ТРАНСПОРТНЫЕ СИСТЕМЫ
2.9.1 – Транспортные и транспортно-технологические системы страны, ее регионов и городов, организация производства на транспорте;
2.9.4. – Управление процессами перевозок;
2.9.6 – Аэронавигация и эксплуатация авиационной техники;
2.9.8 – Интеллектуальные транспортные системы

УДК 004.056:629.7
DOI: 10.26467/2079-0619-2025-28-5-8-21

Метод анализа многомерных сочетаний признаков сетевого трафика для выявления признаков несанкционированного вмешательства в авиационных сетях передачи данных

А.А. Ганичев[1]
[1]*Московский государственный технический университет гражданской авиации, г. Москва, Россия*

Аннотация: В связи с увеличением интенсивности и усложнением сетевого взаимодействия авиационных систем передачи данных существенно возрастает потребность в разработке методов выявления признаков несанкционированного вмешательства в авиационную деятельность. Важность данной проблемы обусловлена необходимостью обеспечения устойчивости авиационной инфраструктуры к разнообразным угрозам, способным привести к критическим нарушениям работы систем управления воздушным движением и повлиять на безопасность полетов воздушных судов. В статье разработан и представлен метод анализа многомерных сочетаний признаков сетевого трафика авиационных систем передачи данных, основанный на модифицированном алгоритме частотного анализа FP-Growth, адаптированном под специфику многомерных данных. Отличительной особенностью предложенного подхода является сохранение контекста признаков и возможность выявления скрытых зависимостей между различными параметрами сетевых событий, которые недоступны традиционным одномерным алгоритмам частотного анализа. Разработана модель представления сетевых событий в виде многомерных транзакций, предложен алгоритм построения многомерного дерева частых признаков и извлечения устойчивых сочетаний признаков с заданной частотой встречаемости. По результатам экспериментальной проверки на реальных данных сетевого трафика подтверждена возможность выявления шаблонов сетевых атак и ранее не регистрируемых аномальных сочетаний признаков. Выполнена количественная оценка производительности предлагаемого метода, подтвердившая его эффективность и пригодность для обработки значительных объемов информации, характерных для авиационных систем передачи данных, в режиме реального времени. Предложенный метод обеспечивает повышение защищенности авиационных сетей и своевременное выявление угроз авиационной деятельности. Разработанный метод может быть использован для повышения устойчивости АСПД систем УВД к угрозам и приоритетного выбора мер защиты для обеспечения безопасности полетов.

Ключевые слова: несанкционированное вмешательство, информационная безопасность, авиационная сеть передачи данных, обнаружение атак, риск, частотный анализ, сетевой трафик.

Для цитирования: Ганичев А.А. Метод анализа многомерных сочетаний признаков сетевого трафика для выявления признаков несанкционированного вмешательства в авиационных сетях передачи данных // Научный вестник МГТУ ГА. 2025. Т. 28, № 5. С. 8–21. DOI: 10.26467/2079-0619-2025-28-5-8-21

Method of analyzing multidimensional combinations of network traffic features for identifying signs of unauthorized intrusion in aviation data transmission networks

A.A. Ganichev[1]
[1]*Moscow State Technical University of Civil Aviation, Moscow, Russia*

Abstract: Due to the increasing intensity and complexity of network interactions in aviation data transmission systems, the need for developing methods to detect signs of unauthorized interference in aviation operations is significantly growing. The importance of this issue is due to the need to ensure control systems and affect the safety of aircraft flights. This article develops and presents a method for analyzing multidimensional combinations of network traffic features in aviation data transmission systems, based on a modified frequent-pattern FP-Growth algorithm adapted specifically for multidimensional data. A distinctive feature of the proposed approach is maintaining the contextual integrity of network event attributes, enabling the identification of hidden dependencies among various parameters of network events that are inaccessible to traditional one-dimensional frequent pattern analysis algorithms. A model for representing network events as multidimensional transactions is formulated, and an algorithm for constructing a multidimensional frequent-pattern tree and extracting stable combinations of features with a predefined frequency of occurrence is proposed. Experimental validation using real network traffic data confirmed the capability of detecting network attack patterns and previously unrecorded anomalous feature combinations. A quantitative evaluation of the proposed method's performance was conducted, confirming its efficiency and suitability for processing substantial data volumes characteristic of aviation data transmission systems in real-time conditions. The developed method provides improved protection for aviation networks and timely identification of threats to aviation operations. The developed method can be applied to enhance the resilience of aviation data transmission systems for air traffic management and prioritize protective measures to ensure flight safety.

Key words: unauthorized interference, information security, network reliability, aviation data network, attack detection, risk, frequency analysis; network traffic.

For citation: Ganichev, A.A. (2025). Method of analyzing multidimensional combinations of network traffic features for identifying signs of unauthorized intrusion in aviation data transmission networks. Civil Aviation High Technologies, vol. 28, no. 5, pp. 8–21. DOI: 10.26467/2079-0619-2025-28-5-8-21

Введение

В современных условиях постоянный рост объема и сложности сетевого трафика на воздушном транспорте требует разработки новых подходов к анализу данных авиационных сетей передачи данных (АСПД) [1–3]. АСПД представляют собой совокупность каналов цифрового обмена между воздушными судами (ВС), объектами наземной инфраструктуры воздушного транспорта и системами управления воздушным движением (УВД), формируемую в рамках информационного взаимодействия в доменах аэронавигационного и оперативного контроля [4]. В соответствии с Федеральным законом № 187-ФЗ «О безопасности критической информационной инфраструктуры Российской Федерации»[1] авиационные системы связи и УВД входят в число критически важных объектов. В соответствии с положениями Федерального закона № 16-ФЗ «О транспортной безопасности»[2] воздушный транспорт отнесен к числу объектов, в отношении которых предусмотрены меры противодействия актам незаконного вмешательства, нарушающим функционирование информационных и технических систем, при этом подход к обеспечению безопасности в авиации должен основываться на международных стандартах, установленных ИКАО. На международном уровне соответствующие направ-

[1] Федеральный закон № 187-ФЗ от 26 июля 2017 г. О безопасности критической информационной инфраструктуры Российской Федерации [Электронный ресурс] // КонсультантПлюс. 2017. URL: https://www.consultant.ru/document/cons_doc_LAW_220885/ (дата обращения: 22.03.2025).

