10. *Shelukhin O.I., Ryabinin V.S., Farmakovskiy M.A.* Obnaruzhenie anomal'nykh sostoyaniy komp'yuternykh sistem sredstvami intellektual'nogo analiza dannykh sistemnykh zhurnalov [Detection of abnormal states of computer systems by means of intelligent analysis of system log data], *Voprosy kiberbezopasnosti* [Cybersecurity Issues], 2018, No. 2 (26), pp. 33-43. DOI: 10.21681/2311-3456-2018-2-33-43. EDN XYHQUP.
11. *Slipenchuk P.V.* Algoritm izvlecheniya kharakternykh priznakov iz dannykh pol'zovatel'skikh aktivnostey [Algorithm for extracting characteristic features from user activity data], Voprosy kiberbezopasnosti [Cybersecurity Issues], 2019, No. 1 (29), pp. 53-58. DOI: 10.21681/2311-3456-2019-1-53-58. EDN YZFWPZ.
12. *Do E.H. and Gadepally V.N*. Classifying Anomalies for Network Security, *ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Barcelona, Spain, 2020,* pp. 2907-2911. DOI: 10.1109/ICASSP40776.2020.9053419.
13. *Wu J., Gan W., Chen Z., Wan S. and Yu P.S*. Multimodal Large Language Models: A Survey, *2023 IEEE International Conference on Big Data (BigData), Sorrento, Italy, 2023,* pp. 2247-2256. DOI: 10.1109/BigData59044.2023.10386743.
14. *Shi S., Han D., & Cui M.* A multimodal hybrid parallel network intrusion detection model, *Connection Science,* 2023, 35 (1). Available at: https://doi.org/10.1080/09540091.2023.2227780.
15. *Ullah F., Turab A., Ullah S., Cacciagrano D., Zhao Y.* Enhanced Network Intrusion Detection System for Internet of Things Security Using Multimodal Big Data Representation with Transfer Learning and Game Theory, *Sensors,* 2024, 24 (13):4152. Available at: https://doi.org/10.3390/s24134152.
16. *Singh A.K., Krishnan S.* ECG signal feature extraction trends in methods and applications, *BioMed Eng OnLine,* 2023, 22. Available at: https://doi.org/10.1186/s12938-023-01075-1.
17. *Kotenko I., Saenko I., Lauta O., Kribel A.* An approach to detecting cyber attacks against smart power grids based on the analysis of network traffic self-similarity, *Energies,* 2020, Vol. 13, No. 19, pp. 5031. DOI: 10.3390/en13195031. EDN YVERBA.
18. *Get'man A.I., Goryunov M.N., Matskevich A.G. [i dr.].* Primenenie glubokogo obucheniya dlya obnaruzheniya komp'yuternykh atak v setevom trafike [Application of deep learning to detect computer attacks in network traffic], *Tr. Instituta sistemnogo programmirovaniya RAN* [Proceedings of the Institute for System Programming of the Russian Academy of Sciences], 2023, Vol. 35, No. 4, pp. 65-92. DOI: 10.15514/ISPRAS-2023-35(4)-3. EDN CSLHAE.
19. *Jogin M., Mohana, Madhulika M.S., Divya G.D., Meghana R.K. and Apoorva S.* Feature Extraction using Convolution Neural Networks (CNN) and Deep Learning, *2018 3rd IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT), Bangalore, India, 2018,* pp. 2319-2323. DOI: 10.1109/RTEICT42901.2018.9012507.
20. *Xiao Y., Xing C., Zhang T. and Zhao Z.* An Intrusion Detection Model Based on Feature Reduction and Convolutional Neural Networks, *in IEEE Access,* 2019, Vol. 7, pp. 42210-42219. DOI: 10.1109/ACCESS.2019.2904620.
21. *Thakkar A., Lohiya R.* A review of the advancement in intrusion detection datasets, *Procedia Comput Sci*., 2020, 167, pp. 636-645.

**Балыбердин Алексей Викторович –** Финансовый университет при Правительстве РФ**;** e-mail:balyberdinav@gmail.com**;** г. Москва, Россия; аспирант.

**Balyberdin Alexey Viktorovich** – Financial University under the Government of the Russian Federation; e-mail:balyberdinav@gmail.com; Moscow, Russia; graduate student.

УДК 004.89 DOI 10.18522/2311-3103-2025-3-16-31

**М.А. Лапина, Р.А. Дымуха, Н.Н. Кучеров, Е.С. Басан**

## ИССЛЕДОВАНИЕ МЕТОДОВ МАШИННОГО ОБУЧЕНИЯ ДЛЯ ОБНАРУЖЕНИЯ СПУФИНГ-АТАК В ДЕЦЕНТРАЛИЗОВАННЫХ СЕТЯХ

*Беспилотные летательные аппараты всё больше и больше появляются в нашей жизни и используются для различных целей, таких как доставка грузов, мониторинг, управление хозяйством, мониторинг и развлечения. Но вместе с ростом их популярности, увеличивается и число людей, которые намеренно хотят помешать работе БВС (беспилотным воздушным судам) и использовать в своих интересах и целях. Они используют различные виды атак, чтобы любыми способами*

*устранить или перехватить автономный летательный аппарат. Спуфинг-атаки являются одним из наиболее распространенных и опасных видов атак, так как позволяют злоумышленникам действовать незаметно, подделывая идентификаторы автономных летательных аппаратов или операторов, выдавая себя за легитимных участников системы. Целью таких атак может быть перехват управления, кража данных, саботаж или использование БВС для выполнения вредоносных действий, таких как шпионаж, нанесение ущерба или сбой в операциях. Но с каждым годом всё сложнее предотвращать атаки, так как они сложны в обнаружении и могут привести к серьезным последствиям, именно поэтому обнаружение спуфинг-атак на беспилотный аппарат при помощи машинного обучения активно исследуется и применяется. В статье рассматриваются спуфинг-атаки на БВС, проведен анализ спуфинга на автономные летательные аппараты, на основе открытого набора данных с помощью платформы Knime проведено исследование методов машинного обучения обнаружения спуфинг-атак. Результаты исследования демонстрируют, что способ обнаружения атак с помощью машинного обучения на основе ансамблевого метода, модели Tree Ensemble Learner и Random Forest Learner,, показавшие результаты 97.110% и 97.039% соответственно, является лучшим среди других методов, что позволит улучшить безопасность беспилотных летательных аппаратов, снижает нагрузку на операторов и повышает надежность системы в целом. В дальнейшем предложенный подход может быть расширен для обнаружения других видов кибератак, что сделает его универсальным методом защиты от воздействий злоумышленников.*

*Машинное обучение; Machine Learning; KNIME; поиск уязвимостей беспилотных воздушных судов; искусственный интеллект; данные; датасет; атаки на БВС; спуфинг.*

