DOI: 10.17725/rensit.2020.12.287

Детектирование DoS атак, использующих CONNECT сообщения протокола MQTT

Дикий Д.И.

Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики (Университет ИТМО), http://www.itmo.ru/

Санкт-Петербург 197101, Российская Федерация

E-mail: dimandikiy@mail.ru

Поступила 25.09.2019, рецензирована 10.03.2020, после доработки 30.03.2020, принята 13.04.2020

Представлена действительным членом РАЕН А.С. Дмитриевым

Аннотация. **Обнаружение DoS-атак в рамках Интернета вещей является актуальной задачей для обеспечения безопасности этой инфраструктуры. При реализации атаки злоумышленник генерирует большое число запросов на подключение к сети Интернет вещей по протоколу MQTT, что делает коммутационный узел недоступным для других пользователей. Рассмотрены средства и методы детектирования атак как для сетей Интернет в целом, так и для сетей Интернет вещей. Для детектирования атак на основе анализа сетевого трафика предложен метод формирования вектора признаков. Вектор признаков состоит из параметров частоты передачи сообщений за интервал времени для устройства, имеющего один и тот же ip-адрес. В качестве классификаторов были рассмотрены многослойный персептрон, алгоритм случайный лес и метод опорных векторов. Была собрана экспериментальная установка, на которой были сгенерированы обучающие и тестовые выборки с заданными параметрами трафика. Эксперимент показал, что для достижения максимального качества классификации увеличение размерности вектора признаков не требуется. Было проведено сравнение выше перечисленных алгоритмов по значению F1-меры, в ходе которого выяснилось, что лучше других с задачей классификации справляется искусственная нейронная сеть в виде многослойного персептрона. При этом интервал времени, на основании которого формируется вектор признаков, должен превышать 1.5 секунды для достижения значения F1-меры более 0.99 при частоте подключения легального устройства один раз в секунду. Исследование показало эффективность применения рассмотренных классификаторов на базе предлагаемого вектора признаков для детектирования атаки на отказ в обслуживании.**

Ключевые слова: **интернет вещей, отказ в обслуживании, MQTT, машинное обучение, случайный лес, многослойный персептрон, метод опорных векторов, телекоммуникации, детектирование атак**

УДК 004.052.3

Благодарности. Исследование выполнено при финансовой поддержке РФФИ в рамках научного проекта №19-37-90051.

Для цитирования: Дикий Д.И. Детектирование DoS атак, использующих CONNECT сообщения протокола MQTT. *РЭНСИТ,* 2020, 12(2):287-296. DOI: 10.17725/rensit.2020.12.287.

Detection of DoS attacks caused by CONNECT messages of MQTT protocol

Dmitrii I. Dikii

St. Petersburg National Research University of Information Technologies, Mechanics and Optics (University ITMO), http://www.itmo.ru

St. Petersburg 197101, Russian Federation

E-mail: dimandikiy@mail.ru

Received September 25, 2019, reviewed on March 10, 2020, after finalization on March 30, 2020 accepted April 13, 2020

Abstract. **Detecting DoS attacks within the Internet of Things is an urgent task to ensure the security of this infrastructure. The malefactor, undertaking the attack, generates a large number of connection requests to the Internet of Things network based on the MQTT protocol. This makes the gateway unavailable for other users. The author discusses the approaches and methods of detecting DoS attacks within the Internet, in general, as well as within the Internet of Things, in particular. The method of feature vector generation for detecting DoS attacks based on the network traffic analysis was suggested. The feature vector consists of parameters of message transmission frequency within a time interval from a device with the same IP-address. The multilayer perceptron, the random forest algorithm, the support vector machine are classifiers in this study. The author constructed an experimental assembly to generate training and testing sets with the supplied parameters. The experiment results showed: in order to achieve maximum classification accuracy, the dimension increase of the feature vector is not required. A comparison of the mentioned above algorithms by the F1-score value was carried out, which proved the artificial neural network – the multilayer perceptron – to be the best classifier. At that, the time interval, on which the feature vector generation is based, must be higher than 1.5 seconds for the accuracy to be over 0.99 under the legal device connection frequency once per second. The research gave positive results of implementing the reviewed classifiers based on the suggested feature vector to detect DoS attacks**

Keywords: **Internet of Things, DoS, MQTT, machine learning, random forest, multilayer perceptron, support vector machine, telecommunication, attack detection**

UDC 004.052.3

Acknowledgments. The reported study was funded by RFBR, project number №19-37-90051

For citation: Dmitrii I. Dikii. Detection of DoS attacks caused by CONNECT messages of MQTT protocol. *RENSIT,* 2020, 12(2):287-296. DOI: 10.17725/rensit.2020.12.287.

