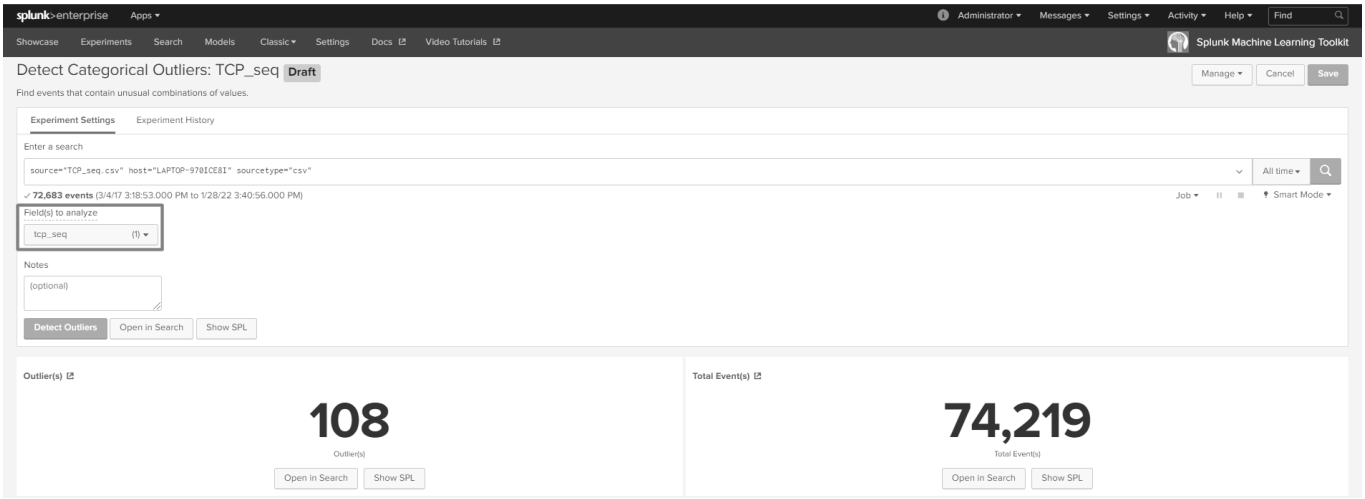

Рис. 4. Результат работы модели Detect Categorial Outliners

канала связи. Данный подход имеет относительно невысокую точность обнаружения модифицированных пакетов. Более детально о реализации данных методов стеганографии можно узнать в [3–6].

Выявление фактов сокрытия данных с применением машинного обучения — это использование существующих метрик и данных совокупно с алгоритмами классификации. Основной недостаток заключается в увеличении требуемой вычислительной мощности с ростом анализируемого трафика, однако эффективность и точность обнаружения в сравнении с остальными методами на порядок выше.

Для реализации метода обнаружения стеганографии с использованием машинного обучения необходимо определиться с данными для анализа. Трафик можно анализировать в режиме реального времени, однако для этого не подойдут обычные рабочие станции — необходимы вычислительные кластеры с высокой производительностью, также имеется возможность производить запись пакетов с помощью программ, таких как Wireshark или Tcpdump и далее производить анализ. В качестве платформы для осуществления механизмов машинного обучения можно создать собственное решение с использованием библиотек, описывающих основные алгоритмы классификации и регрессии, таких как Scikit-learn. Такой подход подойдёт в случае необходимости гибкой настройки, однако он сложен в практической реализации и последующей эксплуатации. Гораздо проще использовать готовые продукты.

Для проведения эксперимента была выбрана платформа Splunk Enterprise с использованием подключаемого модуля Machine learning toolkit. Splunk представляет собой систему сбора и анализа машинных данных IT-инфраструктуры. В рамках возможностей данной системы возможен поиск по данным в режиме реального времени или по архивным данным. Помимо поиска реализованы функции визуализации путём построения графиков-гистограмм, а также доступно формирование автоматизированных отчетов. Модуль Machine learning toolkit позволяет расширить возможности анализа с помощью наиболее популярных алгоритмов машинного обучения таких как Random forest, Kernel ridge, Gaussian NB, Gradient Boosting и других. Для работы с системой предоставлен веб-интерфейс позволяющий не обращаться к использованию скриптового языка программирования, однако в случае необходимости более гибкой настройки имеется возможность ручного ввода команд.