**M.A. Lapina, R.A. Dymuha, N.N. Kucherov, E.S. Basan**

## RESEARCH OF MACHINE LEARNING METHODS FOR DETECTING SPOOFING ATTACKS IN DECENTRALIZED NETWORKS

*Unmanned aerial vehicles are appearing more and more in our lives and are used for various purposes such as cargo delivery, monitoring, household management, exploration and entertainment. But along with their growing popularity, the number of people who intentionally want to interfere with the operation of UAVs and use them for their own interests and purposes is also increasing. They use various types of attacks to eliminate or intercept the drone by any means. Spoofing attacks are one of the most common and dangerous types of attacks, as they allow attackers to act unnoticed, faking the identifiers of autonomous aircraft or operators, posing as legitimate participants in the system. The purpose of such attacks may be to intercept control, steal data, sabotage, or use UAVs to perform malicious actions such as espionage, damage, or malfunction operations. But every year it becomes more difficult to prevent attacks, as they are difficult to detect and can lead to serious consequences, which is why such a solution as detecting spoofing attacks on an unmanned vehicle using machine learning was invented. The article discusses spoofing attacks on UAVs, analyzes spoofing on autonomous aircraft, and studies machine learning methods for detecting spoofing attacks based on a dataset using the Knime platform. The results of the study demonstrate that the method of detecting attacks using machine learning based on the ensemble method, the Tree Ensemble Learner and Random Forest Learner models, which showed results of 97.110% and 97.039%, respectively, is the best among other methods, which will improve the security of unmanned aerial vehicles, reduce the burden on operators and increase the reliability of the system as a whole. In the future, the proposed approach can be expanded to detect other types of cyberattacks, which will make it a universal method of protection against intruders.*

*Machine learning; Machine Learning; KNIME; drone vulnerability search; artificial intelligence; data; dataset; drone attacks; spoofing.*

**Введение.** В настоящее время, использование беспилотных воздушных судов (БВС) влечет за собой не только полезные свойства, но и ряд технических и социальных проблем, таких, как: проблемы с кибербезопасностью, конфиденциальностью и общественной безопасностью. БВС могут также использоваться злоумышленниками для проведения физических и кибератак, угрожающих социуму. При увеличении числа беспилотных летательных аппаратов становится все труднее выявлять и пресекать опасные беспилотные летательные аппараты, которые могут вызвать угрозу. Существуют разные виды атак на БВС, к ним можно отнести: спуфинг-атаки, человек посередине (MitM), атаки отказ в обслуживании, атаки прошивок, фишинговые атаки, атака на GPS-сигналы и атаки на

каналы связи [1, 2]. Исследование данных атак проводились Eldosouky A.R, Khan S.Z, Menaka P.A, Wesson K. [3–6]. В данной статье рассматривается один из самых распространённых видов атак – спуфинг атаки.

Спуфинг-атака – глушение и последующая подмена статического GPS-сигнала со спутника совершенно другим, более мощным, сигналом, который транслируется с наземной станции [7]. С помощью этого нового сигнала можно внести значительные изменения в заданные параметры. Таким образом, из-за получения ошибочных данных устройство быстро теряет ориентировку в пространстве. Такие атаки являются одним из наиболее распространенных и опасных видов атак на БВС, так как они позволяют злоумышленнику действовать незаметно, выдавая себя за легитимного участника системы. Спуфинг-атаки могут быть направлены на различные цели, включая перехват управления, кражу данных, саботаж или даже использование БВС для выполнения вредоносных действий [8].


*Рис. 1. Timeline спуфиг-атак на БВС*

Спуфинг-атаки представляют собой серьезную опасность для БВС. Эти действия были направлены на перехват управления БВС, позволяя себе легитимного оператора. В результате злоумышленник может получить контроль над БВС и использовать его для выполнения конкурентных действий, таких как шпионаж, саботаж или даже физическая атака. Первые спуфинг-атаки на БВС начали появляться в середине 2010-х годов, когда БВС стали более доступными и широко использовались в различных областях, таких как фотография, экономика, логистика и военные операции. Одной из первых известных атак стала атака на БВС, предпринятая для сельскохозяйственных угодий Диптихов. Злоумышленник перехватил управление БВС, изменил его маршрут и использовал его для сбора данных о состоянии полей конкурентов.

Эта атака продемонстрировала уязвимость БВС к спуфингу и необходимость разработки методов защиты. В связи с этим исследователи начали изучать способы обнаружения и предотвращения атак [9].

В 2011 году в Иран представил пресс-релиз, в котором говорилось об успешном перехвате американского БВС типа RQ-170 Sentinel. Среди прочих атак фигурировала и спуфинг-атака, в результате чего судно в автоматическом режиме, ориентируясь по глобальной системе навигации, начало возвращение домой. Поскольку истинный сигнал был заглушён ложным, RQ-170 Sentinel сел на иранский аэродром, приняв его за легитимный [9].

В 2012 году в Техасе американскими учеными была доказана практическая возможность взлома и перехват управления БВС путем GPS-спуфинга. Спуфинг-атака на GPS-атака, которая пытается обмануть GPS-приемник, широковещательно передавая немного более мощный сигнал, чем полученный от спутников GPS, такой, чтобы быть похожим на ряд нормальных сигналов GPS. Эти имитирующие сигналы изменены таким образом, чтобы заставить получателя неверно определять свое местоположение, считая его таким, какое отправит атакующий. Поскольку системы GPS работают, измеряя время, которое требуется для сигнала, чтобы атакующий точно знал, где его цель – так, чтобы имитирующий сигнал мог быть структурирован с надлежащими задержками сигнала [10].

В 2022 году Barak Davidovich, Ben Nassi и Yuval Elovici демонстрируют способность разработанного ими метода защищать БВС от атак с подменой GPS. Результаты исследования показывают, что они могут обеспечить высокий уровень безопасности

БВСа, летящего на высоте 50–100 м над городской местностью со средней скоростью 4 км/ч в условиях низкой освещенности. Предлагаемый метод может обеспечить уровень безопасности, который обнаруживает любую атаку с использованием GPS-спуфинга, при которой поддельное местоположение находится на расстоянии 1-4 м (в среднем 2,5 м) от реального местоположения. Преимущества данного метода включают тот факт, что он не требует никакого дополнительного оборудования или предварительных знаний о районе полета [11].

Схема, представленная на рис. 2, показывает механизм действия спуфинг-атаки. Злоумышленник через Wi-Fi-сигнал передает поддельный сигнал, имитирующий подлинный. В результате атаки соединение между оператором и БВС прерывается, так как БВС подключается к поддельному сигналу, следовательно оператор теряет управление БВС.

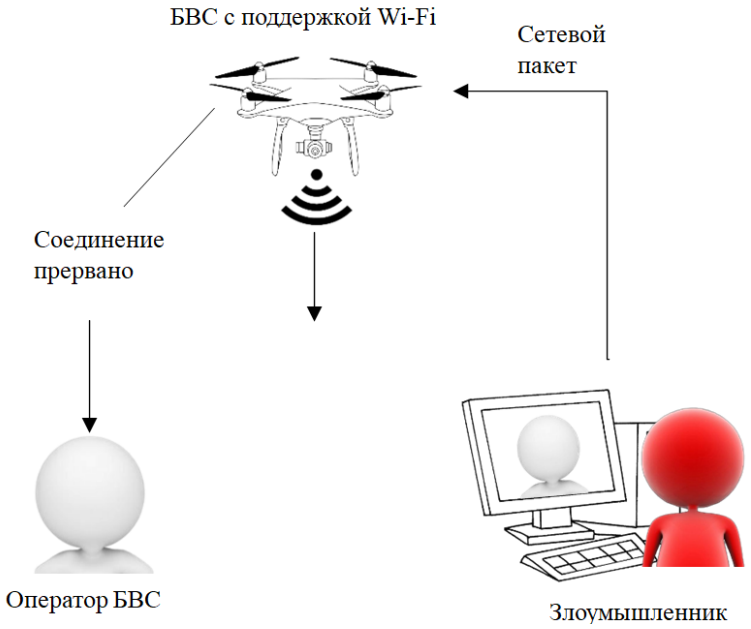


*Рис. 2. Схема спуфиг-атаки на БВС*

Для обнаружения атак на БВС применяется машинное обучение. В статье представлено исследование методов машинного обучения для обнаружения спуфинг-атак на БВС [12].