Содержание

1. ВВЕДЕНИЕ

В последнее время большую популярность набирают технологии, активно использующие сети Интернет. Одной из таких технологий является Интернет вещей [1]. Основная особенность технологии, позволяющей объединить множество проектов под одним названием Интернет вещей, – это возможность коммуникации большого числа устройств, функционирующих без оператора для выполнения одной общей задачи. Устройства, о которых будет идти речь, должны обладать только самым необходимым функционалом, что существенно удешевляет их по сравнению с обычной рабочей станцией (персональный компьютер, смартфон и т.д.). Некоторые устройства сетей Интернет вещей функционируют от автономного источника питания, что накладывает ограничения на их использование с точки зрения экономии электропотребления. Для увеличения срока функционирования от автономного источника питания создаются технологии и протоколы передачи данных, существенно сокращающие энергетические затраты конечного устройства. Разработки в области сетей Интернет вещей охватывают весь стек протоколов модели OSI [2]. Одним из таких протоколов является MQTT (message quality telemetry transport), разработанный альянсом OASIS [3]. На данный момент широко распространена версия 3.1.1 протокола MQTT.

Наряду с тенденцией к упрощению протоколов для устройств Интернет вещей, наблюдается рост угроз информационной безопасности. Информация, циркулирующая в сети Интернет вещей, зачастую находится в незашифрованном виде. Это может привести к большим негативным последствиям со стороны обладателя информации. Наиболее ярким примером в данном вопросе является сфера медицины. В работах [4,5] представлены идеи

авторов по совершенствованию безопасности именно в сфере здравоохранения. В качестве одного из методов предлагается усиленная аутентификация устройств.

Помимо угроз, свойственных именно сетям Интернет вещей, не стоит упускать из внимания угрозы, которые распространены среди всех устройств, имеющих выход в сеть Интернет. Как правило, это атаки вида человек-по-середине, фишинг, вирусы, трояны и др. Одной из таких угроз является распределенная атака на отказ в обслуживании. Она характеризуется большим количеством абонентов сети, отправляющих запросы устройству-жертве. Из-за превышения максимального числа обрабатываемых запросов за момент времени жертва не справляется с нагрузкой и становится недоступной для других устройств. Согласно аналитике от Лаборатории Касперского за первый квартал 2019 года [6] использование вредоносного программного обеспечения на базе ботнета Mirai получило наибольшее распространение по всему миру. Атака на отказ в обслуживании, как часть поведения ботнет сетей, становится одной из самых актуальных угроз.

Таким образом, разворачивая инфраструктуру сети Интернет вещей на предприятии или в собственном доме, стоит учитывать как классические угрозы информационной безопасности, так и специфичные для сетей Интернет вещей, а также их комбинации.

В данной работе будет рассмотрена проблема злоупотребления устройствами сети Интернет вещей возможностей протокола MQTT для реализации атаки вида отказ в обслуживании.

На сегодняшний день выделяют следующие виды атак на отказ в обслуживании:

- атака на истощение полосы пропускания;
- атака на истощение ресурсов жертвы;
- атака на инфраструктуру;
- атака нулевого дня [7].

Злоумышленники чаще всего используют первые два вида атак на отказ в обслуживании. Первый заключается в насыщении полосы передачи информации до такой степени, что сигнал от легитимного источника не проходит до адресата. Второй тип атак эксплуатирует уязвимости протоколов, таким образом, чтобы исчерпать ресурсы сервера: память ОЗУ, процессорное время. Такая атака выполняется как на протоколах сетевого и транспортного уровней, так и прикладного, например, HTTP. Ярким примером является атака вида TCP SYN, когда злоумышленник отправляет запрос на установление соединения по протоколу TCP, но вместо собственного ip-адреса, указывает несуществующий. Сервер ожидает завершение транзакции установления соединения и не получает ответ длительное время. Однако информация о незавершенном соединении сохраняется на стороне сервера, что приводит к истощению ресурсов.

Мероприятия по защите информационных систем от этого вида атак можно разделить на два этапа:

- детектирование;
- противодействие.

Детектирование производится за счет анализа сетевого трафика. Большое распространение получил метод фильтрации пакетов “hop count” [8]. В этом методе оценивается количество TCP пакетов и их параметры: SYN флаг, TTL, адрес отправителя и др. В работе [9] рассмотрен метод определения атаки на отказ в обслуживании и защиты от нее, состоящий из фильтрации MAC-адресов и криптографических преобразований.

Для детектирования атаки все чаще предлагают методы, основанные на искусственном интеллекте и машинном обучении. Так, в работе [10] авторы предлагают использовать роевые алгоритмы. Точность детектирования атаки по предлагаемому методу составила 0.75-0.80.

При защите интернет ресурсов чаще всего анализируется TCP трафик путем оценки времени ответа сервера при обычном трафике и при атаке. При атаке время ответа сервера значительно увеличивается. Этот факт лежит в основе классификации трафика. Например, с помощью алгоритма LS-SVM (модифицированный метод опорных векторов) удалось добиться точности классификации более 0.92 [11].

