Для цитирования
Бурый А.С., Усцелемов В.Н. Гибридная модельная аналитика безопасных информационных технологий / III Научно-практическая конференция «Стандартизация: траектория науки», приуроченная ко Всемирному дню стандартов, Москва, 15 октября 2025 г. // Информационно-экономические аспекты стандартизации и технического регулирования. 2025. № 6(87). С. 513–520.

УДК 004.8+004.056

ГИБРИДНАЯ МОДЕЛЬНАЯ АНАЛИТИКА БЕЗОПАСНЫХ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ

Бурый А.С., *д-р техн. наук, ФГБУ «Институт стандартизации», Россия, г. Москва,*
Усцелемов В.Н.*, аспирант, ФГБУ «Институт стандартизации», Россия, г. Москва*

Современные исследователи нередко сталкиваются с проблемой, когда модели, методы и алгоритмы, разработанные в недавнем прошлом, становятся совершенно беспомощными при обработке больших и разнотипных данных. Требуется переход от традиционной аналитики данных к аналитике больших данных, нацеленной на раскрытие потенциальных ценностей данных – новых знаний, повышающих в том числе и безопасность существующих информационных технологий. Целью работы является рассмотрение инструментария методов предиктивной аналитики в обеспечении безопасности информационных технологий, ориентированных на поддержку принятия решений на основе структуризации гибридных информационно–аналитических моделей.

Ключевые слова: гибридные модели, интеллектуальный анализ данных, предиктивная аналитика, безопасность информационных технологий, системы поддержки принятия решений.

HYBRID MODEL ANALYTICS OF SECURE INFORMATION TECHNOLOGIES

Buryi A.S., *Doctor of Technology, Russian Standardization Institute, Moscow*
Ustselemov V.N., *graduate student, Russian Standardization Institute, Moscow*

Modern researchers often face the problem when models, methods, and algorithms developed in the recent past become completely helpless when processing large and diverse data. A transition from traditional data analytics to big data analytics is required, aimed at uncovering the potential values of data – new knowledge that increases, among other things, the security of existing information technologies. The purpose of the work is to consider the tools of predictive analytics methods in ensuring the security of information technologies focused on decision support based on the structuring of hybrid information and analytical models.

Keywords: hybrid models, data mining, predictive analytics, information technology security, decision support systems.

Введение

Разработка методов обеспечения надежной переработки информации, устойчивой к целенаправленным или случайным воздействиям внешней среды, представляет собой актуальное направление в исследовании современных информационных технологий (ИТ). ИТ занимают все большее значение в производственной сфере [1], в информационном обеспечении киберфизических объектов [2], в реализации коммуникационных и вычислительных функций в сетевых и облачных структурах [3, 4]. Развитие наукоемкого производства и связанных с ним рабочих процессов, их планирование, стратегическое и оперативное управление ими могут потребовать цифровые форматы представления технологий в виде сетевых управляющих кодов, автоматизации, квантификации [5], что в

большинстве своем строится на инженерии данных (их извлечении, очистке, интеграции, интеллектуальном анализе).

Быстрое развитие методов интеллектуального анализа данных (ИАД) обусловлено широким внедрением информационных систем в процесс поддержки принятия решений (ППР) в организационно-технических системах в различных предметных областях, ростом самих данных. Ключевыми тенденциями в области современных ИТ выступают: интеграция методов искусственного интеллекта (ИИ) с технологиями обработки больших данных; методы предиктивной и прескриптивной аналитики, включающие предиктивное моделирование [6], развитие предиктивных методов на основе инструментов машинного обучения (МО) [7], автоматизированной генерации рекомендаций [8] для поддержки принятия решений, в том числе и в составе автоматизированных информационных систем различных предметных областей применения [9].

Методы предиктивной (предсказательной) аналитики из состава инструментария ИАД, применяемые для задач обеспечения информационной безопасности (ИБ) при всем своем разнообразии нацелены прежде всего на обеспечение качества получаемой информации (ее содержательной или семантической составляющей) [10, 11]. При этом основные усилия в обеспечении ИБ концентрируются на задачах обеспечения достоверности, конфиденциальности и сохранности информации (информационных массивов, баз данных), не исключая построения оценок возможных последствий или возможных рисков для критических инфраструктур и технологий [12].

