УДК 004.056 EDN: ITWVQT

# МОДЕЛЬ ПРЕДСТАВЛЕНИЯ ВХОДНОГО НАБОРА СИГНАТУР В ВИДЕ СОКРАЩЕННОЙ ДИАГРАММЫ РЕШЕНИЙ

**Л.Я. Добкач** dobkachleo@mail.ru
**В.Л. Цирлов** v.tsirlov@bmstu.ru

**МГТУ им. Н.Э. Баумана, Москва, Российская Федерация**

**Аннотация**

Формат представления входных данных для последующих операций анализа или обучения алгоритмов распознавания компьютерных атак может оказывать влияние на затраты памяти, производительности и времени при осуществлении указанных процессов. Чаще всего входные данные представляют собой таблицы значений или наборы булевых правил. При этом значения последовательностей из нескольких параметров могут повторяться. Чтобы снизить объем хранимой информации и затраты времени на их обработку, в качестве модели представления входных данных предлагается использовать сокращенную диаграмму решений. Она позволяет не только снизить затраты памяти и повысить быстродействие, но и при обычном сигнатурном анализе обеспечивает прирост точности распознавания атак на 2 %. Несмотря на то что прирост незначительный, он свидетельствует о возможности сокращенной диаграммы решений усилить способность сигнатурного метода, не обладающего свойством гибкости (адаптивности), к распознаванию незнакомых компьютерных атак. В контексте машинного обучения предлагаемая модель представления данных может помочь сократить период повторного обучения или обновления алгоритмов интеллектуальной обработки данных и обеспечить более адекватное реагирование на новые сценарии попыток вторжений

**Ключевые слова**

*Сигнатурный анализ, машинное обучение, CICIDS 2017, обнаружение вторжений, деревья решений*

Поступила 13.06.2023
Принята 03.07.2023


**Введение.** Методы распознавания компьютерных атак предполагают сопоставление входящего сетевого трафика с наборами сигнатур, правил, шаблонов, профилей поведения и других сценариев злонамеренного вмешательства [1]. В настоящей работе рассмотрена проблема представления

входных данных для обучения алгоритмов интеллектуальной обработки данных (искусственных нейронных сетей [2, 3], методов машинного обучения [4, 5]) на основе сигнатур с известными метками классов. Она заключается в затратах времени и памяти на обучение классификаторов, влияет на их производительность. Простой перебор возможных векторов атак в сигнатурном анализе не только может быть сопряжен со снижением реакции на действительную угрозу [6], но и с риском злоумышленников провести атаку так, чтобы она не соответствовала известным сигнатурам и длилась меньше времени, которое потратит классификатор на сопоставление параметров [7].

**Материалы и методы решения задач, принятые допущения.** Сигнатуры могут быть представлены наборами правил, векторами параметров, уникальными величинами и т. п. [8]. Тем не менее в большинстве случаев такое описание применимо не столько к их реализации в том или ином методе распознавания, сколько к их представлению в понятном для человека (пользователя или специалиста по защите информации в общем случае) виде. На прикладном уровне сигнатуры также можно рассматривать в виде совокупности булевых правил (сравнений параметров с определенными значениями; пример такого подхода описан в [9]).

Сигнатуры имеют различные выводы. Во введении в неявном виде смешались понятия распознавания и классификации компьютерных атак. Следует отметить, что здесь под распознаванием компьютерной атаки имеется в виду не только выявление некой аномальной деятельности в нормальном трафике, но и установление класса этой аномалии.

Для некоторых задач достаточно отделить нормальные события от аномальных [6], т. е. решить достаточно простую булеву задачу: например, 0 может означать нормальное (безопасное) событие, 1 — аномальное (потенциально опасное).

Можно распознавать и более глубоко: присваивать каждому аномальному событию вид атаки. Таковых можно выделить десятки [10], что, хотя и представляет собой счетное конечное множество, все же значительно усложняет процесс классификации.

Здесь будем придерживаться среднего уровня классификации, т. е. под распознаванием компьютерных атак будем понимать присвоение класса события из множества $S = \{S_1, S_2, \ldots, S_5\}$, где $S_1$ — класс нормальных событий; $S_2 - S_5$ — классы аномальных событий. Такая классификация представляет собой упрощенное представление множества классов относительно современного набора данных CICIDS 2017 [11], в котором

выделяются 15 типов: один класс нормальных событий (BENIGN) и 14 типов атак [12].