Метод опорных векторов (SVM) для детектирования атак на отказ в обслуживании был рассмотрен во многих других работах. Например, в [12] при использовании SVM на базе данных DARPA удалось добиться 99% верного распознавания атаки. Такую же точность

определения аномального трафика удалось достичь авторам работ [13,14]. Различные модификации SVM позволяют достичь точности классификации более 0.92 [15]. Разница между исследованиями, которые посвящены использованию SVM для детектирования DoS атак, заключается в разных способах формирования вектора признаков. В работе [16] приведено исследование влияния того или иного признака на точность классификации. Выбор вектора признаков (размерность и его состав) является основным фактором, влияющим на точность. Однако во всех работах алгоритм SVM либо его модификации показали очень хорошие результаты .

Другим подходом к детектированию атак является использование искусственных нейронных сетей (ИНС). Модификаций искусственных нейронных сетей существует огромное количество. Самая распространенная модель, многослойный персептрон (MLP), рассмотрена в работе [11]. Сравнение SVM и ИНС показало, что последний алгоритм имеет меньшую точность и требует больше времени на принятие решения [17]. В работе [18] представлены результаты эксперимента по использованию ИНС для детектирования аномалий в трафике по TCP и ICMP протоколам. Подход, предлагаемый авторами, достиг точности детектирования атаки 0.98. Кроме того, применяются ансамбли рекуррентных искусственных сетей [19].

Для детектирования атак используются алгоритмы случайный лес (RF) и деревья решений. Этот подход показывает неплохие результаты. Например, в работе [20] точность детектирования составила более 0.96. Аналогичные работы [21,22,23] также демонстрируют довольно высокую точность классификации. Также применяются подходы, основанные на нечеткой логике [24].

Методы детектирования аномального трафика в сетях Интернет вещей, как правило, основаны на анализе данных сетевых и транспортных протоколов, как предложено в [25,26]. Но в сетях Интернет вещей используют протоколы других уровней, которые уязвимы к атаке на отказ в обслуживании. Это как протоколы прикладного уровня (CoAP, MQTT), так и протоколы более низких уровней (LoRa). Атаки на физическом и канальном уровнях наиболее распространены в беспроводных сенсорных сетях. Например, атака, нацеленная на истощение энергетических ресурсов, описана в [27]. Другой тип атак, свойственный сети Интернет вещей — “blackhole”. Здесь устройство сообщает другим участникам сети о том, что через его узел будет самый короткий путь для доставки пакета. Однако все пакеты, поступающие на этот узел, сбрасываются [28]. Так же выделяют атаки, формирующие помехи для передачи информации по радиоканалам, тем самым, вызывая отказ в обслуживании [29].

2. МАТЕРИАЛЫ И МЕТОДЫ

Применительно к протоколу прикладного уровня MQTT выделяется предрасположенность к атаке на отказ в обслуживании. Обычно атака реализуется за счет увеличения нагрузки на элементы сети таким образом, чтобы коммуникация между устройствами нарушилась. Протокол функционирует по структуре издатель-подписчик. Таким образом, в сети имеется ключевой элемент, называемый шлюзом. Он отвечает за перенаправление сообщений от отправителя к получателю. Так как все сообщения проходят через шлюз, то он наиболее уязвим к потенциальной атаке. Были проведены исследования о влиянии параметров сообщений (флагов, количество сообщений и т.д.) на устойчивость шлюза к большим нагрузкам. В большинстве работ рассматривались только сообщения вида PUBLISH с учетом следующих параметров:

- качество доставки (QoS) [30, 31];
- количество получателей [32];
- размер сообщения [33, 34];
- криптографические преобразования над сообщениями [32].

Из виду упускается процесс подключения устройства к шлюзу. При одновременной отправке большого количества запросов на подключение (CONNECT сообщения), шлюз не справляется с нагрузкой, что не позволяет легальным устройствам подключиться к нему [35]. В сетях, использующих протокол MQTT, необходимо обнаруживать аномальное

поведение устройств на всех стадиях работы протокола.

Целью данной работы является разработка метода обнаружения атаки на отказ в обслуживании, вызываемой аномальным поведением устройств сети при использовании CONNECT сообщений протокола MQTT, с помощью алгоритмов машинного обучения. Для достижения данной цели в первую очередь решается задача выбора оптимального вектора признаков. Вторая задача, которую необходимо решить: определить наиболее эффективный метод классификации. В качестве классификаторов в данной работе были рассмотрены следующие алгоритмы: многослойный персептрон, случайный лес и метод опорных векторов с радиально-базисной функцией ядра, программно реализованные на базе проекта WEKA [36]. Для генерации обучающих и тренировочных наборов данных была создана экспериментальная установка (**рис. 1**), состоящая из шлюза, коммуникационного оборудования и нескольких ЭВМ, моделирующих поведение множества устройств сети Интернет вещей с помощью фреймворка Paho-MQTT [37]. В качестве шлюза использовался микрокомпьютер Raspberry Pi 3 model B с программным исполнением Moquette на языке JAVA [38].