Основное содержание общего подхода к прогнозированию в задачах обеспечения безопасности заключается в определении параметров, влияющих на прогнозируемое событие (угрозу безопасности, уязвимость), или определение признаков, «маяков», косвенно указывающих на риск реализации угрозы ИБ с определенной вероятностью. Целью исследования является рассмотрение инструментария интеллектуального анализа данных в составе методов предиктивной аналитики в обеспечении безопасности информационных технологий, ориентированных на ППР на основе структуризации гибридных моделей.

Актуальность обеспечения безопасности использования ИТ

Состояние защищенности информации (данных), при котором обеспечены ее (их) конфиденциальность, доступность и целостность[1], по сути составляет понятие безопасности информации или данных. Для состояния защищенности информационной технологии важно обеспечение[2]: а) безопасности информации (данных), для обработки которой применяется ИТ; б) безопасности самой информационной системы, в которой ИТ реализована.

Первичным источником информации, как правило, выступают данные (измерительные, результаты обработки первичных данных и др.). Данные зачастую выступают носителями информации. Окончательное формирование информации осуществляется посредством переработки данных, в ходе которой может быть осуществлено одно или несколько следующих действий: категоризация (классификация) данных; контекстуализация; коррекция; агрегирование/декомпозиция (группирование, сжатие, усреднение и др.).

Большие данные, пулы данных, ИАД, аналитика данных, наука о данных – это очень популярные сегодня термины, которые означают, что появилась новая отрасль науки, которая революционизирует большинство ИТ, организационную основу которой составляет координация данных (D), информации (I) и знаний (K), а также адаптация информационных массивов, ориентированная на максимизацию информационной эффективности [13].

Указанные составляющие, известные как элементы «пирамиды знаний» или DIKW – модели (данные, информация, знания, мудрость), представляемой в виде четырехуровневой

[1] ГОСТ Р 50922–2006 Защита информации. Основные термины и определения. – М.: Стандартинформ, 2008. (Дата введения: 2008–02–01); (п. 2.4.5).

[2] ГОСТ Р 53114–2008 Защита информации. Обеспечение информационной безопасности в организации. Основные термины и определения. – М.: Стандартинформ, 2018. (Дата введения: 2009–10–01); (п. 3.1.2).

иерархии, где каждый уровень добавляет определенные атрибуты к сведениям предыдущего уровня[3]. Данные (*data*) – это самый базовый уровень. Информация (*information*) добавляет контекст, знание (*knowledge*) показывает, как их (данные или информацию) использовать, а мудрость (*wisdom*) показывает, когда и зачем их использовать, что чаще применяется на этапе выработки управляющих решений.

Мудрость имеет дело с ценностями. Она предполагает использование логики рассуждений. Все оценки эффективности основаны на логике, которая в принципе может быть запрограммирована в компьютере и автоматизирована. Эти принципы оценки являются безличными. Мы можем говорить об эффективности действия как о степени достижения некоторой цели (исследования, моделирования и др.).

ИАД (Data Mining) – это процесс поддержки принятия решений, основанный на поиске в данных скрытых закономерностей (шаблонов информации). При этом накопленные сведения автоматически обобщаются до информации, которая может быть охарактеризована как знания. ИАД обычно включает следующие этапы:

1) выявление закономерностей (свободный поиск);

2) использование выявленных закономерностей для предсказания неизвестных значений (прогностическое моделирование);

3) анализ исключений, предназначенный для выявления и толкования аномалий в найденных закономерностях.

Особенностью аналитики больших данных является то, что они, с одной стороны выступают источником получаемого многообразия информации, а с другой стороны, источником расширенных возможностей для формирования сложных киберугроз. Вот почему поиск эффективных методов проактивной безопасности является актуальной задачей.

Методы прогнозной аналитики позволяют реализовать гибридные модели, объединяющие типовые информационные процессы, связанные со сбором, хранением, переработкой данных (см. рис. 1), формировать оценочные модели под конкретные типы угроз, а также под предполагаемые кибервоздействия.