[2] Федеральный закон № 16-ФЗ от 9 февраля 2007 г. О транспортной безопасности [Электронный ресурс] // КонсультантПлюс. 2007. URL: https://www.consultant.ru/document/cons_doc_LAW_66069/ (дата обращения: 22.03.2025).

ления закреплены в Глобальном плане обеспечения авиационной безопасности, утвержденном по итогам 41-й сессии Ассамблеи ИКАО[3]. В актуальной редакции документа подчеркивается значимость учета угроз информационной безопасности, связанных с воздействием на сетевую инфраструктуру авиационного транспорта, при модернизации систем и реализации новых технологических решений. НСВ в СПД, обеспечивающие функционирование систем управления ВД, рассматривается как один из возможных векторов реализации подобных угроз.

Согласно ГОСТ Р 57240-2016[4] к числу объектов, подлежащих защите от несанкционированного вмешательства (НСВ), отнесены технические средства связи и сети передачи данных, задействованные в работе воздушного транспорта. Авиационные СПД представляют собой гетерогенные сети, объединяющие различные подсистемы и протоколы, с большим количеством источников событий [5–7]. При этом сохранение безопасности этих сетей осложняется появлением новых угроз несанкционированного вмешательства в работу бортовых и наземных узлов связи [8–10]. Традиционные средства защиты, основанные на фиксированных сигнатурах или простом пороговом контроле, часто не способны своевременно выявлять сложные многоэтапные атаки, маскирующиеся под легитимный трафик [11]. В последние годы исследователи отмечают необходимость корреляционного анализа событий в сетях для обнаружения скрытых зависимостей, указывающих на атакующие воздействия [12–14]. Однако существующие решения, как правило, рассчитаны на однородные данные и не учитывают всю полноту информации, присутствующей в телеметрии авиационных сетей.

Современные исследования указывают на наличие значительного числа уязвимостей в архитектуре АСПД. Нарушители могут искажать маршрутную информацию, осуществлять перехват управляющих команд и нарушать функционирование взаимосвязанных подсистем. В беспилотных авиационных комплексах, как показано в [15, 16], отсутствуют надежные механизмы противодействия НСВ. Отмечены случаи внедрения вредоносных компонентов, захвата элементов управления и деструктивного воздействия на навигационные устройства. Быстрое развитие средств и методов осуществления НСВ на АСПД сопровождается отставанием отраслевых регламентов и нормативных основ. Анализ текущих стандартов, регламентирующих обеспечение информационной безопасности в сфере воздушного транспорта (в том числе международных), показывает, что нормативная база не охватывает всего спектра актуальных векторов атак. Результаты прикладных исследований в области оценки рисков и тестирования подтверждают недостаточную эффективность действующих защитных механизмов при моделировании сложных сценариев воздействия [17, 18].

Одним из перспективных подходов к выявлению признаков атак является анализ часто встречающихся сочетаний признаков сетевого трафика. Основная гипотеза состоит в том, чтобы обнаруживать устойчиво повторяющиеся комбинации атрибутов сетевых событий, которые могут указывать на характерные сценарии поведения, в том числе связанные с НСВ в АСПД. Например, многократное появление определенной последовательности действий в различных подсистемах сети может свидетельствовать о целенаправленной вредоносной активности. Ассоциативный анализ данных, широко применяемый для поиска шаблонов (паттернов) в транзакционных базах, представляет особый интерес в данном контексте. Алгоритмы поиска частых наборов, такие как Apriori и FP-Growth, изначально предназначены для выявления часто совместно встречающихся элементов в массивах транзакций [19, 20]. Применение их идей к задаче анализа сетевого трафика открывает возможность автоматического выявления характерных комбинаций параметров сетевых пакетов, потенциально указывающих на атаки.

[3] Doc 10118: Global Aviation Security Plan: 2nd ed. // ICAO, 2024. 60 p.

[4] ГОСТ Р 57240-2016. Защита информации. Системы обеспечения информационной безопасности воздушного транспорта. Общие положения. М.: Стандартинформ, 2016. 18 с.

Однако прямое использование классических методов частотного анализа в АСПД сталкивается с серьезными затруднениями [21, 22]. С одной стороны, традиционные алгоритмы ассоциативного правила (Apriori, FP-Growth) рассчитаны на одномерные данные, то есть на транзакции, представленные в виде неупорядоченных наборов элементов из одного множества признаков. В задаче анализа сетевых событий мы имеем дело с многомерными записями: каждое сетевое событие характеризуется совокупностью разнородных атрибутов (время, протокол, адреса отправителя и получателя, порт, флаги и проч.). Сведение многомерного события к простому набору элементов приводит к потере структуры: становится неясно, какой атрибут какого значения касается. С другой стороны, раздельный анализ по каждой категории признаков лишает возможности обнаружить взаимосвязи между различными типами атрибутов. В результате традиционные алгоритмы либо теряют контекст, либо не раскрывают междоменные зависимости, которые как раз и могут содержать ключ к распознаванию комплексных атак.

Таким образом, возникает необходимость в специальном методе, учитывающем многомерную природу данных авиационных сетей. В настоящей работе предлагается метод поиска частых наборов признаков, адаптированный для многомерных данных сетевого трафика авиационной СПД. Цель исследования – разработка и демонстрация эффективности метода, способного выявлять многомерные частые паттерны сетевых событий, характерные для НСВ.

Методы

Задача обнаружения признаков НСВ в АСПД систем УВД формулируется следующим образом: требуется на основе журналов сетевых событий выявить устойчиво повторяющиеся комбинации признаков (атрибутов пакетов и соединений), которые статистически встречаются значительно чаще во время атак или аномальных ситуаций, чем при штатной работе системы. Выявленные частые многомерные шаблоны должны лечь в основу правил корреляции событий для последующего оповещения о потенциальных атаках. При этом метод должен работать в условиях гетерогенности источников (разные подсистемы генерируют события с разными форматами) и большого объема данных, поступающих в режиме, близком к реальному времени.

Для поиска часто встречающихся сочетаний элементов в наборах данных наиболее известны два подхода. Алгоритм Apriori [21] выполняет итеративное порождение и проверку наборов-кандидатов, опираясь на принцип антимонотонности: если некоторый набор признаков нечастый, то и любое его надмножество не может быть частым. Apriori последовательно (по возрастанию размера наборов) генерирует комбинации элементов и отфильтровывает те, чья поддержка не достигает минимального порога. Этот метод прост и понятен, однако плохо масштабируется на большие базы с высокой размерностью записей: число кандидатов растет экспоненциально, и алгоритм вынужден многократно сканировать весь набор данных, что приводит к большим затратам памяти и времени при анализе сетевых логов большого объема.