В целях проведения анализа на сетевом интерфейсе персонального компьютера были записаны дампы трафика с вложениями в поля заголовков TCP пакетов, произведённых с помощью Python-библиотеки Scapy.

1. Вложения в поле Source Port протокола TCP

Поле Source Port представляет собой 16 бит числовых данных. Вложения не определяются стандартными средствами, однако могут быть легко выявлены человеком при ручном анализе. В наборе данных имеется 50704 пакета (рис. 1), из которых 3884 являются модифицированными.

Модуль машинного обучения имеет десять стандартных моделей, три из них предназначены непосредственно для определения отклоняющихся значений. Так как номер порта источника представляет собой числовое значение, наиболее подходящей моделью будет Detect Numeric Outliners.

При стандартных параметрах алгоритм выдал результат в 4276 отклонений, при истинном значении 3884 (рис. 2). Для повышения эффективности анализа можно изменять параметр threshold method, который задаёт тип отклонения и мультипликатор огибающей функции.

2. Вложения в поле seq заголовков TCP пакетов

Поле Sequence Number состоит из 32 бит, являющих собой порядковый номер пакета. Вложения в данное поле никак не определяются системно, однако могут привести к сбою в программе и выявлению если значение будет меньше первого sn

Для анализа можно воспользоваться моделью Detect Categorial Outliners. Данная модель основана на алгоритме вероятностных мер. Данный дамп представляет собой 74219 пакетов (рис. 3), из которых 146 с вложениями.

В результате работы алгоритм распознал 108 из 146 модифицированных пакетов (рис. 4). Данный результат достигается благодаря возможности анализа сразу нескольких полей. Также стоит учесть низкую распространённость ICMP пакетов в сетях, что упрощает детектирование аномалий.

Проверка устойчивости стеговложения (проверка работы метода)

Для построения вероятностной модели и проверки работы метода в среде использования данного ЦВЗ в доверенной среде, удобно использовать байесовские сети (далее — БС). БС — это ациклический ориентированный граф, в котором каждая вершина (узел сети) представляет n-значную переменную, дуги обозначают существование непосредственных причинно-следственных зависимостей между соединенными переменными, а сила этих зависимостей количественно выражается в виде условных вероятностей, сопоставленных каждой из переменных. Байесовские сети являются одним из видов вероятностных графических моделей. Строгое формальное определение и теория байесовских сетей доверия построены и развиты в трудах. Байесовские сети представляют собой удобный инструмент для описания достаточно сложных процессов и событий с неопределенностями. Для описания байесовской сети необходимо определить структуру графа и параметры каждого узла. Эта информация может быть получена непосредственно из данных или из экспертных оценок. В данном пункте описан процесс построения байесовской сети для анализа риска обнаружения ЦВЗ после обфускации [11]. Структура байесовской сети отражает возможные последствия использования обфускации, и является инструментом, с помощью которого можно выносить суждения и оценки относительно стойкости ЦВЗ после обфускации. Основные этапы этого процесса: идентификация переменных, определение структуры, определение параметров. После того, как байесовская сеть сконструирована, она готова для того, чтобы с её помощью можно было проводить вычисления. После поступление некоторого количества свидетельств, могут быть вычислены апостериорные вероятности [10].

Байесовская сеть представляет собой ориентированный граф без циклов, вершинами которого являются дискретные и (или) непрерывные случайные величины. При интерпретации дуг считается, что стрелка, ведущая от вершины x к вершине y, означает, что вершина x является родительской вершиной для вершины y и оказывает непосредственное влияние на y. При изображении байесовской сети родительские вершины изображаются на более высоком уровне, чем вершины потомки. Каждая вершина x характеризуется распределением условных вероятностей. Байесовская сеть служит правильным представлением проблемной области, если любая вершина на ней после задания родительских вершин становится условно независимой от других вершин, лежащих на более высоких уровнях.