Для обнаружения спуфинга-атак необходимо определить параметры, которые будут анализироваться и проверяться на наличие аномалий или отклонений, указывающих на потенциальные угрозы.

Проведем анализ датасета, который использовался для исследования: drone_communication_dataset.csv [13].

Набор данных Drone Communication and Network Anomaly Detection содержит многомерные данные, собранные из моделируемой сети связи БВСов за период с 1 ноября 2019 года по 31 декабря 2024 года с почасовыми временными метками. Он включает в себя различные параметры связи, данные GPS, статистику сети и многоцелевые целевые индикаторы для различных сетевых аномалий. Набор данных предназначен для исследований и разработок в таких областях, как кибербезопасность, обнаружение аномалий IoT, связь БВС-БВС (D2D) и БВС-базовая станция (D2BS), а также оптимизация сети [13].

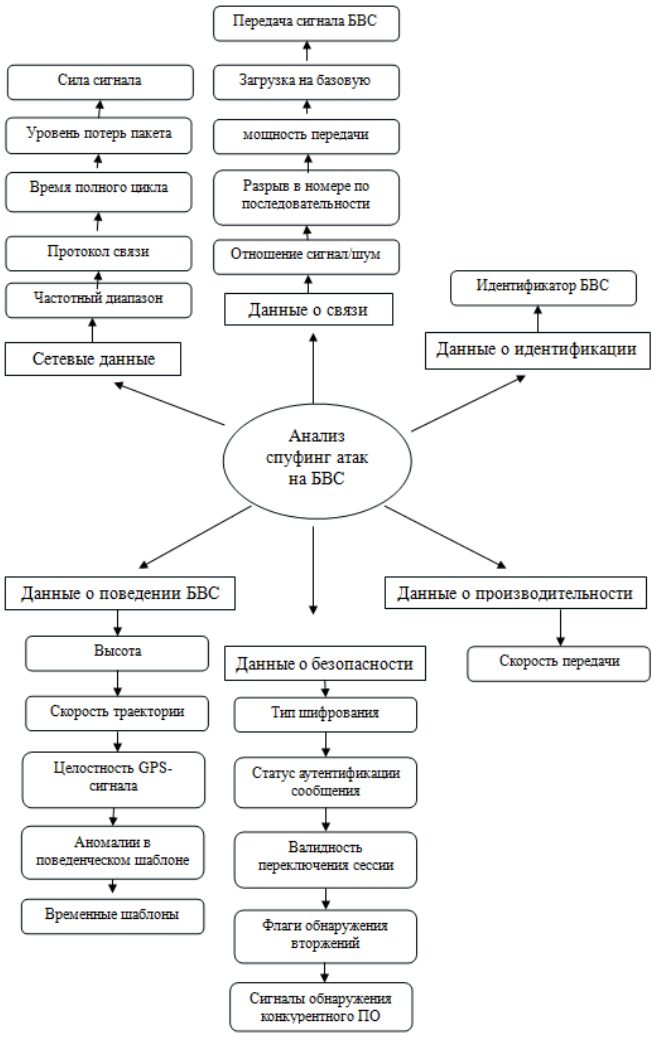


*Рис. 3. Параметры, используемые при обнаружении спуфинг-атак*

Набор данных содержит 45 289 записей и 35 столбцов. В табл. 1 приведено описание набора данных.

Таблица 1

**Анализ набора данных**

| № | Название колонки | Тип данных | Описание | Метод получения |
|---|---|---|---|---|
| 1. | timestamp | datetime | Время записи данных | Фиксируется при каждом измерении или передаче данных, записывается автоматически системой |
| 2. | signal_strength | числовой, int/float | Уровень сигнала (в dBm). Отрицательные значения указывают на силу сигнала (чем ближе к 0, тем сильнее сигнал) | Измеряется сенсорами или приемниками сигнала |

| | | | | |
|---|---|---|---|---|
| 3. | packet_loss_rate | числовой, int/float | Процент потерянных пакетов данных. | Вычисляется как отношение потерянных пакетов к общему числу отправленных |
| 4. | round_trip_time | числовой, int/float | Время (в мс) между отправкой запроса и получением ответа | Измеряется таймерами при передаче данных |
| 5. | communication_protocol | категориальный, string | Протокол связи (например, ZigBee, LoRa, Wi-Fi) | Устанавливается в настройках устройства или определяется автоматически |
| 6. | frequency_band | числовой (float) | Частота (в МГц или ГГц), на которой работает устройство | Указывается в конфигурации устройства |
| 7. | encryption_type | категориальный, string | Тип шифрования данных (например, AES, RSA, Plain-text) | Устанавливается в настройках безопасности |
| 8. | drone_gps_coordinates | категориальный, string | Географические координаты БВСа | Считываются с GPS-модуля БВСа |
| 9. | altitude | числовой, int/float | Высота БВСа (в метрах) | Измеряется барометром или GPS |
| 10. | speed_trajectory | числовой, int/float | Скорость движения БВСа (в м/с или км/ч) | Вычисляется на основе данных GPS или инерциальных датчиков |
| 11. | transmission_power | числовой, int/float | Мощность передачи сигнала (в dBm) | Устанавливается в настройках передатчика |
| 12. | message_authentication_status | бинарный, int | Статус аутентификации сообщения (1 – успешно, 0 – неудача) | Проверяется криптографическ ими алгоритмами |
| 13. | session_key_validity | бинарный, int | Валидность сессионного ключа (1 – действителен, 0 – недействителен) | Проверяется системой безопасности |
| 14. | signal_noise_ratio | числовой, int/float | Отношение сигнал/шум (в dB) | Измеряется приемником сигнала |
| 15. | sequence_number_gap | числовой, int | Разрыв в последовательности номеров пакетов | Анализируется при приеме данных |
| 16. | drone_identification | числовой (integer) | Уникальный идентификатор БВС | Присваивается при регистрации БВС |
| 17. | data_rate | числовой, int/float | Скорость передачи данных (в кбит/с или Мбит/с) | Устанавливается в настройках связи |
| 18. | network_traffic_volume | числовой, int/float | Объем сетевого трафика (в байтах или пакетах) | Измеряется сетевым оборудованием |
| 19. | gps_signal_integrity | бинарный, int | Целостность GPS-сигнала (1 – хорошая, 0 – плохая) | Анализируется GPS-приемником |
| 20. | uplink_downlink_quality | числовой, int/float | Качество связи (в условных единицах) | Измеряется приемопередатчи ком |
| 21. | base_station_load | числовой, int/float | Нагрузка на базовую станцию (в %) | Мониторится базовой станцией |
| 22. | port_scanning_attempts | числовой, int | Количество попыток сканирования портов | Фиксируется системами защиты |