Более производительным решением является алгоритм Frequent Pattern Growth (FP-Growth), предложенный в [22]. Он устраняет необходимость генерировать все комбинации, вместо этого используя компактную древовидную структуру для представления транзакций. На первом проходе по данным FP-Growth вычисляет частоты отдельных элементов и отбрасывает редкие; на втором – строит FP-дерево (FP-tree), в узлах которого хранятся элементы, а пути от корня соответствуют транзакциям. Повторяющиеся префиксы транзакций в таком дереве занимают общее место, благодаря чему достигается сжатие данных. Частые наборы извлекаются из FP-дерева с помощью обхода путей, оканчивающихся на интересующий элемент, и восстановления соответствующих комбинаций. FP-Growth часто работает значительно быстрее Apriori, особенно на данных с боль-

шим количеством общих фрагментов, что свойственно, например, потокам сетевых пакетов, где многие соединения имеют схожие атрибуты.

Несмотря на преимущества FP-Growth, оба рассмотренных алгоритма исходят из предположения о плоской транзакции, то есть из того, что каждая запись – это набор элементов из единого унифицированного множества. В анализе сетевых событий это означало бы, например, что мы заранее перечисляем все возможные значения всех атрибутов (IP-адресов, портов, типов сообщений и т.д.) в одно большое множество и представляем каждое событие как подмножество этих значений. Такой подход не отражает структуру данных: теряется информация о том, какое значение к какому полю относилось. Например, значения «80» в поле «порт» и в поле «длина пакета» для классических алгоритмов неразличимы. Как результат, многомерное событие превращается в бессмысленный набор разрозненных данных. Альтернативный вариант – вести поиск частых сочетаний раздельно по каждому атрибуту тоже неприменим, так как не позволит выявить связь между разными аспектами события (например, между номером порта и флагом протокола). В авиационных сетях такая связь критически важна: комплексные атаки проявляются сочетанием признаков в разных измерениях (временных, протокольных, адресных), и их невозможно обнаружить анализом каждого поля по отдельности. Таким образом, необходимо усовершенствовать метод частотного анализа, чтобы он работал с многомерными транзакциями.

Разработка метода анализа многомерных сочетаний признаков сетевого трафика

Для решения поставленной задачи разработан алгоритм, модифицирующий FP-Growth таким образом, чтобы учитывать принадлежность каждого элемента к определенному атрибуту (измерению) сетевого события. Ключевая идея состоит в кодировании элементов с указанием их измерения, построении на этой основе особого дерева частых паттернов и извлечении сочетаний, учитывающих несколько измерений одновременно. Формально каждое событие (например, единичная запись в журнале сети) представлено вектором признаков

$$E_j = \left(x_1^{(j)}, x_2^{(j)}, \dots, x_m^{(j)}\right), \quad (1)$$

где $x_i^{(j)}$ – значение i-го признака (атрибута) в j-м событии.

Последовательность наблюдаемых сетевых событий за некоторый период можно обозначить как множество (или упорядоченный набор) таких векторов:

$$D = \{E_1, E_2, \dots, E_N\}, \quad (2)$$

где N – общее число событий в рассматриваемой выборке данных. В силу гетерогенности источников и форматов данных перед анализом выполняется нормализация и объединение событий в единую структуру. Это требует удаления дубликатов, синхронизации временных меток, унификации представления атрибутов. Результатом является нормализованная последовательность $\overline{D}$, пригодная для дальнейшего ассоциативного анализа:

$$\overline{D} = \{\bar{E}_1, \bar{E}_2, \dots, \bar{E}_N\}, \quad (3)$$

где $\bar{E}_j$ – вектор признаков j-го события после предварительной обработки (сохранена исходная семантика, но приведен единый формат данных).

На основном этапе реализуется модифицированный метод частотного анализа, обозначаемый как Multi Dimensional Frequent Pattern Growth (MDFP). Его принцип работы следующий. Вместо простых значений метод оперирует парами (атрибут, значение). То есть каждый возможный исходный элемент кодируется как пара вида $a_{attr}: v$, где a_{attr} – идентификатор атрибута (например, номер поля или название параметра), а v – конкретное значение. Например, если первый атри-

бут – это номер порта, а второй – тип протокола, то элемент $a_1: 80$ будет означать значение 80 в поле «номер порта», а $a_1: TCP$ – значение TCP во втором атрибуте (тип протокола). Благодаря этому различаются одинаковые по написанию значения, принадлежащие разным признакам, и данные из разных источников могут быть объединены для анализа без потери контекста.

Метод строит специальное MDFP-дерево (многомерное FP-дерево) для представления всех транзакций $\overline{D}$. Построение происходит в два шага:

1) одно прохождение по данным для вычисления поддержки каждого возможного значения в разрезе каждого измерения, то есть для каждой пары (атрибут a_{attr}, значение v) подсчитывается, сколько раз она встречается в наборе событий. Элементы, не достигающие заданного минимального порога поддержки σ_{min}, считаются редкими и отбрасываются как неинформативные. Остальные частые элементарные признаки сортируются по убыванию частоты;

2) выполняется второе чтение набора данных, в ходе которого строится FP-дерево. Каждый узел дерева содержит пару $a_{attr}: v$ и счетчик вхождений. Транзакции (события) добавляются поочередно: для текущего события берется упорядоченный (согласно шагу 1) список пар $a_{attr}: v$, после чего этот список добавляется в дерево как путь от корня. Если на данном пути некоторые узлы (префикс последовательности) уже существуют в дереве, их счетчики увеличиваются; новые узлы создаются для тех элементов, которые встречаются впервые на данном ветвлении. В результате получается компактная древовидная структура, в которой общие префиксы событий хранятся единым участком. Это экономит память и позволяет затем эффективно извлекать частые комбинации.