Теоремой Байеса является простое соотношение, которое показывает, как условная вероятность зависит от обратной условной вероятности. Согласно теореме Байеса, вероятность события A при условии события B может быть вычислена следующим образом:

$$\mathrm{P}(\mathrm{A}|B) = \frac{\mathrm{P}(B|A)\mathrm{P}(A)}{\mathrm{P}(B)}, \quad (1)$$

где P(A) — вероятность гипотезы A, P(B) — вероятность события B, P(B|A) — вероятность наступления события B при истинности гипотезы A, P(A|B) — вероятность гипотезы A при наступлении события B.

Теорема Байеса показывает, как должна измениться вероятность в свете новых данных. В табл. 1 описаны существующие типы БС.

Выделим следующие характеристики (табл. 2), которые могут быть учтены в параметрах модели:

Выбранные характеристики описываются как непрерывными, так и дискретными величинами, следовательно, для построения вероятностной модели необходимо использовать гибридную БС. Для гибридных БС справедливо следующее ограничение: распределение непрерывной переменной X с дискретными родителями Y и непрерывными родителями Z является нормальным распределением:

Таблица 1. Существующие типы БС

Типы БС	Описание
Дискретные	Переменные узлы являются дискретными величинами. Каждая вершина представляет собой событие, описываемое случайной величиной; Все вершины, связанные с «родительскими», определяются таблицей условных вероятностей; Для вершин без «родителей» вероятности ее состояний являются маргинальными.
Динамические	Значения узлов изменяется со временем. Идеально подходят для моделирования временных процессов. Их преимущество в том, что они используют табличное представления условных вероятностей что, облегчает представление различных нелинейных явлений.
Непрерывные	Переменные узлов сети являются непрерывными величинами. Распределение вероятности для непрерывной случайной величины определяются иначе, чем в дискретном случае и для их описания используются функции распределения вероятностей и плотности распределения вероятностей.
Гибридные	Сети, содержащие как узлы с дискретными переменными, так и с непрерывными. Дискретные переменные не могут иметь непрерывных родителей. Непрерывные переменные должны иметь нормальный закон распределения, условный на значениях родителей. Распределение непрерывной переменной X с дискретными родителями Y и непрерывными родителями Z является нормальным распределением.

Таблица 2. Характеристики модели

Характеристика	Имя переменной	Тип переменной
Замусоривание стегоканала	A1	дискретная
Перестановки блоков стегоканала	A2	дискретная
Для замусоривания использовались те же инструкции, что и для замусоривания	B	дискретная
Инструкция попала в N стегоканала	C	непрерывная
Произошел разрыв M стегоканала	C1	непрерывная
Количество (K) вложений стегоканала	C3	непрерывная
Размер стеговложени	D	непрерывная
Размер секции стегоканала	D1	непрерывная
Длина стеговложений	X	непрерывная
Количество подходящих для вложения инструкций	X1	непрерывная
Количество блоков для перестановок	Y	непрерывная
K Стеговложения обнаружены	Z	непрерывная

$$P(X|Y=y,Z=z)=N\left(\mu_x(\mu_y\mu_z),\sqrt{\sigma_x(\sqrt{\sigma_y}}\right), \quad (2)$$

где μ_x, μ_y, μ_z — математические ожидания, σ_x, σ_y — дисперсии, $\sqrt{\sigma_x}$, $\sqrt{\sigma^y}$ — среднеквадратические отклонения. μ_x линейно зависит от непрерывных родителей, σ_x вообще не зависит от непрерывных родителей. Однако, оба они (μ^x и σ^x) зависят от дискретных родителей. Это ограничение гарантирует возможность точного вывода.

По схеме (рис. 5) сканеры сначала собирают особенности действий пользователя. Затем специальные алгоритмы выявляют в них особенности. Каждый признак проверяется как на наличие в нем аномалий, так и на отнесение его к категории «подозрительных на создание скрытых каналов». Полученные на выходе результаты подаются в интегральный классификатор, который уже дает окончательный ответ — присуще ли такое поведение стегоинсайдеру.