Выделяемые в наборе CICIDS 2017 типы атак разбиваем на четыре семейства, которые будем именовать классами:

1) DOS (Denial of Service) — атаки типа «отказ в обслуживании» (включают в себя DDoS, DoS Hulk, DoS slowloris, DoS Slowhttptest и DoS GoldenEye [13, 14] и Heartbleed [15]);

2) Web Attack — сетевые (удаленные) атаки (типы FTP-Patator, SSH-Patator, Bot, Brute Force, XSS, Sql Injection);

3) Infiltration — атаки, направленные на повышение привилегий [16];

4) PortScan — зондирующие атаки, или атаки сканирования (портов, узлов и т. п. [17]).

Последние два класса содержат по одному типу атак.

**Результаты.** Независимо от того, как именно представлены сигнатуры, каждая из них содержит метку для проверки успешности обучения. При этом метка не считается неотъемлемой частью сигнатуры. По сути, это ожидаемый вывод классификатора, следовательно, когда речь идет о размерности набора параметров, событии сетевого трафика, сигнатуре, имеется в виду размерность без учета этой метки.

Введем размерность $Q'$ для вектора сигнатуры. Для CICIDS 2017 она равна 85, однако целесообразно подготовить набор данных для последующих операций. Не меняя на текущем этапе форму представления входных данных, удаляем из набора столбцы с одинаковыми значениями, дублирующие столбцы и проводим линейную комбинацию однотипных параметров с ограниченным множеством значений [18]. Далее векторы подвергаются снижению размерности и нормализации [19]. Без значительных потерь полезной информации размерность можно сократить до 12 параметров [20].

Всего в той или иной выборке может быть $n$ сигнатур (для CICIDS 2017 значение $n$ превышает 2,5 млн записей). Отсюда следует, что простейший способ представления базы сигнатур — табличный. Очевидно, тогда требуется хранить в памяти запоминающего устройства не менее $Q' \times n$ элементов. Хотя $Q' \ll n$, нельзя пренебречь размерностью.

Если исключить из рассмотрения тот факт, что размерность можно снизить, то все равно остается большое число элементов с равными или близкими значениями. Поскольку наборы данных формировались без необходимости выбирать только максимально уникальные сигнатуры, многие из них можно назвать похожими, а следовательно, они имеют расхождение всего в нескольких параметрах.

Изложенное позволяет предложить такую модель представления, в которой некоторые строки таблицы могут быть объединены в тех случаях, когда, начиная с какого-то столбца, все значения оказываются одинаковыми, т. е. каждый вектор сигнатуры можно представить как неразветвленное дерево, а при слиянии дублей независимые ветви образуют новые, более сложные деревья решений. В общем виде получается граф (рис. 1):

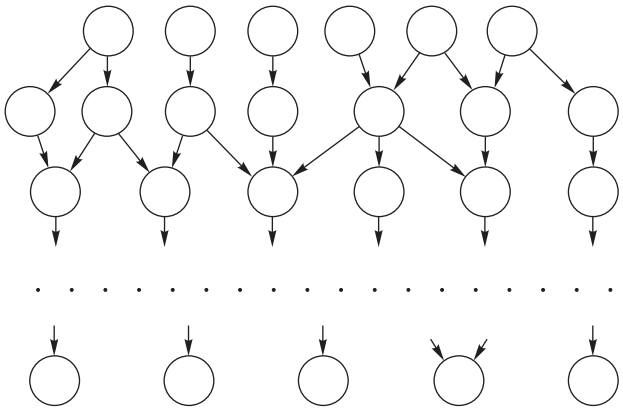

**Рис. 1.** Графическое представление структуры сокращенной диаграммы решений

Такая модель ориентированного ациклического графа, узлами которого выступают булевы правила для отдельных признаков сигнатуры события сетевого трафика, соответствует так называемой сокращенной диаграмме решений [21]. Она позволяет не только значительно сократить число повторяющихся данных, но и оптимизировать ее структуру для обучения искусственных нейронных сетей и алгоритмов машинного обучения.

Сокращенная диаграмма решений реализуется путем последовательного сравнения значений параметров между собой на двух соседних уровнях. Если значение одного из параметров отличается от другого более чем на порядок (более чем в 10 раз), такие пары параметров образуют отдельные ветви, иначе они объединяются и начинается просмотр следующей пары параметров, один из параметров которой задействован в предыдущей паре.