На рис. 1 приведен упрощенный пример построения MDFP-дерева. В этом примере показано постепенное добавление транзакций (сетевых событий) и формирование общей древовидной структуры: повторяющиеся последовательности атрибутов агрегируются, а новые ветви появляются только для новых сочетаний. Как видно, итоговое дерево содержит узлы, помеченные парами $a_{attr}: v$, и каждый узел имеет счетчик, отражающий, сколько раз данный признак встречается в соответствующем контексте. Рис. 1 ясно демонстрирует, что после обработки всех транзакций дерево компактно отражает все частые многомерные комбинации признаков, присутствующие в исходных данных (в примере показано состояние после чтения 10 транзакций). На основе построенного MDFP-дерева затем осуществляется обход для извлечения частых наборов признаков.

Для извлечения частых паттернов из MDFP-дерева применяются стандартные техники, аналогичные используемым в FP-Growth: выполняется поиск всех путей, оканчивающихся данным значением, и реконструируются соответствующие комбинации признаков (с учетом различных измерений). В результате работы алгоритма формируется множество P всех частых многомерных шаблонов:

$$P = \{p_k | \sup(p_k) \geq \sigma_{min}\}, \quad (4)$$

где $\sup(p_k)$ – поддержка (частота встречаемости) паттерна p_k в выборке, σ_{min} – заданный минимальный уровень поддержки. Поддержка рассчитывается как доля транзакций (событий) из формулы (2), содержащих данный набор признаков. Паттерны, попавшие в множество P, представляют собой устойчиво повторяющиеся комбинации признаков событий, потенциально указывающие на присутствие аномальной активности. В контексте сетевой безопасности такие шаблоны могут соответствовать сигнатурам атак или характерному поведению вредоносного трафика.

В отличие от одномерных алгоритмов, в которых значения признаков обрабатываются без учета их принадлежности к конкретным атрибутам, многомерный анализ сохраняет структурную связь между признаками и их измерениями. Это означает, что выявленные паттерны включают информацию о связях между полями сетевого события. Например,

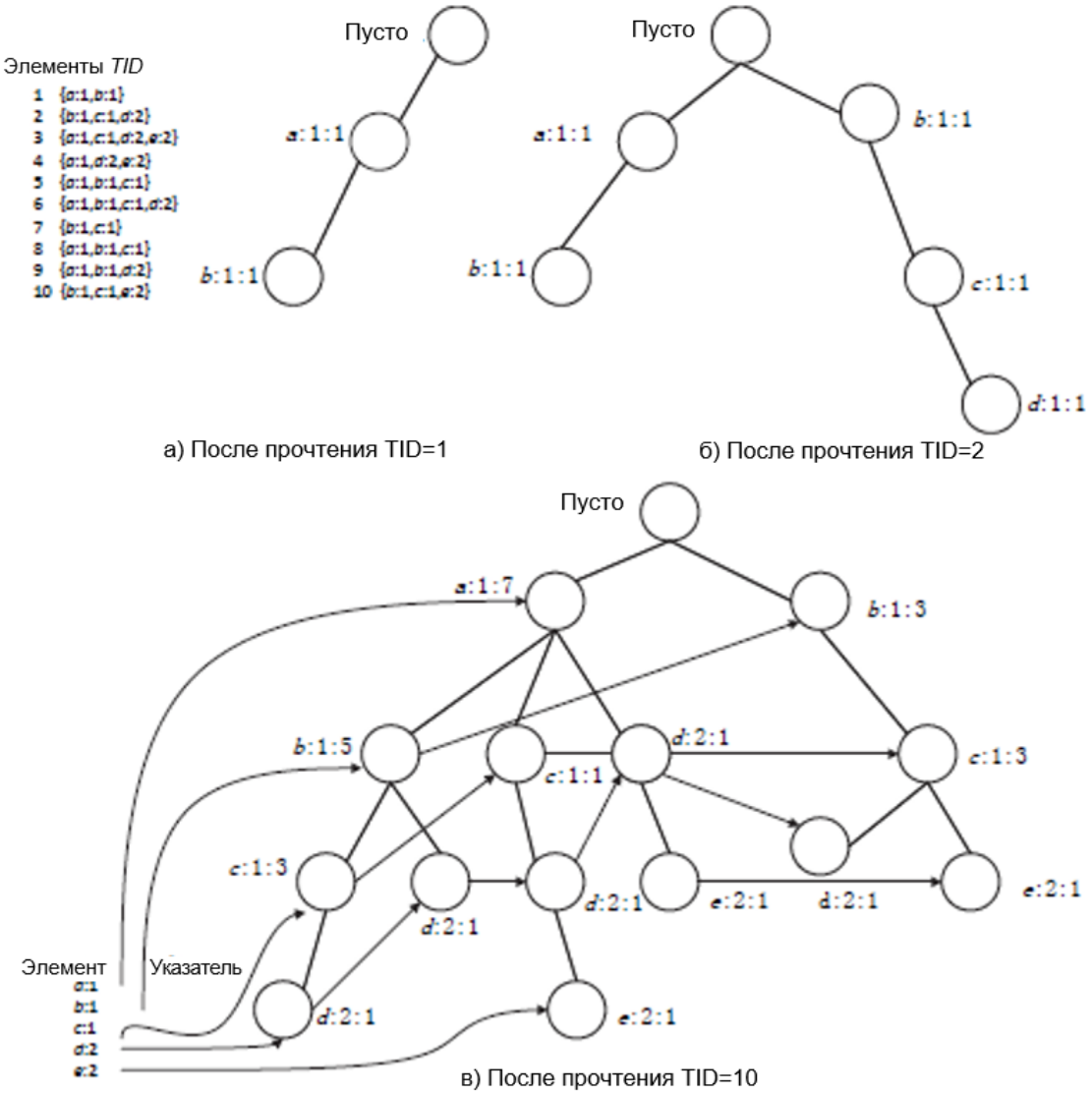

Рис. 1. Построение дерева MDFP
Fig. 1. Construction of the MDFP tree

метод способен обнаружить, что определенное значение порта часто встречается вместе с конкретным значением флага протокола и адреса отправителя, то есть фиксирует комплексный признак, который может быть характерен только для атаки. Классические подходы не видят разницы между комбинацией {порт = 80, флаг = SYN} и {порт = 80, флаг = RST} как между разными явлениями, тогда как в методе это будут разные паттерны, поскольку флаг – отдельное измерение. За счет этого достигается способность выявлять более сложные и значимые корреляции признаков, отражающие потенциальные попытки НСВ.