Активность стего-инсайдеров эмулировали путем создания стего-каналов на внешний сервер с помощью программного обеспечения из табл. 2 для различных типов входных и выходных данных (текст, изображения, звук, видео). Затем информация передавалась по созданному стего-каналу. Таким образом, помимо легального сетевого трафика, у организации были пакеты, формировавшие скрытые каналы. Параметры стего-каналов были разными и зависели от программного обеспечения, используемого для имитации действий инсайдеров. Всего было собрано 800 ГБ сетевого трафика, сгенерированного 200 пользователями в сети, в которую был установлен стенд. Из них около 5 ГБ (т.е. около 0,5%) составляли сгенерированные инсайдером

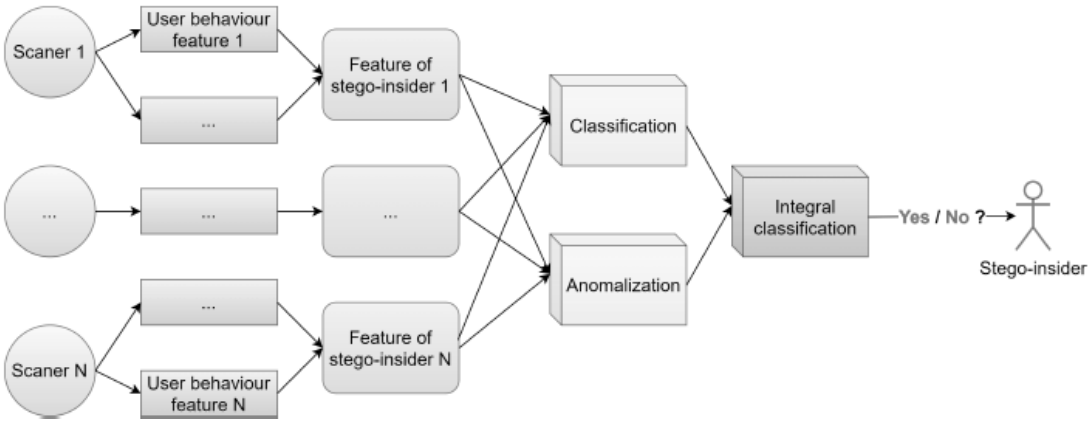

Рис. 5. Общая схема определения стего-инсайдера (Применение БС)

сетевые пакеты — содержащие процесс построения стего-канала и передачи по нему данных.

Заключение

Анализ (особенно в режиме реального времени) сетевого трафика требует работы с большим объемом данных по сбору, хранению, обработке. Эта проблема частично решается следующими способами. Во-первых, с помощью специализированной гибридной базы данных NoSQL. Во-вторых, отсутствие в документах базы данных полного содержания сетевого трафика и передаваемых объектов — поскольку сохраняются только признаки, необходимые для выявления стего-инсайдеров. В-третьих, потенциальным использованием управление ресурсами для параллельных баз данных.

Злоумышленник может использовать целый ряд программ, позволяющих ему встраивать сообщения в отправляемый контейнер. В качестве дальнейшего решения именно этой проблемы можно предположить создание и периодическое обновление сигнатурной (или иной) базы таких средств, ставя их в один ряд с вирусным ПО.

В случае шифрования стего-встраивания, очевидно, маловероятно получение исходного передаваемого сообщения. Однако первоначальной задачей было определение факта наличия скрытых каналов передачи. Получением исходного сообщения считается лишь дополнительная информация, позволяющая, в том числе, составить портрет правонарушителя, оценить причиненный ущерб и т.д.

В некотором смысле предлагаемые функции можно считать достаточно простыми, хотя они и не столь тривиальны, как особенности поведения пользователя. Однако в случае сложных сценариев злоумышленника (например, получение стего-вложений с одной рабочей станции и отправка от другого), они могут оказаться недостаточно эффективными. Решение может быть найдено путем введения новых функций, определяемых базой данных с использованием целого комплекса SQL-запросов. Также возможно добавление интеллектуальных функций, которые определяются не строгими правилами, а с помощью интеллектуальных агентов с возможностью самоорганизации — анализа поведения пользователей и устройств, присвоения им уровней доверия и даже, в некоторых случаях, корректировки прав доступа.

Причина этого кроется в «удобстве» данной среды для построения злоумышленником скрытых каналов. Для повышения эффективности предлагаемой модели и метода в данной среде целесообразно использовать более интеллектуальные алгоритмы определения признаков стего-инсайдера с учетом сложных взаимосвязей между признаками поведения пользователя (т.е. определяется вручную экспертом).