Затраты времени, памяти и производительности не влияют на работу будущего программного продукта, так как преобразование базы сигнатур в сокращенную диаграмму решений происходит на предварительном этапе до начала обучения и тестирования системы обнаружения вторжений [22]. Следует отметить, что при необходимости переобучения систе-

мы на актуальных данных в практических условиях можно добиться сокращения переходного периода на обновленную версию [23].

Соединение ветвей по принципу отличия не более чем на порядок приводит к тому, что в базе сигнатур в формате сокращенной диаграммы решений возникают новые меченые события, причем метки присваиваются на основании пограничных значений. Это означает, что такое представление позволяет, с одной стороны, определять слегка видоизмененные атаки, чье отличие от изначально известных аналогично погрешности вычислений. С другой стороны, новые, неявно заданные сигнатуры могут по каким-то причинам не соответствовать истинному классу события, но это можно полагать лишь поводом для углубленного анализа и добавления в базу сигнатур таких событий, которые выбиваются из описания, с помощью сокращенной диаграммы решений.

Однако некоторые новообразованные с помощью такой модели сигнатуры не столько выбиваются из вычисленных рамок, сколько оказались в них ненамеренно. Другими словами, на каком-то уровне несколько событий могут объединиться, но потом ввиду сильно различающихся (более чем на порядок) значений могут образоваться отдельные ветви и в силу связности графа из первой части сигнатуры можно попасть во вторую часть сигнатуры. Даже если они имеют одинаковые метки, то в этот сегмент дерева могут попасть совершенно неподходящие ветви сигнатур, которые назовем паразитными [24].

Для борьбы с паразитными ветвями установили ограничение на возможность объединения различных сигнатур. Таким образом, схема сокращенной диаграммы решений принимает вид, показанный на рис. 2. По горизонтали отложено число $n$ сигнатур размерностью $m \le Q'$.

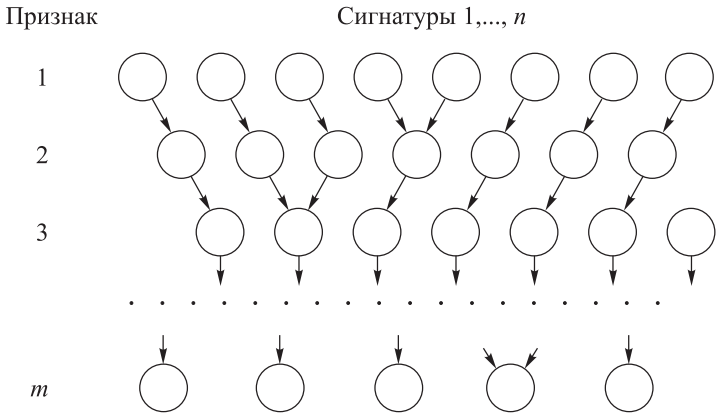


**Рис. 2.** Сокращенная диаграмма решений без паразитных связей

Сигнатуры образуют теперь менее связанный лес в виде сокращенной диаграммы решений, сводящийся к пяти меткам и их копиям. Это увеличивает затраты памяти, но в силу $Q' \ll n$ не настолько значительно, зато позволяет сохранить эффект, заключающийся в повышении точности распознавания атак исключительно сигнатурным методом. Сокращенная диаграмма решений отчасти может противостоять попытке злоумышленника, осведомленного об используемой в защитной системе базе сигнатур, создать такую атаку, которая бы обходила все содержащиеся там векторы событий.

**Обсуждение полученных результатов.** Модель представления наборов сигнатур в виде сокращенной диаграммы решений позволяет, во-первых, значительно сократить число уникальных и фактически повторяющихся значений параметров (почти на 55 %), во-вторых, снизить затраты памяти (приблизительно на 22 %); в-третьих, повысить точность распознавания событий безопасности минимум на 2 % даже без внедрения адаптивных методов классификации, как показали результаты тестирования на наборе CICIDS 2017.

Хотя повышение точности на 2 % кажется незначительным, оно свидетельствует о том, что одно только представление набора входных данных в виде сокращенной диаграммы решений позволяет улучшить возможности применения сигнатурного анализа.