Предлагаемый метод многомерного анализа частых наборов признаков лег в основу общей многоуровневой модели ассоциации событий информационной безопасности в авиационной СПД. На рис. 2 представлена схема этой модели, состоящей из нескольких уровней обработки событий безопасности. Уровень сбора телеметрии агрегирует сырые события из различных подсистем: бортовых шлюзов, приемников протоколов (ACARS, VDL, SATCOM и др.), коммутаторов сетей управления воздушным движением, наземных маршрутизаторов и т. д. На выходе этого уровня формируется поток событий в виде

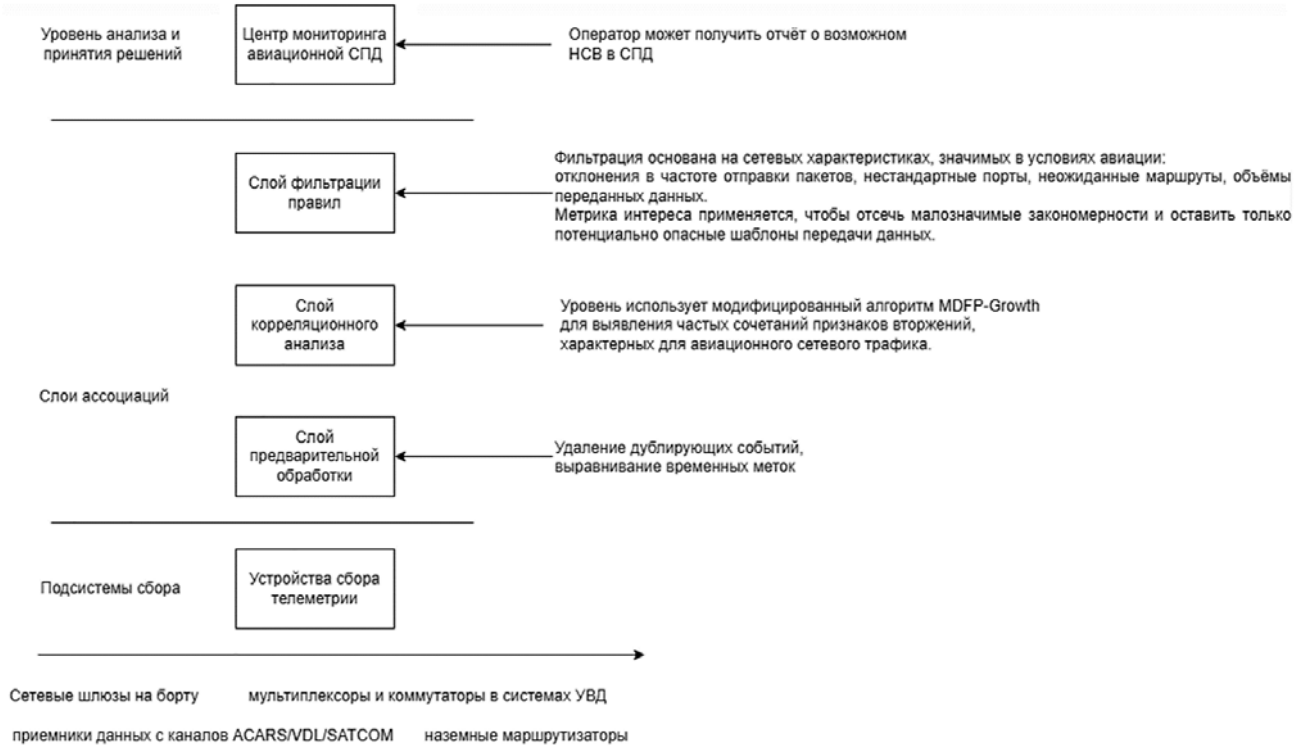

Рис. 2. Модель ассоциации событий информационной безопасности в авиационной СПД
Fig. 2. Model of information security events association in aviation data transmission systems

векторов атрибутов, который соответствует множеству D по формуле (2).

Уровень предварительной обработки выполняет описанные выше процедуры очистки и нормализации данных, приводя их к унифицированному виду, и получается последовательность $\overline{D}$ по формуле (3). Далее следует уровень корреляционного анализа – на нем применяется предлагаемый метод для выявления множества частых многомерных шаблонов P по формуле (4). Полученные паттерны отражают статистически значимые сочетания признаков сетевых событий, среди которых ожидается наличие признаков атак. Уровень фильтрации правил предназначен для отбора самых информативных результатов: на этом этапе каждому частому паттерну может быть сопоставлено ассоциативное правило и рассчитаны показатели его значимости. В частности, вычисляется поддержка паттерна (доля событий, ее определяющих, что соответствует внутреннему критерию частоты), достоверность ($confidence$, условная вероятность появления следствия при наличии основания) и специальная метрика интереса, отражающая степень отклонения обнаруженной зависимости от случайной нормы. Метрика интереса задается методом степени нормы, например в виде относительного превышения фактической совместной поддержки над ожидаемой при независимости событий:

$$Lift(A \Rightarrow B) = \frac{support(A \cup B)}{support(A) \cdot support(B)}. \quad (5)$$

Здесь $A \Rightarrow B$ – ассоциативное правило, соответствующее некоторому паттерну (множество A имплицирует событие B), $support(A \cup B)$ – поддержка объединения A и B (доля совместного появления), а произведение $support(A) \cdot support(B)$ представляет ожидаемую поддержку при условии статистической независимости A и B. Таким образом, $Lift(A \Rightarrow B) > 0$ означает, что связка A, B встречается существенно чаще, чем можно ожидать случайно, то есть правило представляет интерес для анализа. На этапе фильтрации отсеиваются шаблоны с недоста-

Рис. 3. Алгоритм фильтрации правил ассоциаций
Fig. 3. Association rules filtering algorithm

точной поддержкой и достоверностью, а также тривиальные зависимости с близким к единице значением $Lift(A \Rightarrow B)$ – последние считаются отражающими нормальные, заранее известные взаимосвязи и не несут аналитической ценности. Блок-схема алгоритма фильтрации правил ассоциаций представлена на рис. 3. Более подробно алгоритм планируется раскрыть в последующих публикациях.

После фильтрации остаются наиболее значимые правила ассоциаций, потенциально указывающие на НСВ. Они могут быть ис-

пользованы на завершающем уровне анализа и принятия решений – например, переданы в центр мониторинга, где оператор получает уведомление или отчет об обнаруженных подозрительных корреляциях событий.