| | | | | |
|---|---|---|---|---|
| 23. | drone_signal_handoff | бинарный, int | Факт передачи сигнала между станциями (1 – да, 0 – нет) | Логируется при переключении каналов |
| 24. | malware_detection_signals | бинарный, int | Количество сигналов о вредоносном ПО | Анализируется антивирусными системами |
| 25. | anomaly_in_behavioral_pattern | числовой, int | Уровень аномалии в поведении БВСа | Выявляется системами мониторинга |
| 26. | intrusion_detection_flags | бинарный, int | Флаги обнаружения вторжений | Устанавливаются системами безопасности |
| 27. | temporal_patterns | числовой, int | Временные закономерности в данных | Анализируются алгоритмами машинного обучения |
| 28. | label_normal | бинарный, int | Метка нормального состояния (1 – норма, 0 – аномалия) | Присваивается вручную или алгоритмами |
| 29. | label_spoofing | бинарный, int | Метка спуфинга (1 – атака, 0 – нет) | Определяется системами защиты |
| 30. | label_mitm | бинарный, int | Метка атаки "человек посередине" (1 – атака, 0 – нет) | Выявляется криптографическ ими методами |
| 31. | label_ddos | бинарный, int | Метка DDoS-атаки (1 – атака, 0 — нет) | Анализируется сетевыми системами |
| 32. | label_gps_spoofing | бинарный, int | Метка GPS-спуфинга (1 – атака, 0 – нет) | Обнаруживается GPS-приемниками |
| 33. | label_malware | бинарный, int | Метка вредоносного ПО (1 – обнаружено, 0 – нет) | Сканируется антивирусами |
| 34. | label_jamming | бинарный, int | Метка глушения сигнала (1 – атака, 0 – нет) | Фиксируется приемниками сигнала |
| 35. | label_protocol_exploit | бинарный, int | Метка эксплуатации уязвимостей протокола (1 – атака, 0 – нет) | Выявляется системами мониторинга |

Существуют разные подходы и платформы, которые позволяют работать с моделями машинного обучения. В работе проводились исследования с применением платформы для анализа данных Knime [14]. Платформа предоставляет широкий выбор моделей машинного. На рис. 4 приведены модели машинного обучения, которые применялись для исследования.

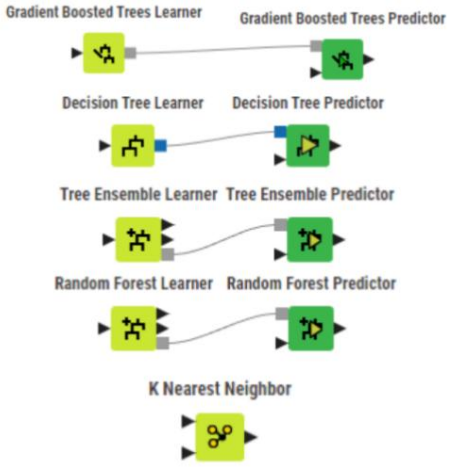


*Рис. 4. Модели машинного обучения*

В табл. 2 представлена сравнительная характеристика способа защиты от спуфинг-атак по критериям: удобство, приватность, скорость, цена, сложность обеспечения и эффективность

Таблица 2

**Сравнительная характеристика способа защиты от спуфинг-атак**

| Методы защиты | Удобство | Приватность | Скорость | Цена | Сложность обеспечения | Эффективность |
|---|---|---|---|---|---|---|
| Криптографические методы | ✔ | | | | | ✔ |
| Аутентификация | | ✔ | | | | ✔ |
| Мониторинг сетевого трафика | | | | | ✔ | ✔ |
| Обнаружение аномалий | | | | | ✔ | ✔ |
| Защита GPS-сигналов | | ✔ | | | | ✔ |
| Регулярное обновление прошивки | ✔ | | | | ✔ | |
| Физическая защита | | | | | | ✔ |

На основе данных из табл. 2, универсального способа защиты от спуфинг-атак не выявлено, однако машинное обучение позволяет анализировать данные и на их основе обнаруживать спуфинг-атаки

На примере Gradient Boosted Trees Learner рассмотрим структурную схему модели машинного обучения CSV Reader загружает данные из CSV-файла. Узел преобразует данные в таблицу, с которой происходить исследование. Затем данные поступают на узел Column Filter, который позволяет выбрать или исключить определенные столбцы из таблицы данных. В табл. 1 будут исключены такие калонки: label_normal, label_mitm, label_ddos, label_gps_spoofing, label_malware, label_jamming, label_protocol_exploit. Колонки убраны, так как в область исследования входит только спуфинг. Следующий нод, One to Many, преобразует данные из формата "один ко многим". Он создает отдельные строки для каждой категории, то есть кодирует данные. В наборе данных нужно закодировать категориальные данные, чтобы не возникало ошибок при исследовании. Поэтому будут кодироваться такие столбцы, как communication_protocol, encryption_type, drone_gps_coordinates. Узел Number to String конвертирует числовые значения в строки. В нем нужно преобразовать только целевую колонку, чтобы она могла использоваться узлом SMOTE. Узел SMOTE используется для балансировки несбалансированных данных. SMOTE создает синтетические примеры для меньшинственного класса, чтобы увеличить его представительство в данных. Узел PCA уменьшает размерность данных, сохраняя при этом наибольшую часть вариации. PCA преобразует данные в новые признаки, которые являются линейными комбинациями исходных признаков. Узел X-Partitioner разделяет данные на обучающую и тестовую выборки. Это важно для оценки производительности модели на данных, которые не использовались в процессе обучения. Подключается Gradient Boosted Trees Learner и к нему предиктор Gradient Boosted Trees Predictor. Узлы выполняют обучение  Узел X-Aggregator объединяет результаты предсказаний обучающей и тестовой выборок в одну таблицу. Наконец, узел Scorer вычисляет метрики качества модели, такие как точность, полнота, F1-score и другие, на основе предсказаний и истинных значений.

На примере Gradient Boosted Trees Learner была представлена методика организации 10 исследований с разными моделями машинного обучения. Ниже представлены 5 моделей, показавших лучшие результаты.

1. **Gradient Boosted Trees Learner** изучает деревья с градиентным усилением с целью классификации. Алгоритм использует очень мелкие деревья регрессии и специальную форму усиления для построения ансамбля деревьев [15].