Рис. 1. Система прогнозной аналитики задач безопасности

Структуризация аналитики данных в контексте безопасности

Смена парадигм в аналитической практике анализа данных сопровождается развитием новых подходов, методов работы с данными. Смена парадигм в аналитической практике постоянно эволюционирует, обеспечивая решение все более сложных задач:

– дескриптивная (описательная) аналитика и составление отчетов;

– диагностическая аналитика (выявление закономерностей);

[3] Ackoff R.L. From data to wisdom // Journal of Applied Systems Analysis. 1989. Vol. 16. P. 3–9.

– предиктивная (прогнозная) аналитика;

– предписывающая аналитика и принятие решений.

Результаты дескриптивной аналитики – получение ретроспективных обобщений и описаний повседневных процессов: сбор и систематизация данных, выявление фактов корреляции между данными, а также обнаружение исключительных ситуаций (отказы, аварии, инциденты), предварительная оценка возможных последствий. Для своих моделей предписывающая аналитика собирает данные из нескольких описательных и прогностических источников и применяет их к процессу принятия решений. Таким образом, мы можем сказать, что она относится как к описательной, так и к прогностической аналитике.

В табл. 1 представлены сравнительные характеристики указанных аналитических методов, составленных на основе [14].

Таблица 1

Различные типы аналитических методов с примерами

Аналитические методы	Построение модели на основе данных	Примеры
Описательная аналитика	Отвечает на вопрос: «Что происходило в прошлом?»	Обобщает прошлые события, например результаты измерений, бизнес-данные, использование социальных сетей, сообщает об общих тенденциях и т.д.
Диагностическая аналитика	Отвечает на вопрос: «Почему это произошло?»	Выявляет: аномалии и случайные связи, потери в бизнес-процессах, влияние программы лечения и т.д.
Прогностическая аналитика	Отвечает на вопрос: «Что произойдет в будущем?»	Прогнозирование: состояния оборудования, предпочтений клиентов, выявления возможных нарушений безопасности, потребностей в ресурсах и др.
Предписывающая аналитика	Отвечает на вопрос: «Какие действия следует предпринять?»	Улучшает: управление бизнесом; техническим состоянием; уход за пациентами и администрирование здравоохранения; маркетинговые стратегии и др.

Динамика взаимодействия этапов в сопоставлении с рабочими (типовыми) процессами ППР представлена на рис. 2.