Если злоумышленник использует нестандартные способы передачи данных, очевидно, что особенности поведения и особенности стего-инсайдера не всегда будут корректно идентифицированы. Однако сам факт использования таких методов свидетельствует о какой-то аномалии в работе пользователя, а, следовательно, является подозрительным (ситуация аналогична шифрованию).

Дальнейшим развитием работы должно стать внедрение машинного обучения (в части окончательной классификации пользователей на правонарушителей и правонарушителей) и полноценное внедрение системы определения стего-инсайдеров. При этом требу-

ется расширить набор возможностей стего-инсайдера с учетом более сложных сценариев его поведения.

Также с научной и практической точки зрения будет востребовано создание отдельных метрик для оценки легальных пользователей с точки зрения их возможного перехода в категорию «стегоинсайдеров». Создание полноценного продукта, в том числе работающего в режиме реального времени, также запланировано на будущие исследования.

ЛИТЕРАТУРА

1. Базовая модель угроз безопасности персональных данных при их обработке в информационных системах персональных данных (выписка). ФСТЭК России, 2008 год
2. ГОСТ Р 53113.1–2008 «Защита информационных технологий и автоматизированных систем от угроз безопасности, реализуемых с использованием скрытых каналов».
3. Ушаков И.А., Обнаружение инсайдеров в корпоративной компьютерной сети на основе технологий анализа Больших данных // Вестник Санкт-Петербургского государственного университета технологии и дизайна. Серия 1: Естественные и технические науки. 2019. № 4. С. 38–43.
4. Г.А. Орлов, А.В. Красов, А.М. Гельфанд, Применение Big Data при анализе больших данных в компьютерных сетях // Наукоемкие технологии в космических исследованиях Земли. — 2020. — Т. 12. — № 4. — С. 76–84. — DOI 10.36724/2409–5419–2020–12–4–76–84.
5. А.В. Красов, С.И. Штеренберг, Д.Р. Голузина, Методика визуализации больших данных в системах защиты информации для формирования отчетов уязвимостей // T-Comm: Телекоммуникации и транспорт. — 2018. — Т. 12. — № 10. — С. 36–40. — DOI 10.24411/2072–8735–2018–10154.
6. V.K. Fedorov, E.G. Balenko, S.I. Shterenberg, A.V. Krasov, Development of a method for building a trusted environment by using hidden software agent steganography // Journal of Physics: Conference Series, Vladivostok, 07–08 октября 2021 года. — Vladivostok, 2021. — P. 012047. — DOI 10.1088/1742–6596/2096/1/012047.
7. А.С. Салита, А.В. Красов, Создание стеганографического канала при помощи полей // Вестник Санкт-Петербургского государственного университета технологии и дизайна. Серия 1: Естественные и технические науки. — 2021. — № 2. — С. 36–40. — DOI 10.46418/2079–8199_2021_2_6.
8. I. Kotenko, K. Izrailov, A. Krasov, I. Ushakov, An approach for stego-insider detection based on a hybrid NOSQL database // Journal of Sensor and Actuator Networks. — 2021. — Vol. 10. — No 2. — DOI 10.3390/jsan10020025.
9. I. Kotenko, A. Krasov, I. Ushakov, K. Izrailov, Detection of stego-insiders in corporate networks based on a hybrid NoSQL database model // ACM International Conference Proceeding Series: 4, St. Petersburg, 26–27 ноября 2020 года. — St. Petersburg, 2020. — P. 3442612. — DOI 10.1145/3440749.3442612.
10. Штеренберг С.И., Стародубцев И.В., Шашкин В.С., Разработка комплекса мер для защиты предприятия от фишинговых атак // Защита информации. Инсайд. 2020. № 2 (92). С. 24–31.
11. Штеренберг С.И., Методика построения поисковой системы для примитивной программы адаптивного действия // Наукоемкие технологии в космических исследованиях Земли. 2015. Т. 7. № 4. С. 52–57.

Журнал «Современная наука: актуальные проблемы теории и практики»