Исследование Gradient Boosted Trees Learner на изменение PCA для определения лучшей точности приведено на рис. 5.

| Gradient Boosted Trees Learner | |
|---|---|
| PCA | F1-score |
| 29 | 0,791 |
| 27 | 0,792 |
| 25 | 0,791 |
| 23 | 0,782 |
| 21 | 0,771 |
| 19 | 0,801 |
| 17 | 0,665 |
| 15 | 0,664 |
| 13 | 0,634 |
| 11 | 0,715 |
| 9 | 0,609 |
| 7 | 0,59 |
| 5 | 0,586 |
| 3 | 0,578 |
| 1 | 0,63 |

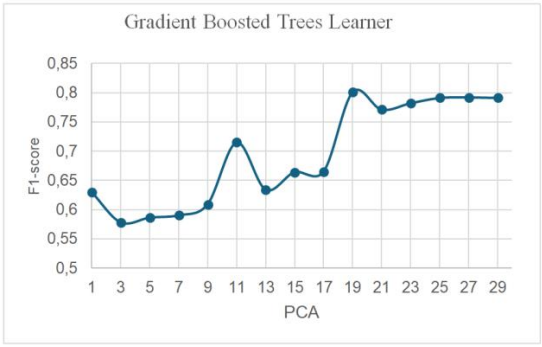


*Рис. 5. Результаты изменения точности PCA модели Gradient Boosted Trees Leamey*

На графике на рис. 5 показан результат изменения точности при изменении числа колонок в PCA. Наилучшим является PCA с 19 колонками и результатом 0,801.

2. **Decision Tree Learner** – этот узел индуцирует дерево решений классификации в основной памяти. Деревья решений строятся путем последовательного разделения данных на подмножества на основе значений признаков. Каждое разделение выбирается так, чтобы максимизировать однородность подмножеств относительно целевой переменной [16].

Исследование Decision Tree Learner на изменение PCA для определения лучшей точности, показано на рис. 6.

| Decision Tree Learner | |
|---|---|
| PCA | F1-score |
| 29 | 0,914 |
| 27 | 0,913 |
| 25 | 0,913 |
| 23 | 0,907 |
| 21 | 0,911 |
| 19 | 0,898 |
| 17 | 0,875 |
| 15 | 0,871 |
| 13 | 0,869 |
| 11 | 0,869 |
| 9 | 0,857 |
| 7 | 0,847 |
| 5 | 0,83 |
| 3 | 0,815 |
| 1 | 0,636 |

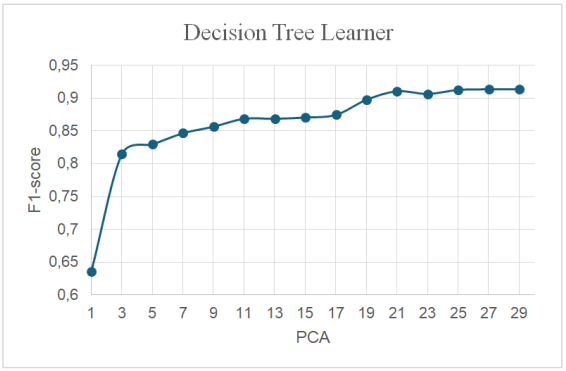


*Рис. 6. Результаты изменения точности PCA модели Decision Tree Leamer*

На графике показан результат изменения точности при изменении числа колонок в PCA. Наилучшим является PCA с 29 колонками и результатом 0,914.

3. **Tree Ensemble Learner** изучает ансамбль деревьев решений (например, варианты случайного леса). Обычно каждое дерево строится с различным набором строк (записей) и/или столбцов (атрибутов). Ансамбль деревьев решений – это метод машинного обучения, который объединяет множество деревьев решений для повышения точности предсказаний [17].

Исследование Tree Ensemble Learner на изменение PCA для определения лучшей точности, показано на рис. 7.

| Tree Ensemble Learner | |
|---|---|
| PCA | F1-score |
| 29 | 0,941 |
| 27 | 0,936 |
| 25 | 0,928 |
| 23 | 0,945 |
| 21 | 0,948 |
| 19 | 0,926 |
| 17 | 0,883 |
| 15 | 0,866 |
| 13 | 0,86 |
| 11 | 0,866 |
| 9 | 0,849 |
| 7 | 0,869 |
| 5 | 0,862 |
| 3 | 0,838 |
| 1 | 0,635 |

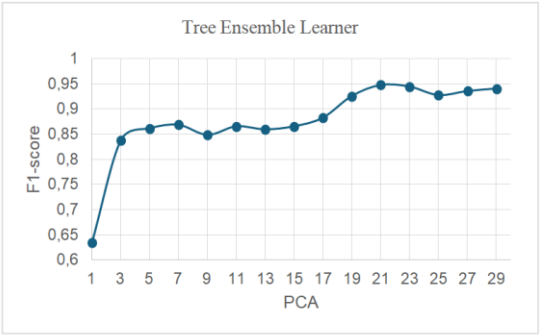


*Рис. 7. Результаты изменения точности PCA модели Tree Ensemble Learner*

На графике показан результат изменения точности при изменении числа колонок в PCA. Наилучшим является PCA с 21 колонками и результатом 0,948.

4. **Random Forest Learner** изучает случайный лес, состоящий из выбранного количества деревьев решений. Каждая из моделей дерева решений строится с различным набором строк (записей), и для каждого разделения в дереве используется случайно выбранный набор столбцов (описывающих атрибуты) [18].

Исследование Random Forest Learner на изменение PCA для определения лучшей точности, показано на рис. 8.

| Random Forest Learner | |
|---|---|
| PCA | F1-score |
| 29 | 0,973 |
| 27 | 0,975 |
| 25 | 0,974 |
| 23 | 0,974 |
| 21 | 0,973 |
| 19 | 0,966 |
| 17 | 0,953 |
| 15 | 0,945 |
| 13 | 0,935 |
| 11 | 0,931 |
| 9 | 0,922 |
| 7 | 0,898 |
| 5 | 0,88 |
| 3 | 0,838 |
| 1 | 0,614 |

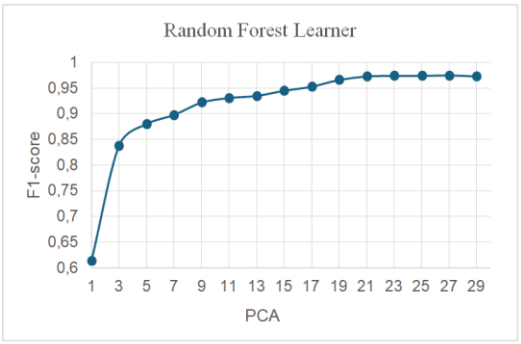


*Рис. 8. Результаты изменения точности PCA Random Forest Learner*

На данном графике показан результат изменения точности при изменении числа колонок в PCA. Наилучшим является PCA с 27 колонками и результатом 0,975.

5. **K Nearest Neighbor** классифицирует набор тестовых данных на основе алгоритма k ближайших соседей с использованием обучающих данных. Базовый алгоритм использует дерево KD и, следовательно, должен демонстрировать разумную производительность [19].

Исследование K Nearest Neighbor на изменение PCA для определения лучшей точности, как показано на рис. 9.