Для того, чтобы правильно классифицировать сообщение, необходимо сформировать вектор признаков. Так как для формирования запроса на подключение злоумышленнику достаточно знать адрес шлюза и номер порта, то такая служебная информация, как идентификатор устройства и имя пользователя, могут быть сгенерированы автоматически, и не рассматриваются. Таким образом, основные параметры, описывающие CONNECT сообщение по протоколу MQTT, это:

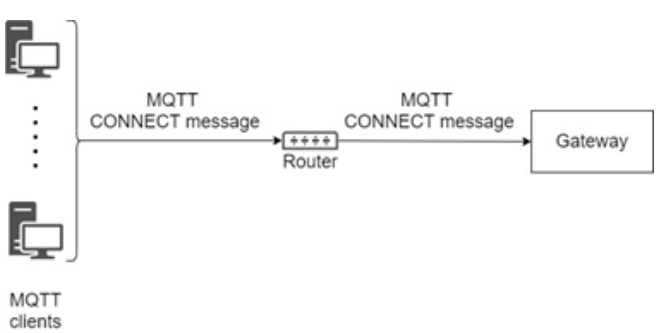

Рис. 1. *Схема экспериментальной установки.*

- адрес отправителя в виде ip-адреса — необходим для ведения черного списка адресов, с которых происходит атака. В данном случае ip-адрес используется в качестве ярлыка;
- количество запросов на подключение за интервал времени.
- математическое ожидание времени между запросами на подключение за интервал времени;
- бинарное значение, определяющее использование криптографических преобразований по протоколу TLS, которое значительно влияет на время подключения к шлюзу: 0 – протокол TLS не используется, 1 – протокол TLS используется.

Таким образом, вектор признаков сообщения о подключении состоит из трех основных параметров за один анализируемый интервал времени (далее **m**).

Выбор интервала времени **m** играет важную роль в формировании вектора признаков. Также это может быть не один интервал времени, а их совокупность. Тогда размерность вектора признаков вычисляется по формуле:

$$W = 1+3k, \quad (1)$$

где k – количество рассматриваемых интервалов времени **m**.

Под интервалом времени **m** понимается промежуток на временной оси между моментом получения сообщения на шлюзе и моментом заданного числа миллисекунд до этого события. Пример формирования совокупности из трех интервалов времени представлен на **рис. 2**.

Легальный трафик был создан на основе

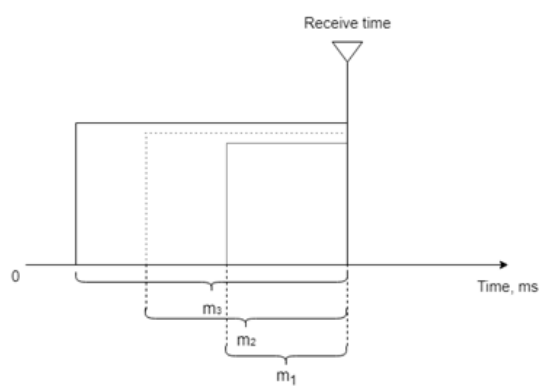

Рис. 2. *Изображение интервалов времени m, на основании которых формируется вектор признаков,* $m_3 > m_2 > m_1$.

моделирования поведения устройств сети с учетом установления незащищенных и защищенных по протоколу TLS соединений. Моделирование легального трафика зависит от условий практического применения сети, поэтому тренировочный набор данных будет отличаться в зависимости от проектируемой максимальной нагрузки сети. Аномальный трафик моделировался за счет генерации большого потока сообщений с запросами на подключение, как по открытому каналу, так и по защищенному. Выбор защищённости канала определялся случайным образом для каждого подключения.

Чтобы определить лучший классификатор, был сгенерирован тестовый набор данных, содержащий примеры легитимного и аномального трафика. Для этого использовался сценарий работы сети в штатном режиме, содержащий десять тысяч подключений, и сценарий при потенциальной атаке, состоящий из пяти тысяч последовательно отправленных CONNECT сообщений.

Актуальным остался вопрос о выборе интервалов времени, по которым будет сформирован вектор признаков. Для выбора оптимальных наборов временных интервалов предлагается следующая методика.

На первом шаге экспертной оценкой определяется конечное множество интервалов времени **M** с натуральными значениями ($\mathbf{M} \in \mathbf{N}$).

На втором шаге для каждого значения $\mathbf{m} \in \mathbf{M}$ производится обучение трех классификаторов (MLP, RF, SVM) и проверка классификатора на тестовом наборе данных.

На третьем шаге производится оценка качества классификации путем расчета F1-меры - средневзвешенного значения точности и полноты. Данная метрика широко используется при оценке качества бинарной классификации для методов машинного обучения, как показано в работах [24, 39]. Для вычисления этой метрики используются результаты о количестве правильно и неправильно классифицированных сообщений сообщений на тестовом наборе данных (**Таблица 1**, где TP – количество легитимных сообщений, распознанных верно; TN – количество сообщений атаки, распознанных верно; FP – количество аномальных сообщений, распознанных неверно; FN – количество легальных сообщений, распознанных неверно).