Следует отметить, что в рамках данной статьи основное внимание уделяется этапу выделения частых многомерных паттернов (уровень ассоциаций на рис. 2). Этапы формирования ассоциативных правил и их фильтрации упомянуты для полноты картины, однако детальный их анализ выходит за рамки обсуждения. В частности, расчет показателей достоверности и интереса (формула (5)) служит для дальнейшей работы с правилами корреляции и будет предметом отдельного исследования.

Экспериментальные результаты

Для проверки работоспособности и эффективности предложенного метода был проведен эксперимент на данных сетевого трафика, содержащих как нормальные, так и атакующие воздействия. В качестве исходных данных использовалась выборка из общедоступного набора CICIDS2017 [23], который имитирует трафик сети с различными атаками (DoS/DDoS, сканирование портов, Bruteforce и др.) наряду с легитимной активностью. Данный набор был выбран ввиду наличия меток атак и разнообразия представленных признаков сетевых соединений. Из него были извлечены журналы сетевых событий, включающие основные атрибуты пакетов: временная метка, продолжительность соединения, адрес источника, адрес назначения, номер порта назначения, используемый протокол, количество переданных байт, число пакетов, флаги TCP и др. Всего рассматривалось $m = 83$ ключевых признака, характеризующих каждое событие. Были отфильтрованы избыточные и малозначимые поля (например, идентификаторы записей), данные разных подсистем объединены по времени в единую последовательность событий. Общее число проанализированных событий составило порядка $N \approx 10^5$. Для предлагаемого метода был задан порог минимальной поддержки $\sigma_{min} \approx 0{,}005$ (0,5 % от всех событий) – порог достаточно малый, чтобы выявить даже редкие, но регулярно повторяющиеся атаки.

По результатам работы метода выделено несколько сотен частых паттернов. Большинство из них соответствуют легитимным закономерностям (например, регулярные служебные обмены данными между узлами, периодические сигнальные сообщения авиационного оборудования и т. п.), что отражает нормальную работу сети. Однако среди полученных результатов особый интерес представляют шаблоны, характерные для известных атак, имевшихся в выборке. Так, метод обнаружил паттерн, соответствующий атакам типа PortScan: сочетание признаков, включающее непривилегированные номера портов в широком диапазоне, один и тот же IP-адрес источника и малое число пакетов в каждом соединении. Этот паттерн был выявлен как частый, поскольку в сетевом трафике набора данных присутствовала серия сканирований портов, генерирующая множество однотипных коротких соединений от одного источника. Другой примечательный найденный шаблон отразил распределенную DoS-атаку на определенный узел: в частых паттернах оказалась комбинация {адрес назначения = dst.ip, флаг TCP = SYN} – она указывает, что адрес X (конкретный IP сервера) многократно фигурировал в качестве цели соединений с флагом SYN (начало TCP-сессии) без последующего установления – именно так проявляется SYN-Flood DDoS. Примечательно, что выявленное сочетание признаков было обнаружено исключительно при многомерном анализе: при рассмотрении атрибутов по отдельности адрес назначения dst.ip не относился к часто встречающимся значениям, а флаг SYN также широко представлен в штатном трафике. Однако их совместное появление демонстрировало статистическую значимость именно в интервалах, соответствующих реализации атаки.

Помимо явных признаков атак, метод позволил выявить аномальные корреляции между признаками, ранее не наблюдавшимися в

нормальной работе. Например, обнаружено, что в периоды, соответствующие внедрению вредоносного ПО, часто встречалось сочетание {тип протокола = HTTP, нетипичный высокий номер порта, больший объем переданных данных}. Данная комбинация может указывать на нежелательную активность (возможно, экс-фильтрацию данных через нестандартный порт HTTP). Классические системы мониторинга, построенные на фиксированных правилах, не регистрировали данную ситуацию, поскольку каждый признак в отдельности оставался в пределах допустимых значений. Однако при многомерном анализе было выявлено, что одновременное присутствие этих признаков наблюдается с повышенной частотой именно в периоды нарушений. Полученные результаты демонстрируют, что предлагаемый метод позволяет выявлять скрытые многомерные закономерности в структуре трафика, отражающие характерные признаки различных типов атак.

Обсуждение результатов

Таким образом, апробация показала корректность и результативность разработанного метода. На реальных данных сетевого трафика были обнаружены частые многомерные шаблоны, соответствующие известным атакам, а также новые нетривиальные корреляции, потенциально указывающие на аномалии. Метод пригоден для применения в системе мониторинга АСПД и способен обрабатывать большие потоки событий в приемлемые сроки.

В перспективе планируется развитие подхода в направлении оптимизации хранения и обработки больших объемов поступающих данных. В частности, необходимо уделить внимание эффективной организации хранения промежуточных структур (таких как MDFP-дерево) и реализовать метод компактного сохранения и обновления информации о сетевом трафике. Кроме того, дальнейшая работа будет направлена на совершенствование механизмов фильтрации и ранжирования выявленных паттернов по степени важности, а также на интеграцию алгоритма с модулями реального времени для применения в системах УВД.

Заключение

В работе представлен метод анализа сетевого трафика АСПД, основанный на выявлении часто встречающихся многомерных сочетаний признаков сетевых событий. Предложенный метод модифицирует классический метод частотного анализа FP-Growth с учетом многомерной структуры данных, что позволило сохранить контекст признаков и выявить сложные зависимости между различными атрибутами сетевых событий, недоступные при применении традиционных одномерных алгоритмов. Разработана формальная модель представления многомерных сетевых событий и алгоритм извлечения устойчивых сочетаний признаков с заданным уровнем поддержки (формулы (1)–(4)). Экспериментальное исследование на наборе данных CICIDS2017 подтвердило работоспособность метода и показало, что он эффективно выделяет признаки, характерные для сетевых атак типа PortScan и SYN-Flood DDoS, а также выявляет нетривиальные аномальные комбинации, ранее не наблюдавшиеся стандартными системами мониторинга. Научная новизна состоит в том, что впервые для АСПД разработан частотный анализ многомерных данных с сохранением структурных связей между атрибутами. Практическая ценность заключается в возможности автоматического обнаружения устойчивых сочетаний признаков, характерных для атак без предварительно заданных сигнатур, что повышает точность выявления НСВ и позволяет сформировать основу для разработки правил корреляции событий и последующего мониторинга безопасности АСПД. Разработанный метод может быть использован для повышения устойчивости АСПД систем УВД к угрозам и приоритетного выбора мер защиты для обеспечения безопасности полетов.