На данном графике показан результат изменения точности при изменении числа колонок в PCA. Наилучшим является PCA с 21 колонками и результатом 0,864.

| K Nearest Neighbor | |
|---|---|
| PCA | F1-score |
| 29 | 0,863 |
| 27 | 0,863 |
| 25 | 0,862 |
| 23 | 0,863 |
| 21 | 0,864 |
| 19 | 0,859 |
| 17 | 0,858 |
| 15 | 0,858 |
| 13 | 0,855 |
| 11 | 0,849 |
| 9 | 0,842 |
| 7 | 0,825 |
| 5 | 0,785 |
| 3 | 0,759 |
| 1 | 0,636 |

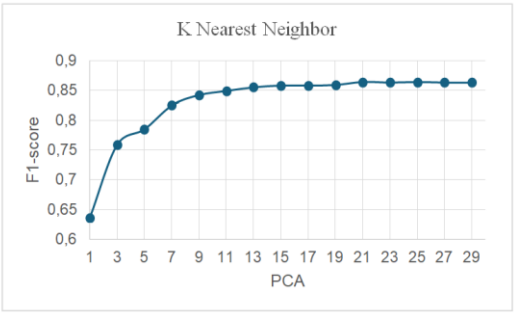


*Рис. 9. Результаты изменения точности PCA K Nearest Neighdor*

В табл. 3 представлены данные результатов исследования для выбранных моделей обучения.

Таблица 3

**Сравнительная таблица результатов PCA**

| № | Модель | PCA | F1-score |
|---|---|---|---|
| 1 | Gradient Boosted Trees Learner | 19 | 0,801 |
| 2 | Decision Tree Learner | 29 | 0,914 |
| 3 | Tree Ensemble Learner | 21 | 0,948 |
| 4 | Random Forest Learner | 27 | 0,975 |
| 5 | K Nearest Neighbor | 21 | 0,864 |

По результатам выбора наилучшей конфигурации PCA проведена настройка моделей машинного обучения. В табл. 4 представлены две модели машинного обучения на основе деревьев: Tree Ensemble Learner и Random Forest Learner

Таблица 4

**Результаты исследования моделей Tree Ensemble Learner и Random Forest Learner**

| Tree Ensemble Learner | Дерево | Random Forest Learner | Дерево |
|---|---|---|---|
| Глубина дерева | Accuracy, % | Глубина дерева | Accuracy, % |
| 20 | 83,926 | 20 | 84,598 |
| 25 | 87,774 | 25 | 87,652 |
| 30 | 90,878 | 30 | 90,388 |
| 35 | 93,112 | 35 | 92,396 |
| 40 | 94,855 | 40 | 93,992 |
| 45 | 95,897 | 45 | 95,208 |

*Окончание табл. 4*

| | | | |
|---|---|---|---|
| 50 | 96,465 | 50 | 96,172 |
| 55 | 96,799 | 55 | 96,663 |
| 60 | 97,072 | 60 | 97,011 |
| 65 | 97,103 | 65 | 97,031 |
| 70 | 97,11 | 70 | 97,039 |

Таким образом, две модели машинного обучения на основе деревьев: Tree Ensemble Learner и Random Forest Learner, показали близкий результат. При изменении настроек, обе модели выдавали лучшие результаты по мере увеличения глубины дерева, но при этом, достигнув определенного уровня, у моделей изменения стали минимальны. Это значит, что дальнейшее увеличение глубины может привести к тому, что деревья могут начать запоминать данные, следовательно, к переобучению.

В табл. 5 представлены результаты исследования модели машинного обучения K Nearest Neighbor.

Таблица 5

**Результаты исследования модели K Nearest Neighbor**

| K Nearest Neighbor | |
|---|---|
| K | Accuracy, % |
| 1 | 88,73 |
| 3 | 85,497 |
| 5 | 83,561 |
| 7 | 81,949 |
| 9 | 80,666 |

Таким образом, из таблицы видно, что наилучшим результатом является 3 ближайших соседа, которых следует учитывать.

В табл. 6 представлены результаты исследования модели машинного обучения Gradient Boosted Trees Learner.

Таблица 6

**Результаты исследования модели Gradient Boosted Trees Learner**

| Gradient Boosted Tree Learner | |
|---|---|
| Глубина дерева | Accuracy, % |
| 20 | 84,595 |
| 25 | 92,576 |
| 30 | 90,92 |
| 35 | 90,711 |
| 40 | 90,319 |

Таким образом, из таблицы видно, что наилучшим результатом будет глубина дерева равная 25.

В табл. 7 представлены результаты исследования модель машинного обучения Decision Tree Learner.

Таблица 7

**Результаты исследования модели Decision Tree Learner**

| Decision Tree Learner | |
|---|---|
| Глубина дерева | Accuracy,% |
| 1 | 0,863 |
| 3 | 0,863 |
| 5 | 0,862 |
| 7 | 0,863 |
| 9 | 0,864 |

Таким образом, из таблицы видно, что наилучшим результатом будет глубина дерева равная 1.

Таблица 8

**Сравнительная таблица результатов исследования моделей**

| № | Модель | Accuracy, % |
|---|---|---|
| 1 | K Nearest Neighbor | 85,497 |
| 2 | Tree Ensemble Learner | 97,11 |
| 3 | Random Forest Learner | 97,039 |
| 4 | Decision Tree Learner | 91,863 |
| 5 | Gradient Boosted Trees Learner | 92,576 |

Рассмотрим механизмы борьбы с переобучением. Исследуя модели машинного обучения, было замечено, что модель переобучается, поэтому возникла необходимость найти способ решения этой проблемы избегая замены датасета. В Knime есть отдельные блоки, которые предотвращают этот процесс, они позволяют выполнить кросс-валидацию, которая является стандартным методом для борьбы с переобучением. X-Partitioner и X-Aggregator – цикл перекрестной проверки, настроенный на пятикратное повторение. Это означает, что он делит набор данных на пять равных частей, и в каждой интеграции он использовал четыре части для обучения (80% данных) и одну часть для тестирования (20% данных). Узел X-Aggregator собирает все прогнозы на основе тестовых данных и предоставляет обобщенную оценку производительности модели [20].

**Заключение.** В настоящее время беспилотные летательные аппараты используют в различных сферах жизни: фермерство, доставка, строительство, видеосъемка и многое другое. Но использование БВС приносит человеку не только выгоду и удобство, но и проблемы. Например, проблемы с кибербезопасностью, конфиденциальностью и общественной безопасностью. БВС используются злоумышленниками для проведения физических и кибератак, что приводит к потере контроля над БВС и утечке важной информации. С ростом числа беспилотных систем усложняется идентификация и нейтрализация потенциально опасных устройств, создающих угрозу безопасности.

В работе проведено исследование методов обнаружения спуфинг-атак на БВС на основе машинного обучения. Наилучший результат показали модели Tree Ensemble Learner и Random Forest Learner, основанные на ансамблевом методе машинного обучения, показавшие результаты 97.110% И 97.039% соответственно. Это связанно с тем, что их особенностью является моделирование сложности задач в зависимости от данных, устойчивостью к шуму и высокой гибкостью. Предложенный подход отличается простотой реализации благодаря импользованию платформы Knime, которая позволяет решать алгоритмические блоки с предустановленным кодом. Это минимизирует необходимость глубоких знаний в программировании, позволяя сосредоточиться на оптимизации параметров моделей для повышения качества детектирования атак.