**Заключение.** Предложена модель представления базы сигнатур событий безопасности в виде сокращенной диаграммы решений. По сравнению с табличным представлением такая модель позволяет уменьшить затраты памяти и, следовательно, затраты времени на обработку меньшего объема входных данных, а также немного повысить (на 2 %) точность выявления компьютерных атак даже без применения адаптивных классификаторов. В то же время модель пригодна как для обучения искусственных нейронных сетей и алгоритмов машинного обучения, так и для проверки (тестирования) ансамблей этих классификаторов.

## ЛИТЕРАТУРА

[1] Добкач Л.Я. Обзор методов распознавания компьютерных атак. *Безопасность информационных технологий. Сб. тр. Десятой междунар. науч.-тех. конф.* М., Изд-во МГТУ им. Н.Э. Баумана, 2019, с. 124–129.

[2] Евглевская Н.В., Ракицкий С.Н. Выбор метода обнаружения компьютерных атак. *Известия ТулГУ. Технические науки*, 2021, № 5, с. 247–249.

[3] Королев И.Д., Попов В.И., Рева Д.И. Обзор методов прогнозирования целенаправленных угроз информационной безопасности. *Информационная безопасность: вчера, сегодня, завтра. III Междунар. науч.-практ. конф.* М., РГГУ, 2020, с. 163–170.

[4] Liu H., Lang B. Machine learning and deep learning methods for intrusion detection systems: a survey. *Appl. Sci.*, 2019, vol. 9, iss. 20, pp. 4396–4423. DOI: https://doi.org/10.3390/app9204396

[5] Шелухин О.И., Раковский Д.И. Бинарная классификация многоатрибутных размеченных аномальных событий компьютерных систем с помощью алгоритма SVDD. *Наукоемкие технологии в космических исследованиях Земли*, 2021, т. 13, № 2, с. 74–84. DOI: https://doi.org/10.36724/2409-5419-2021-13-2-74-84

[6] Кусакина Н.М. Методы анализа сетевого трафика как основа проектирования системы обнаружения сетевых атак. *International Scientific Review of the Problems and Prospects of Modern Science and Education. XLI Междунар. науч.-практ. конф.* Иваново, Проблемы науки, 2018, с. 28–31.

[7] Оралбаев Е.А. Обнаружения DDoS-атак ботнетов в сетях доступа IoT. В кн.: *Актуальные вопросы современной науки и образования*. Пенза, Наука и Просвещение, 2021, с. 190–200.

[8] Акушуев Р.Т. Модель обнаружения сигнатур. *Modern Science*, 2020, № 7-1, с. 330–332.

[9] Абрамов Е.С., Тарасов И.В. Применение комбинированного нейросетевого метода для обнаружения низкоинтенсивных DDoS-атак на Web-серсисы. *Инженерный вестник Дона*, 2017, № 3, ст. 59. URL: http://vww.ivdon.ru/ru/magazine/archive/N3y2017/4354

[10] Кузьмин В.Н., Менисов А.Б. Исследование путей и способов повышения результативности выявления компьютерных атак на объекты критической информационной инфраструктуры. *Информационно-управляющие системы*, 2022, № 4, с. 29–43. DOI: https://doi.org/10.31799/1684-8853-2022-4-29-43

[11] Васильев И.Н. Исследование методов обнаружения и устранения киберугроз для корпоративных сетей. *ОРВСЭУ–2022*. Переславль-Залесский, ИПС РАН, 2022, с. 92–99.

[12] Ландызин А.Н., Шелухин О.И. Методика предварительной обработки набора данных для бинарной и многоклассовой классификации атак. *Телекоммуникации и информационные технологии*, 2022, т. 9, № 1, с. 46–57.

[13] Yin Y., Jang-Jaccard J., Sabrina F., et al. Improving multilayer-perceptron (MLP)-based network anomaly detection with birch clustering on CICIDS-2017 dataset. *arXiv:2208.09711*. DOI: https://doi.org/10.48550/arXiv.2208.09711

[14] Al-Harbi A., Jabeur R. An efficient method for detection of DDoS attacks on the web using deep learning algorithms. *IJATCSE*, 2021, vol. 10, no. 4, pp. 2821–2829. DOI: https://doi.org/10.30534/ijatcse/2021/271042021

[15] Abdulraheem M.H., Ibraheem N.B. A detailed analysis of new intrusion detection dataset. *J. Theor. Appl. Inf. Technol*, 2019, vol. 97, no. 17, pp. 4519–4537.