Таблица 1

Матрица ошибок

	Легальные сообщения	Нелегальные сообщения
Правильно классифицированные сообщения	TP	TN
Неправильно классифицированные сообщения	FN	FP

Рассчитывается точность классификаций по формуле:

$$\text{Precision} = TP/(FP+TP) \quad (2)$$

и полнота классификации по формуле:

$$\text{Recall} = TP/(TP+FN). \quad (3)$$

Зная эти значения, вычисляется F1-мера по формуле:

$$F = (2\times\text{Precision}\times\text{Recall})/(\text{Precision}+\text{Recall}). \quad (4)$$

На четвертом шаге формируется конечное множество **S** из уникальных комбинаций неповторяющихся элементов $\mathbf{m} \in \mathbf{M}$ длиной l, такой, что $2 \leq l \leq |\mathbf{M}|$.

На пятом шаге повторяются шаги 2-3 с интервалами времени из множества **S**. По окончанию всех вычислений определяется лучшая комбинация временных интервалов, при которой достигается наибольшее значение F1-меры для того или иного метода.

3. РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЯ

Для смоделированной сети был определен следующий набор начальных интервалов времени для множества **M** {20, 50, 100, 150, 200, 250, 500, 1000, 1500, 2000, 3000}.

Обучающая выборка состоит из двух массивов данных. Первый содержит информацию о сообщениях при легальном потоке сообщений о подключении к шлюзу, второй – о потоке, схожем с атакой на отказ в обслуживании. Для генерации легитимного потока данных была определена следующая модель. Определяется интервал времени **I**, в течение которого гарантированно отправляется хотя бы одно сообщение на подключение. Момент отправки **i** выбирается случайным образом (по равномерному распределению вероятностей, так чтобы в наборе данных присутствовали примеры как с небольшими, так и близкими к максимальным из интервала **I** значениями

задержки между сообщениями) из интервала **I** (**i** ∈ **I**). Таким образом, разность времени между отправкой двух последовательных сообщений определяется как:

$$\Delta T = i + t, \quad (5)$$

где t – время необходимое на обработку сообщения и получение ответа от шлюза (по незащищенному каналу в среднем составляет не более 50 мс, по защищенному – не более 700 мс для данной экспериментальной установки), **i** – случайная задержка времени не больше максимального значения интервала **I**.

Чтобы определить влияние обучающей выборки, содержащей легитимный трафик, на качество классификации было рассмотрено два интервала **I** = [0,1000] мс и **I** = [0,500] мс.

Для генерации массива данных, содержащего данные об аномальном трафике, был смоделирован большой поток сообщений на подключение к шлюзу за короткий промежуток времени. Обучающая выборка состоит из десяти тысяч сообщений для легитимного потока данных и пяти тысяч сообщений для моделирования атаки на отказ в обслуживании. Результаты классификации на тестовой выборке при использовании одного интервала времени **m** ∈ **M** приведены на **рис. 3**.

Из результатов проведенного эксперимента следует, что классификатор, построенный на модели многослойного персептрона, показал лучшие результаты. При увеличении интервала времени, в течение которого собирается статистика о трафике, наблюдается увеличение значения F1-меры. Например, при интервале в две секунды это значение достигает 0.9989 ± 0.0001.

Для алгоритма случайного леса наблюдается сложная динамика. При увеличении интервала времени с 20 мс до 1500 мс значение F1-меры также увеличивается. Однако последующее увеличение интервала времени до трех секунд приводит к ухудшению характеристик классификатора из-за повышения количества ложно отрицательных классификаций. Максимальное значение F1-меры достигается при интервале **m** = 1500 мс и составляет более 0.9934 ± 0.0027.

Хуже других алгоритмов с задачей классификации справился метод опорных векторов. Значение F1-меры при каждом из рассматриваемых интервалов времени меньше, чем у других алгоритмов. Максимальное значение F1-меры достигается при **m** = 500, при этом значение составляет 0.985 ± 0.0021. При увеличении интервала времени до трех секунд результаты классификации ухудшаются.

Применение совокупности из нескольких интервалов **m** ∈ **M** не дало значительного положительного эффекта. На **рис. 4** представлены значения F1-меры для рассматриваемых алгоритмов при использовании следующих наборов интервалов: {200, 250, 500}, {250, 500}, {200,500}. Для сравнения на графике также приведены значения для интервалов {200}, {250}, {500}. Можно сделать вывод о том, что применение совокупности интервалов не повышает точность классификации, а зачастую ухудшает ее. Например, при использовании метода опорных векторов наблюдается явная отрицательная динамика изменения значения F1-меры при увеличении признакового пространства. В большинстве других случаев значения F1-меры меньше, либо незначительно больше, показателей при классификации с использованием одного наибольшего интервала.

Таким образом, использование совокупности интервалов для формирования вектора признаков большей размерности является нецелесообразным.