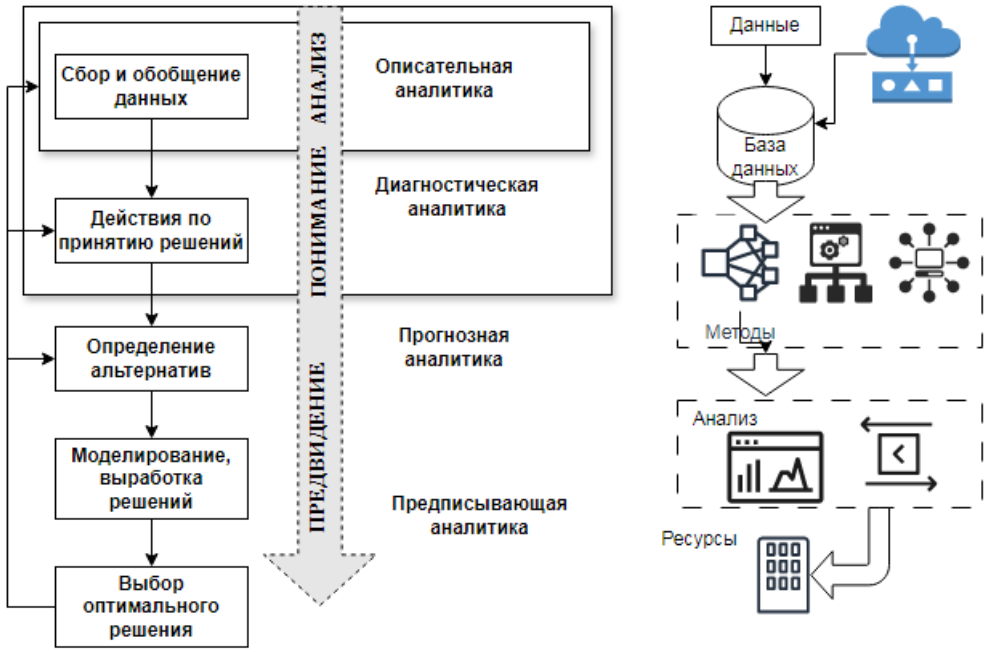

Рис. 2. Этапы и задачи рабочего процесса принятия решений

Предиктивная аналитика (ПдА) – это основной этап, на котором предсказываются будущие события, результаты, будущие риски и исключительные ситуации. В части аналитики на этом этапе активно применяются методы МО. В ходе предиктивного моделирования применяются следующие методы: классификация; нейронные сети, метод опорных векторов; байесовские методы; методы на основе паттернов и ряд других.

Предписывающая (прескриптивная) аналитика (ПсА) уже успешно применяется во многих промышленных и исследовательских проектах. Исследователю (ЛПР – лицу, принимающему решение) предлагается наилучший вариант адаптации и формирования планов в соответствии с прогнозируемым будущим. Это позволяет ЛПР не только выявлять проблемы и возможности (анализируя прошлое, настоящее или будущее), но и непосредственно предписывать наилучшие варианты решений в соответствии с определенными целями и оценивать их результаты (оценка моделей – см. рис. 1).

Хотя методы оптимизации уже являются хорошо зарекомендовавшим себя и широко применяемым способом решения задач принятия решений с использованием математических инструментов, именно сочетание прогнозирования и оптимизации открывает новые возможности для ППР. Более того, поскольку ПсА часто применяется в реальных случаях с большой долей неопределенности, оптимизация в значительной степени зависит от точности прогнозов а в некоторых случаях от способности количественно оценить неопределенность прогнозов. В подобных ситуациях оптимизация должна учитывать факторы неопределенности, присущие предметной области, обусловленные (комбинированным) использованием статистических и имитационных моделей, т.е. в составе гибридных моделей.

Состав этапов процесса принятия решений (см. рис. 2) в сочетании с аналитическими задачами:

1. Сбор и обобщение данных. Данные собираются и хранятся, например, в СУБД, таблицах данных, распределенных файловых системах и т.д. Очистка и консолидация данных для устранения ошибок и несоответствий или ИАД для извлечения информации и функций являются обычными операциями в рамках этой задачи.

2. Получение предварительных прогнозов на основе алгоритмов принятия решений. Используя методы МО (например, прогнозирование временны́х рядов, марковские модели), целью этой задачи являются выполнение исследовательского анализа события и тенденции, которые влияют на процесс принятия решений путем анализа исторических данных. По сути, это промежуточное действие, которое аналитик может выполнять для получения начальной точки для модели оптимизации (или прогноза) относительно будущего процесса ППР.

3. Определение альтернативных решений. На основании имеющихся целей, ограничений, правил (критериев оптимизации) задачи принятия решения. Для поддержки этой задачи выполняется анализ процесса принятия решений с использованием методов МО, интеллектуального анализа данных и процессов. Цель – определить альтернативные варианты принятия решений, а также соответствующие функции затрат/выгод и связанные с ними ограничения.

4. Моделирование, выработка решений. Необходимо оценить влияние предлагаемого варианта решения на процесс в целом. Эта задача в математическом обеспечении связана с задачей п. 2, где могут быть использованы модели прогнозирования и имитационного моделирования, которые позволят смоделировать поведение системы при различных настройках (альтернативные варианты применения).

5. Выбор оптимального решения. На основании предварительного анализа (задачи п. 2) строятся варианты оценок в ходе моделирования (задача п. 4). Используя методы оптимизация, теории игр, осуществляют ранжирование полученных вариантов, из которых на основании имеющихся критериев точности, оперативности, экономичности выбирают наиболее рациональный вариант (план действий, сценарий управления).

Если ни один из вариантов не устраивает исследователей или появились новые инциденты, процесс принятия решений переходит к новой итерации (переход к задаче п. 1).

Таким образом, рассмотренные этапы формирования оптимального решения можно трактовать как процесс обучения модели принятия решений, где ресурсы – это базы данных, инцидентов, базы знаний – это имеющиеся методы интеллектуального анализа.