***Благодарность:*** *Исследование выполнено за счет гранта Российского научного фонда № 25-71-30007, https://rscf.ru/project/25-71-30007/.*

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. *Савинкова С.А.* Разработка метода отслеживания перемещений объектов // Вестник современных исследований. – 2021. – №. 1-6. – С. 28-36.
2. GPS: глушилки, спуфинг и уязвимости // SavePearlHarbor. – URL: https://savepearlharbor.com/?p=264385 (дата обращения: 03.04.2025).
3. *Eldosouky A.R., Ferdowsi A., Saad W.* Drones in distress: A game-theoretic countermeasure for protecting UAVs against GPS spoofing // IEEE Internet of Things Journal. – 2019. – Vol. 7, No. 4. – P. 2840-2854.
4. *Khan S.Z., Mohsin M., Iqbal W.* On GPS spoofing of aerial platforms: a review of threats, challenges, methodologies, and future research directions // PeerJ Computer Science. – 2021. – Vol. 7. – P. e507.
5. *Arthur M.P.* Detecting signal spoofing and jamming attacks in UAV networks using a lightweight IDS // 2019 international conference on computer, information and telecommunication systems (CITS). – IEEE, 2019. – P. 1-5.
6. *Wesson K.D., Shepard D.P., Bhatti J.A., Humphreys T.E.* An evaluation of the vestigial signal defense for civil GPS anti-spoofing // Proceedings of the 24th International Technical Meeting of the Satellite Division of The institute of navigation (ION GNSS 2011). – 2011. – P. 2646-2656.
7. *Савинкова С.А.* Разработка метода отслеживания перемещений объектов // Вестник современных исследований. – 2021. – №. 1-6. – С. 28-36.
8. Спуфинг БВСов // Спуфинг БВСов (БВС). – URL: https://protectionsystem.ru/spoofing (дата обращения: 03.04.2025).
9. Иранские хакеры смогли получить управление американским БПЛА и посадить его на своей территории. – https://habr.com/ru/articles/135150/ (дата обращения: 05.06.2025).
10. *Humphreys T.* Statement on the vulnerability of civil unmanned aerial vehicles and other systems to civil GPS spoofing // University of Texas at Austin (July 18, 2012). – 2012. – P. 1-16.
11. *Davidovich B., Nassi B., Elovici Y.* Towards the detection of GPS spoofing attacks against drones by analyzing camera's video stream // Sensors. – 2022. – Vol. 22, No. 7. – P. 2608.
12. What is machine learning? // IBM. – URL: https://www.ibm.com/think/topics/machine-learning (дата обращения: 03.04.2025).
13. Drone Communication Dataset // kaggle. – URL: https://www.kaggle.com/datasets/datasetengineer/drone-communication-dataset (дата обращения: 04.04.2025).
14. Что такое KNIME и как его использовать // Skypro. – URL: https://sky.pro/media/chto-takoe-knime-i-kak-ego-ispolzovat/ (дата обращения: 03.04.2025).
15. Gradient Boosted Trees Learner // Knime. – URL: https://hub.knime.com/knime/extensions/org.knime.features.ensembles/latest/org.knime.base.node.mine.treeensemble2.node.gradientboosting.learner.classification.GradientBoostingClassificationLearnerNodeFactory2 (дата обращения: 03.04.05).
16. Decision Tree Learner // Knime. – URL: https://hub.knime.com/knime/extensions/org.knime.features.base/latest/org.knime.base.node.mine.decisiontree2.learner2.DecisionTreeLearnerNodeFactory3 (дата обращения: 03.04.05).
17. Tree Ensemble Learner // Knime. – URL: https://hub.knime.com/knime/extensions/org.knime.features.ensembles/latest/org.knime.base.node.mine.treeensemble2.node.learner.classification.TreeEnsembleClassificationLearnerNodeFactory2 (дата обращения: 03.04.05).
18. Random Forest Learner // Knime. – URL: https://hub.knime.com/knime/extensions/org.knime.features.ensembles/latest/org.knime.base.node.mine.treeensemble2.node.randomforest.learner.classification.RandomForestClassificationLearnerNodeFactory2 (дата обращения: 03.04.05).
19. K Nearest Neighbor // Knime. – URL: https://hub.knime.com/knime/extensions/ org.knime.features.base/latest/org.knime.base.node.mine.knn.KnnNodeFactory2 (дата обращения: 03.04.05).
20. *Zul M.I., Yulia F., Nurmalasari D.* Social media sentiment analysis using K-means and naïve bayes algorithm // 2018 2nd International conference on electrical engineering and informatics (ICon EEI). – IEEE, 2018. – P. 24-29.

REFERENCES

1. *Savinkova S.A.* Razrabotka metoda otslezhivaniya peremeshcheniy ob"ektov [Development of a method for tracking the movement of objects], *Vestnik sovremennykh issledovaniy* [Bulletin of Modern Studies], 2021, No. 1-6, pp. 28-36.
2. GPS: glushilki, spufing i uyazvimosti [GPS: jammers, spoofing and vulnerabilities], *SavePearlHarbor*. Available at: https://savepearlharbor.com/?p=264385 (accessed 03 April 2025).
3. *Eldosouky A.R., Ferdowsi A., Saad W.* Drones in distress: A game-theoretic countermeasure for protecting UAVs against GPS spoofing, *IEEE Internet of Things Journal,* 2019, Vol. 7, No. 4, pp. 2840-2854.