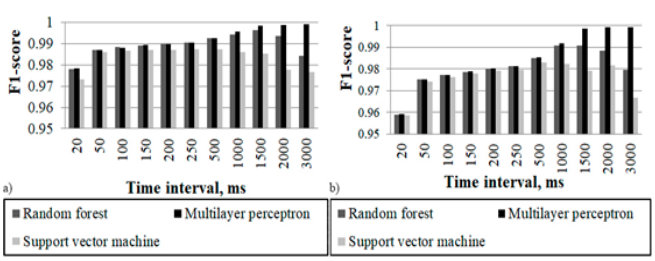

Рис. 3. *Результаты классификации при частоте легальных сообщений из интервала а) I = [0, 500] мс, b) I = [0, 1000] мс.*

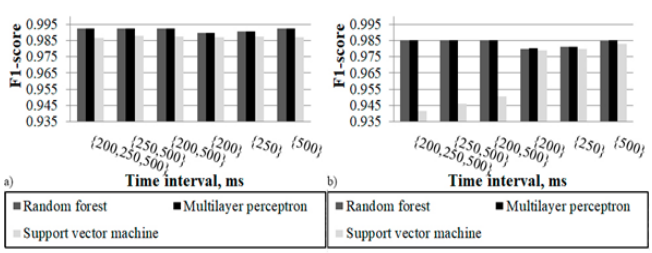

Рис. 4. *Результаты классификации при частоте легальных сообщений из интервала а) I = [0, 500] мс, б) I = [0, 1000] мс.*

4. ОБСУЖДЕНИЕ

В рамках проведенного исследования были проанализированы методы детектирования атак на отказ в обслуживании, применяемые как в сетях Интернет, так и в сетях Интернет вещей. Для детектирования атаки, реализуемой посредством злоупотребления сообщениями вида CONNECT по протоколу MQTT, был предложен вектор признаков, состоящий из трех основных параметров: количество сообщений за интервал времени, математическое ожидание времени между двумя последовательными сообщениями, среднее значение параметра, отвечающего за использование протокола TLS при создании криптографически защищенного канала. Еще один параметр используется в качестве ярлыка: ip-адрес отправителя.

В качестве классификаторов были рассмотрены следующие алгоритмы: многослойный персептрон, алгоритм случайный лес, метод опорных векторов. Эксперимент на основе сгенерированных обучающих и тестовых выборках показал, что все алгоритмы справляются с задачей классификации трафика с точностью более 0.90. Наиболее лучшим по качеству классификации является искусственная нейронная сеть в виде многослойного персептрона. При увеличении значения интервала времени, в течение которого собирается статистика о трафике и формируется вектор признаков, значение F1-меры также увеличивается в отличие от других методов. Алгоритм случайный лес чуть хуже справился с задачей классификации. Увеличение интервала времени до 1.5 секунд положительно сказывается на значении F1-меры. Однако при дальнейшем увеличении этого интервала значение F1-меры уменьшается. Хуже других алгоритмов справился метод опорных векторов. Динамика F1-меры аналогична алгоритму случайного леса. Максимальное значение F1-меры наблюдается при интервале 500 мс. Использование векторов признаков большей размерности признано нецелесообразным, так как в таком случае характеристики классификации могут не только не улучшиться, но и ухудшиться.

5. ЗАКЛЮЧЕНИЕ

Таким образом, среди всех рассмотренных подходов и алгоритмов для детектирования атак на отказ в обслуживании по предлагаемому вектору признаков (вектор признаков в таком случае будет иметь размерность четыре) рекомендуется использовать многослойный перцептрон, т.к. этот классификатор показал наилучшие результаты среди всех рассмотренных методов. При этом качество классификации повышается при увеличении интервала времени, в течение которого собирается статистика о трафике. Но стоит отметить, что увеличение этого интервала приведет к большим вычислительным и временным затратам на обучение модели и принятие решения. Качество классификации на основе алгоритма случайный лес или метода опорных векторов с радиально-базисной функцией ядра хуже по сравнению с многослойным перцептроном, однако значения F1-меры достаточно высоки, что также позволяет их использовать.

Дальнейшие исследования будут направлены на изучение атак на отказ в обслуживании, вызванные злоупотреблением другими видами сообщений протокола MQTT.

ЛИТЕРАТУРА

1. Ashton K. That 'Internet of Things' Thing. *RFID Journal*, 2009, 22:97–114.
2. Стандарт "ISO/IEC 7498-1:1994 [ISO/IEC 7498-1:1994] Information technology — Open Systems Interconnection — Basic Reference Model: The Basic Model". *ISO/IEC Information Technology Task Force (ITTF) web site*, 1994.
3. Стандарт ISO/IEC 20922:2016 Information technology – Message Queuing Telemetry Transport (MQTT) v3.1.1. *ISO/IEC Information Technology Task Force (ITTF) web site,* 2016.
4. Albalawi U, Joshi S. Secure and Trusted Telemedicine in Internet of Things IoT. *Proceedings of 2018 IEEE 4th World Forum on Internet of Things* (WF-IoT), 2018, pp. 30-34. DOI: 10.1109/WFIoT.2018.8355206.
5. Wazid M, Kumar Das A, Khurram Khan M, Al Dhawailie AlGhaiheb A, Kumar N, Vasilakos AV. Secure Authentication Scheme for Medicine Anti-Counterfeiting System in IoT Environment. *IEEE Internet of Things Journal,* 2017, 4(5):1634-