В табл. 2 представлен сравнительный анализ традиционной и предиктивной аналитики на основе основных факторов, учитываемых в работе аналитиков в ходе обеспечения ИБ [15]. Большинство факторов характерны для условий использования больших данных.

Таблица 2

Сравнение традиционной и предиктивной аналитики на примере задач ИБ

Основные факторы	Традиционная аналитика	Прогнозирующая аналитика
Подход	Реагирующий, фокусируется на анализе прошлых событий, чтобы понять, что произошло	Проактивный, фокусируется на выявлении закономерностей и прогнозировании будущих угроз
Источники данных	Использует структурированные статические данные, такие как системные журналы и исторические отчеты	Использует как структурированные, так и неструктурированные данные, включая информацию в режиме реального времени, поведенческие модели и данные
Обнаружение угроз ИБ	Ограничивается известными угрозами с использованием предопределенных сигнатур и правил	Позволяет обнаруживать как известные, так и неизвестные угрозы с помощью МО и обнаружения аномалий
Время отклика	Отложенный ответ, поскольку анализ выполняется после того, как угроза материализовалась	Позволяет обнаруживать угрозы в режиме реального времени и быстрее реагировать на них, сводя к минимуму их воздействие
Адаптивность (по типу модели)	Статические модели, которые требуют обновления вручную для устранения новых угроз	Динамические модели, которые развиваются с изменением ландшафта угроз и со временем совершенствуются благодаря непрерывному обучению
Масштабируемость	Проблемы с обработкой больших и сложных наборов данных	Использует технологии больших данных для эффективного анализа обширных и разнообразных наборов данных
Принятие решений	В значительной степени зависит от анализа ЛПР	Автоматизирует процессы принятия решений с помощью искусственного интеллекта и МО
Примеры использования в задачах ИБ	Обнаружение вредоносных программ с помощью сопоставления сигнатур	Обнаружение мошенничества, прогнозирование внутренних угроз и обнаружение аномалий в режиме реального времени
Эффективность в борьбе с возникающими угрозами	Ограниченная, поскольку требуется адаптация к возникающим угрозам и существенное ручное вмешательство	Высокая, поскольку модели постоянно обучаются выявлять и адаптировать к меняющимся тактикам, методам и процедурам кибер-вмешательства

Заключение

Предиктивная аналитика может существенно повлиять на развитие технологий, связанных с данными, выявлением знаний из данных, с пониманием вероятностей будущих событий, предсказанием исключительных событий, сценариев развития чрезвычайных ситуаций.

Прогнозная аналитика играет центральную роль в анализе киберугроз, фокусируя имеющийся арсенал средств на выявлении и снижении потенциальных киберрисков до того, как они материализуются. Анализируя тенденции, поведение и аномалии в сетевом трафике и

активности пользователей, прогнозная аналитика позволяет предвидеть атаки, эффективно распределять ресурсы и минимизировать время простоя. Интеграция возможностей больших данных, методов машинного обучения, технологий на основе нейронных сетей не только повышает точность обнаружения угроз, но и сокращает время на реагирование, что является критическим фактором для минимизации ущерба.

Наука о данных трансформирует отрасли промышленности во всем мире. Это критически важно для будущего информационных систем и технологий, а также сервисов ИБ, поскольку «безопасность – это прежде всего данные», которые определяют успех практически в любой современной предметной области.

Следует отметить, что существующие тенденции и подходы к вопросам безопасности все больше концентрируются не на безопасности концептуального пространства (защите структур и хранилищ данных и информации с помощью шифрования, контроля доступа и других технологий), а на безопасности семантического пространства, связанного с содержательными свойствами информации, на безопасности когнитивного пространства, связанного с выявлением новых знаний, получаемых в результате переработки данных на всех уровнях, например DIKW – моделей.

Часть вопросов в данном направлении снимает предиктивная аналитика, но многие из вопросов требуют дальнейших исследований путем развития комплексных подходов, интеграции знаний, определяя целевые факторы технологического развития.

