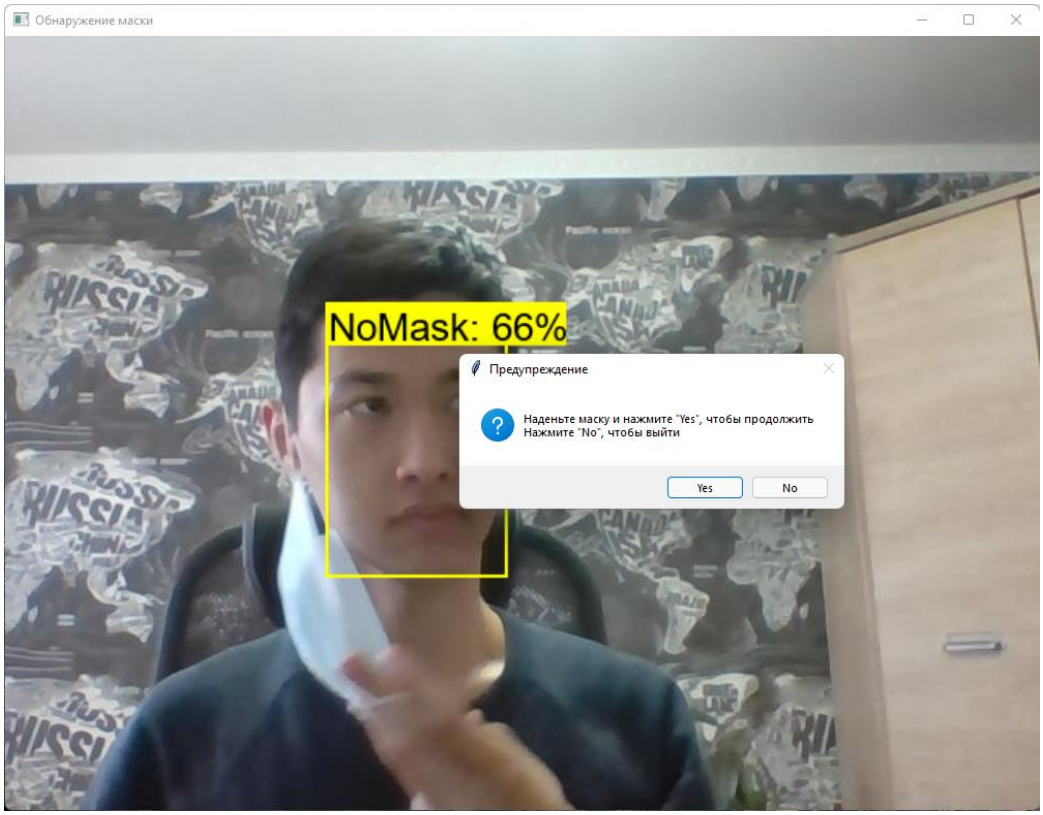

Рисунок 2 – Проверка обнаружения и предупреждения при отсутствии маски

Обучить свою собственную модель искусственного интеллекта и подстроить её под свои нужды в 2022 году значительно проще и доступнее, чем это было 10, или даже 5 лет назад. Созданные в данной области инструменты, автоматизируют многие аспекты разработки и обучения моделей, тем самым снижая порог входа в сферу машинного обучения.

Ссылка на разработанную программу: https://github.com/Meirbek-dev/face-mask_detector.

Список использованных источников

1. Джоши, Пратик. Искусственный интеллект с примерами на Python.: Пер. с англ. - СПб.: ООО "Диалектика", 2019. -448 с. -Парал. тит. англ.

2. OpenCV: OpenCV modules [Электронный ресурс]. – Режим доступа: https://docs.opencv.org/4.x/index.html

3. Жерон, Орельен. Прикладное машинное обучение с помощью Scikit-Learn и TensorFlow: концепции, инструменты и техники для создания интеллектуальных систем. Пер. с англ. - СПб.: ООО "Альфа-книга': 2018. - 688 с.: ил. - Парал. тит. англ.