Список литературы

1. Ганичев А.А., Пителинский К.В., Бритвина В.В. Статистический анализ потенциальных угроз информационной безопасности в бортовой сети воздушного судна // Вопросы защиты информации. 2024. № 1 (144). С. 11–22. DOI: 10.52190/2073-2600_2024_1_11

2. Кротова Е.Л., Андреев Р.А., Андреева П.А. Big data в авиационной отрасли: варианты применения // Международный научно-исследовательский журнал. 2021. № 5-1 (107). С. 6–9. DOI: 10.23670/IRJ.2021.107.5.001

3. Liu D. Deep learning aided packet routing in aeronautical ad-hoc networks relying on real flight data: from single-objective to near-Pareto multi-objective optimization / D. Liu, J. Zhang, J. Cui, S.-X. Ng, R.G. Maunder, L. Hanzo [Электронный ресурс] // Networking and Internet Architecture. 2021. DOI: 10.48550/arXiv.2110.15145 (дата обращения: 22.03.2025).

4. Hillebrecht A., Marks T., Gollnick V. An aeronautical data communication demand model for the North Atlantic oceanic airspace // CEAS Aeronautical Journal. 2023. Vol. 14. Pp. 553–567. DOI: 10.1007/s13272-023-00651-4

5. Adamopoulou E., Daskalakis E. Applications and technologies of big data in the aerospace domain [Электронный ресурс] // Electronics. 2023. Vol. 12, iss. 10. ID: 2225. DOI: 10.3390/electronics12102225 (дата обращения: 22.03.2025).

6. Secera J., Novak A. The future of data communication in Aviation 4.0 environment // INCAS Bulletin. 2021. Vol. 13, iss. 3. Pp. 165–178. DOI: 10.13111/2066-8201.2021.13.3.14

7. Dou X. Big data and smart aviation information management system [Электронный ресурс] // Cogent Business & Management. 2020. Vol. 7, iss. 1. DOI: 10.1080/23311975.2020.1766736 (дата обращения: 22.03.2025).

8. Hu W. Security monitoring of heterogeneous networks for big data based on distributed association algorithm / W. Hu, J. Li, J. Cheng, H. Guo, H. Xie // Computer Communications. 2020. Vol. 152. Pp. 206–214.

9. Ганичев А.А. Модель угроз несанкционированного вмешательства в беспроводных информационных системах авионики / А.А. Ганичев, К.В. Пителинский, С.А. Кесель, В.А. Пиков // Вопросы защиты информации. 2024. № 4 (147). С. 35–43. DOI: 10.52190/2073-2600_2024_4_35

10. Петров В.И. Методика анализа программного обеспечения бортовых компьютеров воздушного судна на отсутствие недекларированных возможностей сигнатурно-эвристическим способом // Научный вестник МГТУ ГА. 2017. Т. 20, № 1. С. 186–193.

11. Shawly T. Architectures for detecting interleaved multi-stage network attacks using hidden Markov models / T. Shawly, A. Elgharian, J. Kobes, A. Ghafoor // IEEE Transactions on Dependable and Secure Computing. 2019. Vol. 18, no. 5. Pp. 2316–2330. DOI: 10.1109/TDSC.2019.2948623

12. Kotenko I., Gaifulina D., Zelichenok I. Systematic literature review of security event correlation methods // IEEE Access. 2022. Vol. 10. Pp. 43387–43420. DOI: 10.1109/ACCESS.2022.3168976

13. Maosa H., Ouazzane K., Ghanem M.C. A hierarchical security event correlation model for real-time threat detection and response [Электронный ресурс] // Network. 2024. Vol. 4, no. 1. Pp. 68–90. DOI: 10.3390/network4010004 (дата обращения: 22.03.2025).

14. Cheng Q. STEP: Spatial-temporal network security event prediction / Q. Cheng, Y. Shen, D. Kong, C. Wu [Электронный ресурс] // Cryptography and Security. 2021. DOI: 10.1109/TIFS.2024.1234567 (дата обращения: 22.03.2025).

15. Исрафилов А. Современные вызовы в области кибербезопасности беспилотных авиационных систем [Электронный ресурс] // Universum: технические науки. 2024. № 2 (119). URL: https://7universum.com/ru/tech/archive/item/16760 (дата обращения: 22.03.2025).

16. Лянгузов Д.А., Плюснин Н.И. Безопасность и уязвимость сетей беспилотных летательных аппаратов: обзор // Известия Тульского государственного университета. Технические науки. 2023. № 7. С. 528–529. DOI: 10.24412/2071-6168-2023-7-528-529

17. Corretjer P.J. A cybersecurity analysis of today's commercial aircrafts and aviation industry systems: A thesis master of science. USA. NY: Utica College, 2018. 22 p.

18. Kessler G.C., Craiger J.P. Aviation cybersecurity: An overview [Электронный ресурс] // NTAS. 2018. URL: https://commons.erau.edu/ntas/2018/presentations/37/ (дата обращения: 22.03.2025).

19. Liu L.J. Research and application of improved Apriori algorithm // Computer Engineering and Design. 2017. Vol. 38, no. 12. Pp. 3324–3328.

20. Wang J.M., Yuan W. Improved FP-Growth algorithm based on node table // Computer Engineering and Design. 2018. Vol. 39, no. 1. Pp. 140–145.

21. Srinadh V. Evaluation of Apriori, FP-Growth and Eclat association rule mining algorithms // International Journal of Health Sciences. 2022. Vol. 6, no. S 2. Pp. 7475–7485. DOI: 10.53730/ijhs.v6nS2.6729

22. Srivastava A., Sinha D. FP growth-based zero-day attack signature extraction & detection model for high-volume attacks on real-time data stream [Электронный ресурс] // SSRN. 2023. 38 p. DOI: 10.2139/ssrn.4701527 (дата обращения: 22.03.2025).