[16] Stiawan D., Idris M.Y.B., Bamhdi A.M., et al. CICIDS-2017 dataset feature analysis with information gain for anomaly detection. *IEEE Access*, 2020, vol. 8, pp. 132911–132921. DOI: https://doi.org/10.1109/ACCESS.2020.3009843

[17] Güven E.Y., Gülgün S., Manav C., et al. Multiple classification of cyber attacks using machine learning. *Electrica*, 2022, vol. 22, iss. 2, pp. 313–320.
DOI: https://doi.org/10.54614/electrica.2022.22031

[18] Добкач Л.Я. Создание модуля распознавания атак для систем обнаружения вторжений. *Всерос. студ. конф. «Студенческая научная весна»*. М., Научная библиотека, 2019, с. 36–37.

[19] Сакулин С.А., Алфимцев А.Н., Квитченко К.Н. и др. Выявление аномалий сетевого трафика с использованием ансамбля классификаторов. *Вестник компьютерных и информационных технологий*, 2020, т. 17, № 10, с. 38–46.
DOI: https://doi.org/10.14489/vkit.2020.10.pp.038-046

[20] Abdulhammed R., Faezipour M., Musafer H., et al. Efficient network intrusion detection using PCA-based dimensionality reduction of features. *2019 IEEE ISNCC*, 2019.
DOI: https://doi.org/10.1109/ISNCC.2019.8909140

[21] Бибило П.Н., Романов В.И. Минимизация многоуровневых представлений систем полностью определенных булевых функций с использованием разложений Шеннона и алгебраических представлений кофакторов. *Информатика*, 2021, т. 18, № 2, с. 7–32. DOI: https://doi.org/10.37661/1816-0301-2021-18-2-7-32

[22] Леонтьев А.Л. Использование теории графов и нейронных сетей для обнаружения уязвимости информационно-технологических систем. *Фундаментальные проблемы информационной безопасности в условиях цифр. Трансформации. II Всерос. науч. конф.* Ставрополь, СКФУ, 2020, с. 207–213.

[23] Лебедев И.С. Адаптивное применение моделей машинного обучения на отдельных сегментах выборки в задачах регрессии и классификации. *Информационно-управляющие системы*, 2022, № 3, с. 20–30.
DOI: https://doi.org/10.31799/1684-8853-2022-3-20-30

[24] Rauzy A., Yang L. Decision diagram algorithms to extract minimal cutsets of finite degradation models. *Information*, 2019, vol. 10, iss. 12, pp. 368–395.
DOI: https://doi.org/10.3390/info10120368

**Добкач Леонид Яковлевич** — аспирант кафедры «Информационная безопасность» МГТУ им. Н.Э. Баумана (Российская Федерация, 105005, Москва, 2-я Бауманская ул., д. 5, стр. 1).

**Цирлов Валентин Леонидович** — канд. техн. наук, доцент кафедры «Информационная безопасность» МГТУ им. Н.Э. Баумана (Российская Федерация, 105005, Москва, 2-я Бауманская ул., д. 5, стр. 1).

**Просьба ссылаться на эту статью следующим образом:**

Добкач Л.Я., Цирлов В.Л. Модель представления входного набора сигнатур в виде сокращенной диаграммы решений. *Вестник МГТУ им. Н.Э. Баумана. Сер. Приборостроение*, 2024, № 1 (146), с. 93–103. EDN: ITWVQT

# MODEL FOR PRESENTING THE INPUT SET OF SIGNATURES IN THE FORM OF A REDUCED DECISION DIAGRAM

**L.Ya. Dobkach** dobkachleo@mail.ru
**V.L. Tsirlov** v.tsirlov@bmstu.ru

**Bauman Moscow State Technical University, Moscow, Russian Federation**

**Abstract**

The format for presenting the input data for subsequent analysis operations or learning algorithms in identifying the computer attacks could affect the memory cost, performance and time within these processes. Most often, the input data appears to be tables of values or sets of the Boolean rules. In this case, the certain parameter values could be repeated. To reduce the amount of stored information and time spent on its processing, the paper proposes to use a reduced decision diagram as the model representing the input data. It makes it possible to not only reduce the memory costs and increase performance, but also provides a 2 % increase in the attack recognition accuracy at the conventional signature analysis. Despite the fact that the increase is insignificant, it indicates a possibility of the reduced decision diagram to enhance the signature method ability not having the flexibility (adaptability) property to identify the unfamiliar computer attacks. In the machine-learning context, the proposed data representation model is able to assist in reducing the retraining period or updating the data mining algorithms, and to provide a more adequate response to renewed scenarios of the intrusion attempts