4. Шакла Нишант. Машинное обучение и TensorFlow. - СПб.: Питер, 2019. - 336 с.: ил. - (Серия«Библиотека программиста»).

УДК 004.49, 004.93

ПРИМЕНЕНИЕ МАШИННОГО ОБУЧЕНИЯ ДЛЯ ДЕТЕКТИРОВАНИЯ РАСПРЕДЕЛЕННЫХ АТАК НА ОТКАЗ ОТ ОБСЛУЖИВАНИЯ

Бисенбаева Назерке Кобыландиевна
n.kobylandievna@gmail.com

Докторант 3-го курса специальности 8D06104 -Вычислительная техника и программное обеспечение, ЕНУ им. Л.Н.Гумилева, Нур-Султан, Казахстан

Нуржаубаев Акниет Алибиевич
nurzhaubaev.akniet@gmail.com
Магистрант 2-го курса ОП 7М06306 -Системы информационной безопасности,
ЕНУ им. Л.Н.Гумилева, Нур-Султан, Казахстан
Научный руководитель – Д.Сатыбалдина

Сетевые атаки типа «отказ в обслуживании» (Denial-of-service, DoS), которые бывают разных форм, представляют собой попытки заблокировать доступ законных пользователей к веб-серверу или онлайн сервису за счет переполнения их полосы пропускания [1]. Используя то, что сетевые ресурсы имеют ограничения по количеству запросов, которые они могут обслуживать одновременно, злоумышленники отправляют на атакуемую систему большое количество запросов с целью превысить её способность обрабатывать их все. Это приводит к значительному замедлению работы или полному отказу в обслуживании легитимных пользователей. Разновидностью подобных сетевых атак является распределенная атака типа «отказ в обслуживании» (Distributed Denial of Service, DDoS-атака), в которой на целевой вычислительный ресурс отправляется большое число запросов из сети зараженных компьютеров (ботнетов). Конечными целями подобных сетевых атак является полное прекращение работы веб-ресурса, которое может нарушить бесперебойное функционирование бизнес-процессов, привести к прямым финансовым потерям или дискредитировать имидж компании.

На рисунке 1 представлена инфографика по статистике инцидентов в 1-м квартале 2022 года в Республике Казахстан, представленная Службой реагирования на компьютерные инциденты KZ-CERT [2]. Как видно из рисунка 1, количество инцидентов, связанных с атакой типа «отказ в обслуживании» невелико, всего 42 из 5325. Однако, на порядок выше выявлено инцидентов «Отсутствие доступа к информационным ресурсам» (332) и «Ботнеты» (333), которые могут быть связаны с DoS- и DDoS-атаками. Таким образом, общее количество выявленных инцидентов, связанных с атаками типа «отказ в обслуживании» достаточное большое, к тому же реализация этих сетевых атак может приводить к значительным ущербам. Это приводит к необходимости как применять методы защиты против атак данного класса на практике, так и повышает актуальность научных исследований по разработке методов раннего выявления признаков DoS- и DDoS-атак, чтобы повысить степень оперативности реагирования на них и снизить ущербы.

Для защиты от DoS- и DDoS-атак используются разные подходы: фильтрация сетевого трафика, устранение уязвимостей сервера и программного обеспечения, наращивание ресурсов и рассредоточение посредством использования продублированных систем, которые продолжат обслуживать пользователей, внедрение систем обнаружения вторжений (Intrusion Detection System, IDS). Новые направления IDS направлены на использование моделей машинного обучения для разработки более надежных систем с более высоким уровнем обнаружения и более низким уровнем ложных срабатываний.

В настоящей работе представлены результаты исследований по выявлению распределенных атак на отказ от обслуживания на основе алгоритмов машинного обучения и их апробирования с использованием эталонного набора данных UNSW-NB15 [2].