4. *Khan S.Z., Mohsin M., Iqbal W.* On GPS spoofing of aerial platforms: a review of threats, challenges, methodologies, and future research directions, *PeerJ Computer Science,* 2021, Vol. 7, pp. e507.
5. *Arthur M.P.* Detecting signal spoofing and jamming attacks in UAV networks using a lightweight IDS, *2019 international conference on computer, information and telecommunication systems (CITS)*. IEEE, 2019, pp. 1-5.
6. *Wesson K.D., Shepard D.P., Bhatti J.A., Humphreys T.E.* An evaluation of the vestigial signal defense for civil GPS anti-spoofing, *Proceedings of the 24th International Technical Meeting of the Satellite Division of The institute of navigation (ION GNSS 2011),* 2011, pp. 2646-2656.
7. *Savinkova S.A.* Razrabotka metoda otslezhivaniya peremeshcheniy ob"ektov [Development of a method for tracking the movement of objects], *Vestnik sovremennykh issledovaniy* [Bulletin of modern studies], 2021, No. 1-6, pp. 28-36.
8. Spufing BVSov [Spoofing of UAVs], *Spufing BVSov (BVS)* [Spoofing of UAVs (UAVs)]. Available at: https://protectionsystem.ru/spoofing (accessed 03 April 2025).
9. Iranskie khakery smogli poluchit' upravlenie amerikanskim BPLA i posadit' ego na svoey territorii [Iranian hackers were able to gain control of an American UAV and land it on their territory]. Available at: https://habr.com/ru/articles/135150/ (accessed 05 June 2025).
10. *Humphreys T.* Statement on the vulnerability of civil unmanned aerial vehicles and other systems to civil GPS spoofing, University of Texas at Austin (July 18, 2012), 2012, pp. 1-16.
11. *Davidovich B., Nassi B., Elovici Y.* Towards the detection of GPS spoofing attacks against drones by analyzing camera's video stream, *Sensors,* 2022, Vol. 22, No. 7, pp. 2608.
12. What is machine learning?, *IBM*. Available at: https://www.ibm.com/think/topics/machine-learning (accessed 04 April 2025).
13. Drone Communication Dataset, *kaggle*. Available at: https://www.kaggle.com/datasets/datasetengineer/drone-communication-dataset (accessed 04 April 2025).
14. Chto takoe KNIME i kak ego ispol'zovat' [What is KNIME and how to use it], *Skypro*. Available at: https://sky.pro/media/chto-takoe-knime-i-kak-ego-ispolzovat/ (accessed 03 April 2025).
15. Gradient Boosted Trees Learner, *Knime*. Available at: https://hub.knime.com/knime/extensions/org.knime.features.ensembles/latest/org.knime.base.node.mine.treeensemble2.node.gradientboosting.learner.classification.GradientBoostingClassificationLearnerNodeFactory2 (accessed 03 April 2025).
16. Decision Tree Learner, *Knime*. Available at: https://hub.knime.com/knime/extensions/org.knime.features.base/latest/org.knime.base.node.mine.decisiontree2.learner2.DecisionTreeLearnerNodeFactory3 (accessed 03 April 2025).
17. Tree Ensemble Learner, *Knime*. Available at: https://hub.knime.com/knime/extensions/org.knime.features.ensembles/latest/org.knime.base.node.mine.treeensemble2.node.learner.classification.TreeEnsembleClassificationLearnerNodeFactory2 (accessed 03 April 2025).
18. Random Forest Learner, *Knime*. Available at: https://hub.knime.com/knime/extensions/org.knime.features.ensembles/latest/org.knime.base.node.mine.treeensemble2.node.randomforest.learner.classification.RandomForestClassificationLearnerNodeFactory2 (accessed 03 April 2025).
19. K Nearest Neighbor, *Knime*. Available at: https://hub.knime.com/knime/extensions/org.knime.features.base/latest/org.knime.base.node.mine.knn.KnnNodeFactory2 (accessed 03 April 2025).
20. *Zul M.I., Yulia F., Nurmalasari D.* Social media sentiment analysis using K-means and naïve bayes algorithm, *2018 2nd International conference on electrical engineering and informatics (ICon EEI)*. IEEE, 2018, pp. 24-29.

**Лапина Мария Анатольевна** – Северо-Кавказский федеральный университет; e-mail: mlapina@ncfu.ru; г. Ставрополь, Россия; к.ф.-м.н.; доцент кафедры вычислительной математики и кибернетики; ORCID: 0000-0001-8117-9142.

**Дымуха Регина Андреевна** – Северо-Кавказский федеральный университет; e-mail: dymuharegina@gmail.com; г. Ставрополь, Россия; кафедра информационной безопасности автоматизированных систем; студент; ORCID: 0009-0005-2107-4636.

**Кучеров Николай Николаевич** – Северо-Кавказский федеральный университет; e-mail: nik.bekesh@gmail.com; г. Ставрополь, Россия; к.т.н.; ведущий научный сотрудник департамента науки Северо-Кавказского федерального университета; ORCID: 0000-0003-0337-0093.

**Басан Елена Сергеевна** – Южный федеральный университет; e-mail: ele-barannik@yandex.ru; г. Таганрог, Россия; к.т.н.; доцент кафедры безопасности информационных технологий им. О.Б. Макаревича,ORCID: 0000-0001-6127-4484.

**Lapina Maria Anatolyevna** – North Caucasus Federal University; e-mail: mlapina@ncfu.ru; Stavropol, Russia; cand. of phys. and math. sc.; associate professor of the Department of Computational Mathematics and Cybernetics; ORCHID: 0000-0001-8117-9142.

**Dymuha Regina Andreevna** – North Caucasus Federal University; e-mail: dymuharegina@gmail.com; Stavropol, Russia; the Department of Information Security of Automated Systems; student; ORCHID: 0009-0005-2107-4636.

**Kucherov Nikolay Nikolaevich** – North Caucasus Federal University; e-mail: nik.bekesh@gmail.com; Stavropol, Russia; cand. of eng. sc.; leading researcher at the Department of Science; ORCHID: 0000-0003-0337-0093.

**Basan Elena Sergeevna** – Southern Federal University; e-mail: ele-barannik@yandex.ru; Taganrog, Russia; cand. of eng. sc.; associate professor of the Information Technology Security Department named after O.B. Makarevich; ORCHID: 0000-0001-6127-4484.

УДК 004.891.2 DOI 10.18522/2311-3103-2025-3-31-41

**А.Е. Анпилогова, В.А. Анпилогов**

## СИСТЕМА АВТОМАТИЗАЦИИ ДОКУМЕНТООБОРОТА И МОНИТОРИНГА ИНЦИДЕНТОВ ЭКОНОМИЧЕСКОЙ БЕЗОПАСНОСТИ НА ОСНОВЕ ТЕХНОЛОГИЙ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

*Автоматизация документооборота - ключевой элемент оптимизации процессов и повышения эффективности. Автоматизация документооборота на базе искусственного интеллекта улучшает управление инцидентами экономической безопасности, оптимизируя рабочие процессы и снижая затраты. Переход на автоматизированный документооборот в России связан со сложной нормативно-правовой базой и масштабными затратами на внедрение на предприятиях. Автоматизация помогает соблюдать требования законодательства и снижает риски юридических и финансовых последствий. Интеграция цифровых подписей повышает эффективность утверждения документов. Внедрение систем автоматизации поддерживает национальные цели цифровой трансформации. Автоматизация документооборота сокращает зависимость от бумажных процессов и способствует созданию централизованных цифровых хранилищ. Внедрение систем автоматизации документооборота требует стратегического подхода и тщательного планирования. Автоматизация документооборота обеспечивает экономию времени, сокращение ошибок и повышение соответствия нормативным стандартам. В статье рассмотрены теоретические основы BPM, интеграция цифровых технологий и нормативные аспекты, специфичные для России. Предложенная система сочетает мониторинг с ИИ и IoT, обеспечивает обработку данных в реальном времени, автоматизирует создание юридических документов и отчетов. Система автоматизации рабочих процессов базируется на интеграции данных, технологиях искусственного интеллекта и seamless-решениях. Система объединяет технологии мониторинга, алгоритмы распознавания лиц и анализа поведения, централизованную базу данных и модуль связи. Система формирует отчеты и юридические документы, заверенные QES, и обеспечивает взаимодействие с правоохранительными органами и службами безопасности. Результаты внедрения: снижение операционных расходов на 30–40% и уменьшение потерь на 50%. Система соответствует стандартам цифровой трансформации и поддерживает модернизацию национальной экономики.*

*Инцидент экономической безопасности; автоматизация документооборота; распознавание образов; анализа поведения; технлогии искусственного интеллекта.*

**A.E. Anpilogova, V.A. Anpilogov**

## A SYSTEM FOR AUTOMATING DOCUMENT FLOW AND MONITORING ECONOMIC SECURITY INCIDENTS BASED ON ARTIFICIAL INTELLIGENCE TECHNOLOGIES

*Automation of document flow is a key element of process optimization and efficiency improvement. Automation of document flow based on artificial intelligence improves the management of economic security incidents by optimizing work processes and reducing costs. The transition to automated document flow in Russia is associated with a complex regulatory framework and large-scale implementation costs at*