23. Ali H. Imbalance class problems in data mining: A review / H. Ali, M.N.M. Salleh, R. Saedudin, K. Hussain, M.F. Mushtaq // Indonesian Journal of Electrical Engineering and Computer Science. 2019. Vol. 14, no. 3. Pp. 1552–1563. DOI: 10.11591/ijeecs.v14.i3.pp1552-1563

References

1. Ganichev, A.A., Pitelinskiy, K.V., Britvina, V.V. (2024). Statistical analysis of potential information security threats in aircraft onboard networks. *Information Security Questions*, vol. 1 (144), pp. 11–22. DOI: 10.52190/2073-2600_2024_1_11 (in Russian)

2. Krotova, E.L., Andreev, R.A., Andreeva, P.A. (2021). Big data in the aviation industry: application options. *International Research Journal*, 2021, no. 5-1 (107), pp. 6–9. DOI: 10.23670/IRJ.2021.107.5.001 (in Russian)

3. Liu, D., Zhang, J., Cui, J., Ng, S.-X., Maunder, R.G., Hanzo, L. (2021). Deep learning aided packet routing in aeronautical ad-hoc networks relying on real flight data: from single-objective to near-Pareto multi-objective optimization. *Networking and Internet Architecture*. DOI: 10.48550/arXiv.2110.15145 (accessed: 22.03.2025).

4. Hillebrecht, A., Marks, T., Gollnick, V. (2023). An aeronautical data communication demand model for the North Atlantic oceanic airspace. *CEAS Aeronautical Journal*, vol. 14, pp. 553–567. DOI: 10.1007/s13272-023-00651-4

5. Adamopoulou, E., Daskalakis, E. (2023). Applications and technologies of big data in the aerospace domain. *Electronics*, vol. 12, issue 10, ID: 2225. DOI: 10.3390/electronics 12102225 (accessed: 22.03.2025).

6. Secera, J., Novak, A. (2021). The future of data communication in Aviation 4.0 environment. *INCAS Bulletin*, vol. 13, issue 3, pp. 165–178. DOI: 10.13111/2066-8201.2021. 13.3.14

7. Dou, X. (2020). Big data and smart aviation information management system. *Cogent Business & Management*, vol. 7, issue 1. DOI: 10.1080/23311975.2020.1766736 (accessed: 22.03.2025).

8. Hu, W., Li, J., Cheng, J., Guo, H., Xie, H. (2020). Security monitoring of heterogeneous networks for big data based on distributed association algorithm. *Computer Communications*, vol. 152, pp. 206–214.

9. Ganichev, A.A., Pitelinskiy, K.V., Kesel, S.A., Pikov, V.A. (2024). Threat model of unauthorized interference in wireless avionics information systems. *Information Security Questions*, no. 4 (147), pp. 35–43. DOI: 10.52190/2073-2600_2024_4_35 (in Russian)

10. Petrov, V.I. (2017). The technique of analysis of software of on-board computers of air vessel to absence of undeclared capabilities by signature-heuristic way. *Civil Aviation High Technologies*, vol. 20, no. 1, pp. 186–193. (in Russian)

11. Shawly, T., Elghariani, A., Kobes, J., Ghafoor, A. (2019). Architectures for detecting interleaved multi-stage network attacks using hidden Markov models. *IEEE Transactions on Dependable and Secure Computing*, vol. 18, no. 5, pp. 2316–2330. DOI: 10.1109/TDSC. 2019.2948623

12. Kotenko, I., Gaifulina, D., Zelichenok, I. (2022). Systematic literature review of security event correlation methods. *IEEE Access*, vol. 10, pp. 43387–43420. DOI: 10.1109/ACCESS.2022.3168976

13. Maosa, H., Ouazzane, K., Ghanem, M.C. (2024). A hierarchical security event correlation model for real-time threat detection and response. *Network*, vol. 4, no. 1, pp. 68–90. DOI: 10.3390/network4010004 (accessed: 22.03.2025).

14. Cheng, Q., Shen, Y., Kong, D., Wu, C. (2021). STEP: Spatial-temporal network security event prediction. *Cryptography and Security*. DOI: 10.1109/TIFS.2024.1234567 (accessed: 22.03.2025).

15. Israfilov, A. (2024). Contemporary challenges in cybersecurity of unmanned aerial systems. *Universum: Technical Sciences*, no. 2 (119). Available at: https://7universum.com/ru/tech/archive/item/16760. (accessed: 22.03.2025). (in Russian)

16. Lyanguzov, D.A., Plusnin, N.I. (2023). Security and vulnerability of unmanned aerial vehicle networks: review. *Izvestiya Tulskogo gosudarstvennogo universiteta. Tekhnicheskiye nauki*, no. 7, pp. 528–529. DOI: 10.24412/2071-6168-2023-7-528-529

17. Corretjer, P.J. (2018). A cybersecurity analysis of today's commercial aircrafts and aviation industry systems: A thesis master of science. USA. NY: Utica College, 22 p.

18. Kessler, G.C., Craiger, J.P. (2018). Aviation cybersecurity: An overview. *NTAS*. Available at: https://commons.erau.edu/ntas/2018/presentations/37/ (accessed: 22.03.2025).

19. Liu, L.J. (2017). Research and application of improved Apriori algorithm. *Computer Engineering and Design*, vol. 38, no. 12, pp. 3324–3328.

20. Wang, J.M., Yuan, W. (2018). Improved FP-Growth algorithm based on node table. *Computer Engineering and Design*, vol. 39, no. 1, pp. 140–145.

21. Srinadh, V. (2022). Evaluation of Apriori, FP-Growth and Eclat association rule mining algorithms. *International Journal of Health Sciences*, vol. 6, no. S 2, pp. 7475–7485. DOI: 10.53730/ijhs.v6nS2.6729

22. Srivastava, A., Sinha, D. (2023). FP growth-based zero-day attack signature extraction & detection model for high-volume attacks on real-time data stream. *SSRN*, 38 p. DOI: 10.2139/ssrn.4701527 (accessed: 22.03.2025).

23. Ali, H., Salleh, M.N.M., Saedudin, R., Hussain, K., Mushtaq, M.F. (2019). Imbalance class problems in data mining: A review. *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 14, no. 3, pp. 1552–1563. DOI: 10.11591/ijeecs.v14.i3.pp1552-1563

Сведения об авторе

Ганичев Александр Александрович, старший преподаватель кафедры основ радиотехники и защиты информации МГТУ ГА, alexunderlich@gmail.com.

Information about the author

Alexandr A. Ganichev, Senior Lecturer, Fundamentals of Radio Engineering and Information Security Chair, Moscow State Technical University of Civil Aviation, alexunderlich@gmail.com.

Поступила в редакцию	22.04.2025	Received	22.04.2025
Одобрена после рецензирования	04.06.2025	Approved after reviewing	04.06.2025
Принята в печать	25.09.2025	Accepted for publication	25.09.2025