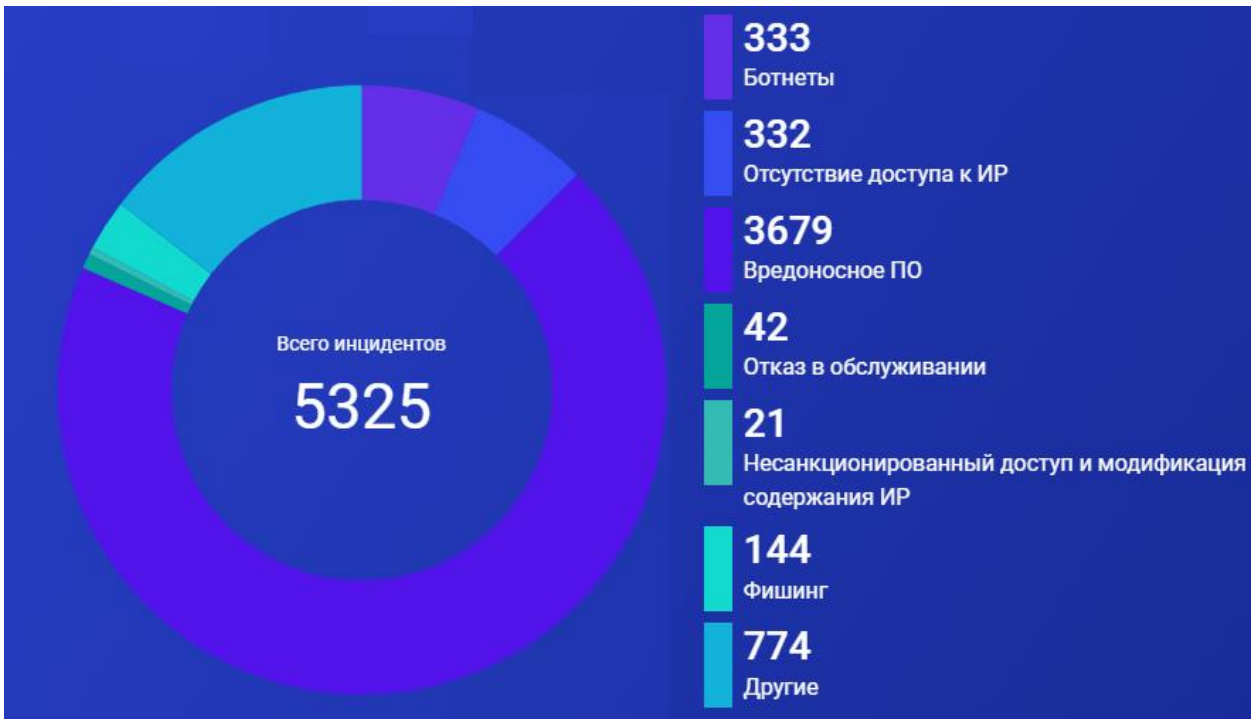

Рисунок 1 – Статистик инцидентов в 1-м квартале 2022 года в Республике Казахстан (https://cert.gov.kz/press_club/infographics).

База данных UNSW-NB15 была создана в лаборатории Cyber Range ACCS [3] с использованием генератора трафика IXIA PerfectStorm [4], который имитирует аномальный сетевой трафик, создаваемый 9-ю видами сетевых атак (см. Таблица 1).

Таблица 1. Распределение записей по категориям атак в базе данных UNSW-NB15

Типы трафиков/атак	Количество записей	Description
Нормальный трафик	2 218 761	естественные данные нормального трафика
Трафик, создаваемый программой-фаззером (Fuzzer)	24 246	попытка приостановить работу сети путем передачи случайно сгенерированных данных
Трафик, создаваемый программой-анализатором	2 677	трафик содержит различные атаки сканирования портов, html-файлов и проникновение спам
Бэкдоры (Backdoors)	2 329	техника, при которой механизм безопасности системы скрытно обходит для доступа к компьютеру или его данным
DoS -атаки	16 353	злонамеренная попытка сделать сервер или сетевой ресурс недоступным для пользователей, обычно путем временного прерывания или приостановки служб хоста, подключенного к Интернету
Эксплойты (Exploits)	44 525	злоумышленник знает о проблеме безопасности в операционной системе или части программного обеспечения и использует эти знания, используя уязвимость
Атака против шифров (Generic)	215 481	общий метод работает против всех блочных шифров (с заданным размером блока и ключа),