**Keywords**

*Signature analysis, machine learning, CICIDS 2017, intrusion detection, decision trees*

Received 13.06.2023
Accepted 03.07.2023


## REFERENCES

[1] Dobkach L.Ya. [The review of methods for identification of computer attacks]. *Bezopasnost informatsionnykh tekhnologiy. Sb. tr. Desyatoy mezhdunar. nauch.-tekh. konf.* [Safety of Information Technologies. Proc. 10th Int. Sci.-Tech. Conf.]. Moscow, BMSTU Publ., 2019, pp. 124–129 (in Russ.).

[2] Evglevskaya N.V., Rakitskiy S.N. Choice of computer attacks detection method. *Izvestiya TulGU. Tekhnicheskie nauki* [News of the Tula State University. Technical Sciences], 2021, no. 5, pp. 247–249 (in Russ.).

[3] Korolev I.D., Popov V.I., Reva D.I. [Overview of methods targeted threat's forecasting in cyber security]. *Informatsionnaya bezopasnost: vchera, segodnya, zavtra. III Mezhdunar. nauch.-prakt. konf.* [Information Safety: Yesterday, Today, Tomorrow. III Int. Sci.-Pract. Conf.]. Moscow, RSUH, 2020, pp. 163–170 (in Russ.).

[4] Liu H., Lang B. Machine learning and deep learning methods for intrusion detection systems: a survey. *Appl. Sci.*, 2019, vol. 9, iss. 20, pp. 4396–4423.
DOI: https://doi.org/10.3390/app9204396

[5] Shelukhin O.I., Rakovskiy D.I. Binary classification of multi-attribute tagged data about anomalous events in computer systems using the SVDD algorithm. *Naukoemkie tekhnologii v kosmicheskikh issledovaniyakh Zemli* [High Technologies in Earth Space Research], 2021, vol. 13, no. 2, pp. 74–84 (in Russ.).
DOI: https://doi.org/10.36724/2409-5419-2021-13-2-74-84

[6] Kusakina N.M. [Network traffic analysis methods as a basis for designing a network attack detection system]. *International Scientific Review of the Problems and Prospects of Modern Science and Education. XLI Mezhdunar. nauch.-prakt. konf.* [XLI Int. Sci.-Pract. Conf.]. Ivanovo, Problemy nauki, 2018, pp. 28–31 (in Russ.).

[7] Oralbaev E.A. Obnaruzheniya DDoS-atak botnetov v setyakh dostupa IoT. V kn.: *Aktualnye voprosy sovremennoy nauki i obrazovaniya* [Detecting DDoS attacks of botnets in IoT access networks. In: Topical Issues of Modern Science and Education]. Penza, Nauka i Prosveshchenie Publ., 2021, pp. 190–200 (in Russ.).

[8] Akushuev R.T. Signature detection model. *Modern Science*, 2020, no. 7-1, pp. 330–332 (in Russ.).

[9] Abramov E.S., Tarasov I.V. Application of the combined neural network method for web-oriented low-rate DDOS-attack detection. *Inzhenernyy vestnik Dona* [Engineering Journal of Don], 2017, no. 3, art. 59 (in Russ.).
Available at: http://vww.ivdon.ru/ru/magazine/archive/N3y2017/4354

[10] Kuzmin V.N., Menisov A.B. A study of ways and solutions to increase the efficiency of detecting computer attacks on the objects of critical information infrastructure. *Informatsionno-upravlyayushchie sistemy* [Information and Control Systems], 2022, no. 4, pp. 29–43 (in Russ.). DOI: https://doi.org/10.31799/1684-8853-2022-4-29-43

[11] Vasilyev I.N. [Study of cyber threat detection and remediation techniques for enterprise networks]. *ORVSEU–2022*. Pereslavl-Zalesskiy, IPS RAS Publ., 2022, pp. 92–99 (in Russ.).

[12] Landyzin A.N., Shelukhin O.I. Data set pre-processing technique for binary and multiclass attack classification. *Telekommunikatsii i informatsionnye tekhnologii*, 2022, vol. 9, no. 1, pp. 46–57 (in Russ.).