1646. DOI: 10.1109/JIOT.2017.2706752.
6. Чебышев В, Синицын Ф, Паринов Д, Ларин Б, Купреев О, Лопатин Е. Развитие информационных угроз в первом квартале 2019 года. Статистика. *Kaspersky Security Bulletin 2019. Статистика.* Отчеты об угрозах URL: https://securelist.ru/it-threat-evolution-q1-2019-statistics/94021/ (дата обращения: 20.08.2019).
7. Mahjabin T, Xiao Y, Sun G, Jiang W. A survey of distributed denial-of-service attack, prevention, and mitigation techniques. *International journal of distributed senssor networks*, 2017, 13(12):1-32. DOI: 10.1177/1550147717741463.
8. Jin C., Wang H., Shin K.G Hop-Count Filtering: An effective defense against spoofed DDoS traffic. *Proc. of the ACM Conf. on Computer and Communications Security,* 2003, pp. 30-41. DOI: 10.1145/948109.948116.
9. Prakash A, Satish M, Sri Sai Bhargav T, Bhalaji N. Detection and Mitigation of Denial of Service Attacks Using Stratified Architecture. *Proc. of the 4th Intern. Conf. on Recent Trends in Computer Science & Engineering Detection Procedia Computer Science,* 2016, 87:275-280. DOI: 10.1016/J.PROCS.2016.05.161.
10. Sharma S, Gupta A, Agrawal S. An Intrusion Detection System for Detecting Denial-of-Service Attack in Cloud Using Artificial Bee Colony. *Proc. of the Intern. Congress on Information and Communication Technology, Advances in Intelligent Systems and Computing,* 2016, pp. 137-145. DOI: 10.1007/978-981-10-0767-5_16.
11. Sahi A, Lai D, LI Y, Diykh M. An Efficient DDoS TCP Flood Attack Detection and Prevention System in a Cloud Environment. *IEEE Access,* 2017, 5:6036-6048. DOI: 10.1109/ACCESS.2017.2688460.
12. Mukkamala S, Sung AH. Detecting Denial of Service Attacks Using Support Vector Machines. *Proc. of the 12th IEEE Intern. Conf. on Fuzzy Systems,* 2003, pp.1231-1236. DOI: 10.1109/FUZZ.2003.1206607.
13. Manuel S. Hoyos Ll, Gustavo AIE, Jairo IV, Castillo OL. Distributed Denial of Service (DDoS) Attacks Detection Using Machine Learning Prototype. *Proc. of the 13th Intern.Conf., Advances in Intelligent Systems and Computing,* 2016, pp. 33-41. DOI: 10.1007/978-3-319-40162-1_4.
14. Kim D, Lee KY. Detection of DDoS Attack on the Client Side Using Support Vector Machine. *Intern. J. of Applied Engineering Research,* 2017, 12(20):pp. 9909-9913.
15. Xu X, Wei D, Zhang Y. Improved Detection Approach for Distributed Denial of Service Attack Based on SVM. *Proc. of the 3th Pacific-Asia Conference on Circuits, Communications and System (PACCS)*, 2011, pp. 1-3. DOI: 10.1109/PACCS.2011.5990284.
16. Chan APF, Ng WWY, Yeung DS, Tsang ECC. Refinement of rule-based intrusion detection system for denial of service attacks by support vector machine. *Proc. of the 13rd Intern. Conf. on Machine Learning and Cybernetics*, 2004, pp. 4252-4256. DOI: 10.1109/ICMLC.2004.1384585.
17. Tsang GCY; Chan PPK; Yeung DS; Tsang ECC. Denial of service detection by support vector machines and radial-basis function neural network. *Proc. of Intern. Conf. on Machine Learning and Cybernetics (IEEE Cat. No.04EX826)*, 2004, pp. 4263-4267. DOI: 0.1109/ICMLC.2004.1384587.
18. Saied A, Overill RE, Radzik T. Detection of known and unknown DDoS attacks using Artificial Neural Networks. *Neurocomputing,* 2016, 172:385-393. DOI: 10.1016/j.neucom.2015.04.101.
19. Al Islam ABMA, Sabrina T. Detection of various Denial of Service and Distributed Denial of Service Attacks using RNN Ensemble. *Proc. of 12th Intern. Conf. on Computer and Information Technology (ICCIT 2009),* 2009, pp. 603-608. DOI: 10.1109/ICCIT.2009.5407308.
20. Lakshminarasimman S; Ruswin S; Sundarakantham K. Detecting DDoS attacks using decision tree algorithm. *Proc. of 4th Intern. Conf. on Signal Processing, Communication and Networking (ICSCN)*, 2017, pp.1-6. DOI: 10.1109/ICSCN.2017.8085703.
21. Chen L, Zhang Y, Zhao Q, Geng G, Yan Z. Detection of DNS DDoS Attacks with Random Forest Algorithm on Spark. *Proc. of 2nd Intern. Workshop on Big Data and Networks Technologies Procedia Computer Science,* 2018, 134:310-315. DOI: 10.1016/j.procs.2018.07.177.
22. Idhammad M, Afdel K, Belouch M. Detection System of HTTP DDoS Attacks in a Cloud Environment Based on Information Theoretic Entropy and Random Forest. *Security and Communication Networks*, 2018, 2018:1-13. DOI: 10.1155/2018/1263123.
23. Cheng J, Li M, Tang X, Sheng VS, Liu Y, Guo