		без учета структуры блочного шифра
Разведка (Reconnaissance)	13 987	содержит все попытки проникновения в сеть, которые могут имитировать атаки со сбором информации
Двоичный вредоносный код (Shellcode)	1 511	небольшой фрагмент кода, используемый в качестве полезной нагрузки при эксплуатации уязвимости в программном обеспечении
Вредоносный код типа «Worms»	174	вредоносный код, который копирует себя, чтобы распространиться на другие компьютеры; часто использует компьютерную сеть для распространения, полагаясь на сбои безопасности на целевом компьютере для доступа к нему

Для захвата сетевого трафика в виде пакетов использовался инструмент tcpdump [5]. Период моделирования составлял 16 часов 22 января 2015 г. и 15 часов 17 февраля 2015 г. для захвата 100 ГБ. Далее каждый pcap файл был разбит на фрагменты по 1000 МБ с помощью инструмента tcpdump. Для создания из pcap-файлов CSV-файлов, содержащих ключевые характеристики как надежные признаки (атрибуты) для классификации нескольких типов атак), используются инструменты Argus [6] и Bro-IDS [7]. Ключевые характеристики базы данных UNSW-NB15 представляют собой 4 CSV-файла, содержащих данные для реального нормального трафика и аномального трафика, полученных в процессе имитации сетевых атак. Выделено 49 ключевых характеристик (атрибутов для классификации атак), которые разделены на несколько групп: основные (Basic), контентные (Content) и временные (Time). Пример основных характеристик приведен в Таблице 2.

Таблица 2. Примеры атрибутов потоков

№ атрибута	**Обозначение**	**Тип данных** (N – номинальный; I- целое число, F- число с плавающей запятой, T- метка времени; B- двоичный	**Описание**
1	srcip	N	Source IP address
2	sport	I	Source port number
3	dstip	N	Destination IP address
4	dsport	I	Destination port number
5	proto	N	Transaction protocol
48	attack_cat	N	Название категории атаки (Fuzzers, Analysis, Backdoors, DoS, Exploits, Generic, Reconnaissance, Shellcode, Worms) см. Таблицу 1
49	Label	B	0 для записи нормального трафика и1 для записи трафика при сетевой атакеs

Проведенный эксперимент представляет собой построение интеллектуальной системы классификации аномального и нормального трафика для выделенной категории атак типа «отказ в обслуживании». Для реализации экспериментов была использована платформа для анализа данных Splunk Enterprise с использованием расширения Machine Learning Toolkit [8].

Данная библиотека содержит более 30 распространенных алгоритмов машинного обучения с открытым исходным кодом на языке Python. Для задач классификации применяются следующие алгоритмы машинного обучения:

- дерево решений (a decision tree, DT);
- случайный лес (a *random forest, RF*);
- метод опорных векторов (a support vector machine, SVM).

В качестве обучающей и тестовой базы данных использован описанный выше набор данных UNSW-NB15: записи файла UNSW-NB15_1.csv использованы как обучающая выборка, а файла UNSW-NB15_2.csv – как тестовая. Схема эксперимента представлена на рисунке 2.

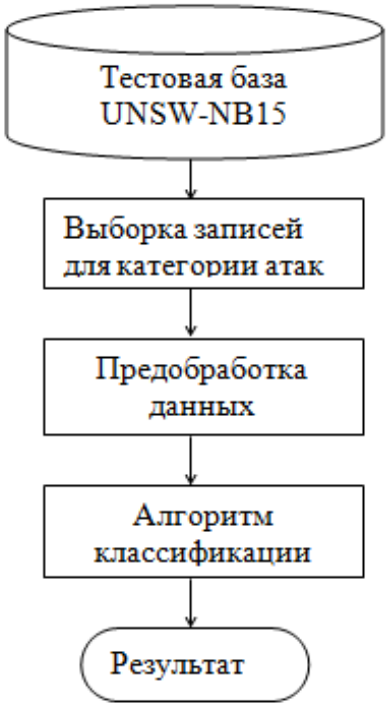

Рисунок 2. Схема эксперимента по классификации атак

с использованием трех алгоритмов машинного обучения

Выборка нужной категории атак осуществлялась посредством команды Splunk *«search»*: ***search attack_cat=Normal OR attack_cat=Dos.***

На этапе предобработки преобразуются нечисловые поля (строки или символы) в числовые, используя однократное кодирование, ввиду того что, алгоритмы машинного обучения работают с числовыми данными. Производится нормирование данных, в итоге данные записи сетевого трафика преобразуются в числовую матрицу из вещественных значений атрибутов, которые подаются на алгоритм машинного обучения.