[13] Yin Y., Jang-Jaccard J., Sabrina F., et al. Improving multilayer-perceptron (MLP)-based network anomaly detection with birch clustering on CICIDS-2017 dataset. *arXiv:2208.09711*. DOI: https://doi.org/10.48550/arXiv.2208.09711

[14] Al-Harbi A., Jabeur R. An efficient method for detection of DDoS attacks on the web using deep learning algorithms. *IJATCSE*, 2021, vol. 10, no. 4, pp. 2821–2829.
DOI: https://doi.org/10.30534/ijatcse/2021/271042021

[15] Abdulraheem M.H., Ibraheem N.B. A detailed analysis of new intrusion detection dataset. *J. Theor. Appl. Inf. Technol*, 2019, vol. 97, no. 17, pp. 4519–4537.

[16] Stiawan D., Idris M.Y.B., Bamhdi A.M., et al. CICIDS-2017 dataset feature analysis with information gain for anomaly detection. *IEEE Access*, 2020, vol. 8, pp. 132911–132921. DOI: https://doi.org/10.1109/ACCESS.2020.3009843

[17] Güven E.Y., Gülgün S., Manav C., et al. Multiple classification of cyber attacks using machine learning. *Electrica*, 2022, vol. 22, iss. 2, pp. 313–320.
DOI: https://doi.org/10.54614/electrica.2022.22031

[18] Dobkach L.Ya. [Creation of an attack recognition module for intrusion detection systems]. *Vseros. stud. konf. "Studencheskaya nauchnaya vesna"* [Russ. Stud. Conf. Student Scientific Spring]. Moscow, Nauchnaya biblioteka Publ., 2019, pp. 36–37 (in Russ.).

[19] Sakulin S.A., Alfimtsev A.N., Kvitchenko K.N., et al. Network traffic anomalies detection using an ensemble of classifiers. *Vestnik kompyuternykh i informatsionnykh tekhnologiy* [Herald of Computer and Information Technologies], 2020, vol. 17, no. 10, pp. 38–46 (in Russ.). DOI: https://doi.org/10.14489/vkit.2020.10.pp.038-046

[20] Abdulhammed R., Faezipour M., Musafer H., et al. Efficient network intrusion detection using PCA-based dimensionality reduction of features. *2019 IEEE ISNCC*, 2019. DOI: https://doi.org/10.1109/ISNCC.2019.8909140

[21] Bibilo P.N., Romanov V.I. Minimization of binary decision diagrams for systems of completely defined Boolean functions using Shannon expansions and algebraic representations of cofactors. *Informatika* [Informatics], 2021, vol. 18, no. 2, pp. 7–32 (in Russ.). DOI: https://doi.org/10.37661/1816-0301-2021-18-2-7-32

[22] Leontyev A.L. [Using graph theory and neural networks to detect vulnerability of information technology systems]. *Fundamentalnye problemy informatsionnoy bezopasnosti v usloviyakh tsifr. Transformatsii. II Vseros. nauch. konf.* [Fundamental Problems of Information Security in Digital Environment. Transformations. II Russ. Sci. Conf.]. Stavropol, SKFU Publ., 2020, pp. 207–213 (in Russ.).

[23] Lebedev I.S. Adaptive application of machine learning models on separate segments of a data sample in regression and classification problems. *Informatsionno-upravlyayushchie sistemy* [Information and Control Systems], 2022, no. 3, pp. 20–30 (in Russ.). DOI: https://doi.org/10.31799/1684-8853-2022-3-20-30

[24] Rauzy A., Yang L. Decision diagram algorithms to extract minimal cutsets of finite degradation models. *Information*, 2019, vol. 10, iss. 12, pp. 368–395.
DOI: https://doi.org/10.3390/info10120368

**Dobkach L.Ya.** — Post-Graduate Student, Department of Information Security, Bauman Moscow State Technical University (2-ya Baumanskaya ul. 5, str. 1, Moscow, 105005 Russian Federation).

**Tsirlov V.L.** — Cand. Sc. (Eng.), Assoc. Professor, Department of Information Security, Bauman Moscow State Technical University (2-ya Baumanskaya ul. 5, str. 1, Moscow, 105005 Russian Federation).

**Please cite this article in English as:**

Dobkach L.Ya., Tsirlov V.L. Model for presenting the input set of signatures in the form of a reduced decision diagram. *Herald of the Bauman Moscow State Technical University, Series Instrument Engineering*, 2024, no. 1 (146), pp. 93–103 (in Russ.). EDN: ITWVQT