W. Flow Correlation Degree Optimization Driven Random Forest for Detecting DDoS Attacks in Cloud Computing. *Security and Communication Networks,* 2018, 2018:1-14. DOI: 10.1155/2018/6459326.

24. Haripriya AP, Kulothungan K. Secure-MQTT: an efficient fuzzy logic-based approach to detect DoS attack in MQTT protocol for internet of things. *Journal on Wireless Communications and Networking*, 2019, Vol. 90. DOI: 10.1186/s13638-019-1402-8.
25. Doshi R, Apthorpe N, Feamster N. Machine Learning DDoS Detection for Consumer Internet of Things Devices. *Proc. of IEEE Symposium on Security and Privacy Workshops,* 2018, pp. 29-35. DOI 10.1109/SPW.2018.00013.
26. Meidan Y, Bohadana M, Mathov Y, Mirsky Y, Breitenbacher D, Shabtai A, Elovici Y. N-BaIoT: Network-based Detection of IoT Botnet Attacks Using Deep Autoencoders. *IEEE Pervasive computing,* 2018, 13(9):1-8. DOI: 10.1109/MPRV.2018.03367731.
27. Mallikarjunan KN, Muthupriya K, Shalinie SM. A survey of Distributed Denial of Service attack. *Proc. of 10th Intern. Conf. on Intelligent Systems and Control (ISCO)*, 2016, pp. 1-6. DOI: 10.1109/ISCO.2016.7727096.
28. Cetinkaya A, Ishii H, Hayakawa T. An Overview on Denial-of-Service Attacks in Control Systems: Attack Models and Security Analyses. *Entropy*, 2019, 21:1-29. DOI:10.3390/E2102021029.
29. Wood AD, Stankovic JA. Denial of Service in Sensor Networks. *Computer,* 2002, 35(10):54-62. DOI: 10.1109/MC.2002.1039518.
30. Chifor B, Patriciu V. Mitigating DoS attacks in publish-subscribe IoT networks. *Proc. of Conf.: Electronics, Computers and Artificial Intelligence,* 2017, pp. 1-6. DOI: 10.1109/ECAI.2017.8166463.
31. Handosa M, Gracanin D. Performance evaluation of mqtt-based internet of things system. *Proc. of Winter Simulation Conference,* 2017, pp. 4544-4545. DOI: 10.1109/WSC.2017.8248196.
32. Fehrenbach P. Messaging Queues in the IoT Under Pressure-Stress Testing the Mosquitto MQTT Broker. *Fakultät Informatik. Hochschule Furtwangen University*, 2017. URL: https://blog.it-securityguard.com/wp-content/uploads/2017/10/IOT_Mosquitto_Pfehrenbach.pdf.
33. Firdous SN, Baig Z, Valli C, Ibrahim A. Modelling and Evaluation of Malicious Attacks against the IoT MQTT Protocol. *Proc. IEEE Intern.Conf. on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData)*, 2017, pp. 748-755. DOI: 10.1109/ITHINGS-GREENCOM-CPSCOM-SMARTDATA.2017.115.
34. Bao C, Guan X, Sheng Q, Zheng K, Huang X. A Tool for Denial of Service Attack Testing in IoT. *Proc. 8th Intern. Conf. on Information Technology in Medicine and Education (ITME)*, 2016, pp. 1-6.
35. Официальный сайт WEKA project. URL: https://www.cs.waikato.ac.nz/ml/weka/32 (дата обращения: 11.03.2020).
36. Дикий Д.И. Анализ протокола MQTT на атаки «отказ в обслуживании». Научно-технич. вестник информац. технологий, механики и оптики ИТМО. 2020, 2(2). DOI: 10.17586/2226.1494.2020.20.2.
37. Официальный сайт клиента paho-mqtt. URL: https://pypi.org/project/paho-mqtt/1.3.0/ (дата обращения: 20.08.2019).
38. Официальный сайт брокера Moquette. URL: https://projects.eclipse.org/projects/iot.moquette (дата обращения: 20.08.2019).
39. Hasan M, Islam M, Zarif I, Hashem M. Attack and anomaly detection in IoT sensors in IoT sites using machine learning approaches. *Internet of Things,* 2019, 7:1–14. DOI: 10.1016/J.IOT.2019.100059.

Дикий Дмитрий Игоревич
аспирант
Университет ИТМО
49, Кронверкский пр., СПб 197101, Россия
dimandikiy@mail.ru.