В Таблице 3 представлены результаты тестирования обученной системы классификации DoS-атак, с указанием точности распознавания.

Таблица 3. Сравнение результатов классификации при использовании разных алгоритмов машинного обучения

Алгоритм классификации	Точность (Precision) (%)
Случайный лес (RF)	99,79
Дерево решений (DT)	99,65
Метод опорных векторов (SVM)	83, 705

Экспериментальные результаты показывают, что наилучшие результаты по детектированию Dos-атак по точности классификации и параметру ложного срабатывания дает алгоритм «случайный лес» по сравнению с другими использованными методами машинного обучения. В будущих работах планируется провести эксперимент с обученной моделью классификации сетевых атак, на основе реальных данных анализа сетевого трафика в программном комплексе для оперативного центра информационной безопасности, построенном на платформе Splunk Enterprise [9].

Список использованных источников

1. Zhijun W. et al. Low-rate DoS attacks, detection, defense, and challenges: a survey //IEEE Access. – 2020. – Т. 8. – С. 43920-43943.
2. https://cert.gov.kz
3. Moustafa N., Slay J. UNSW-NB15: a comprehensive data set for Network intrusion detection systems (UNSW-NB15 network data set) //2015 military communications and information systems conference (MilCIS). – IEEE, 2015. – С. 1-6.
4. http://www.accs.unsw.adfa.edu.au/
5. http://www.ixiacom.com/products/perfectstorm
6. http://www.tcpdump.org/
7. http://qosient.com/argus/index.shtml
8. Splunk. About the ML-SPL API. 2021. Available online: https://docs.splunk.com/Documentation/MLApp/5.3.0/User/AboutMLTK/ (Accessed:2022-03-23)
9. Абдиев Б.С., Сатыбалдина Д.Ж., Бисенбаева Н.К. Программный комплекс для оперативного центра информационной безопасности // Свидетельство о государственной регистрации прав на программу для ЭВМ, №10063 от 21.05.2020 г.

ӘОК 004.42; 004.43; 004.6; 004.7

КОМПЬЮТЕРЛІК ЖЕЛІДЕ АҚПАРАТТЫҚ ҚАУІПСІЗДІКТІ ҚАМТАМАСЫЗ ЕТУГЕ АРНАЛҒАН ЖЕЛІЛІК БАРЛАУ ҚҰРАЛДАРЫ

Ешмұхаметов Е.Н., Күлімханов Ә.
yeshmukhametov.yn@gmail.com
Л.Н.Гумилев атындағы Еуразия ұлттық университеті, Ақпараттық қауіпсіздік жүйелері, Нұр-Сұлтан, Қазақстан
Ғылыми жетекшісі – ф.-м.ғ.к., доцент м.а. А. Оспанова

Аңдатпа: Бұл мақалада кәсіпорындағы компьютерлік желілерді қорғауға арналған желілік барлау құралдары қарастырылып, талданады, осы құралдардың негізінде олардың тиімді қолданылуы туралы қорытынды жасалады. Күрделі хакерлік шабуылдарға қарсы тұру үшін вирусқа қарсы қорғау құралдарын, желіаралық экрандарды және басып кіруді болдырмау жүйесін қолдану жеткіліксіз, ең заманауи тәсілдерді ескеретін осалдықты анықтау үшін технологияларды пайдалану қажет.

Тақырыптың өзектілігі мен маңыздылығы ақпараттық технологиялардың кеңінен таралуына, қазіргі қоғамдағы және бизнестегі ақпараттың маңыздылығына және