Список литературы

1. Левенцов В.А., Костецкий Д.Ю., Аркина К.Г. Разработка интегрированного стандарта обеспечения цифровыми двойниками наукоемкого производства // Известия Санкт-Петербургского государственного экономического университета. 2021. № 1(127). С. 105–115. /
   Leventsov V.A., Kostetskij D.Yu., Arkina K.G. Razrabotka integrirovannogo standarta obespecheniya cifrovymi dvojnikami naukoemkogo proizvodstva. Izvestiya Sankt-Peterburgskogo gosudarstvennogo ekonomicheskogo universiteta. 2021;1(127):105–115. (In Russ.).
2. Грошев А.В. Стратегия алгоритмического повышения точностных характеристик и информационной надежности инерциально-спутниковых навигационных систем в составе беспилотных летательных аппаратов // Труды МАИ. 2019. № 104. С. 16. /
   Groshev A.V. Strategiya algoritmicheskogo povysheniya tochnostnyh harakteristik i informacionnoj nadezhnosti inercial'no-sputnikovyh navigacionnyh sistem v sostave bespilotnyh letatel'nyh apparatov. Trudy MAI. 2019;104:16. (In Russ.).
3. Бурый А.С. Информационное пространство сетевого взаимодействия в клиентской среде // Транспортное дело России. 2011. № 8. С. 156–157. /
   Buryi A.S. Informacionnoe prostranstvo setevogo vzaimodejstviya v klientskoj srede. Transport Business of Russia. 2011;8:156–157. (In Russ.).
4. Бурый А.С. Тенденции развития распределенных информационных систем на основе облачных технологий // Транспортное дело России. 2013. № 6. С. 160–162. /
   Buryi A.S. Tendencii razvitiya raspredelennyh informacionnyh sistem na osnove oblachnyh tekhnologij. Transport Business of Russia. 2013;6:160–162. (In Russ.).
5. Цао Л. Образ мышления в науке о данных: Наступающая научно-техническая и экономическая революция. – СПб.: Изд-во Европейского университета в Санкт-Петербурге, 2022. – 552 с. /
   Cao L. Data Science Thinking the Next Scientific, Technological and Economic Revolution. St. Petersburg: Publ. House of the European University, 2022, 552 p. (In Russ.).
6. Колесников И.Н., Финогеев А.Г. Проактивный мониторинг событий на основе предиктивного анализа временных рядов // Модели, системы, сети в экономике, технике, природе и обществе. 2020. № 1(33). С. 111–125. /
   Kolesnikov I.N., Finogeev A.G. Proaktivnyj monitoring sobytij na osnove prediktivnogo analiza vremennyh ryadov. Modeli, sistemy, seti v ekonomike, tekhnike, prirode i obshchestve. 2020;1(33):111–125. (In Russ.).
7. Леонов Н.В. Противодействие уязвимостям программного обеспечения. Ч. 1. Онтологическая модель // Вопросы кибербезопасности. 2024. № 2(60). С. 87–92. – DOI 10.21681/2311–3456–2024–2–87–92.
   Leonov N.V. Protivodejstvie uyazvimostyam programmnogo obespecheniya. Part 1. Ontologicheskaya model. Cybersecurity issues. 2024;2(60):87–92. (In Russ.).
8. Бурый А.С., Усцелемов В.Н. Развитие систем поддержки принятия решений на основе прецедентов в распределенных информационных структурах // Правовая информатика. 2024. № 1. С. 88–95. /
   Buryi A.S., Ustselemov V.N. Razvitie sistem podderzhki prinyatiya reshenij na osnove precedentov v raspredelennyh informacionnyh strukturah. Legal Informatics. 2024;1:88–95. (In Russ.).

9. Бурый А.С., Усцелемов В.Н. Информационная безопасность автоматизированных систем // Информационно-экономические аспекты стандартизации и технического регулирования. 2023. № 2(72). С. 31–37. /
Buryi A.S., Ustselemov V.N. Information security of automated systems. Information and Economic Aspects of Standardization and Technical Regulation. 2023;2(72):31–37. (In Russ.).
10. Бурый А.С., Полоус А.И. Качество информации в организационно-технических системах управления // Транспортное дело России. 2012. № 6-2. С. 82–87. /
Buryi A.S., Polous A.I. Kachestvo informacii v organizacionno-tekhnicheskih sistemah upravleniya. Transport Business of Russia. 2012;6-2:. 82–87. (In Russ.).
11. Ловцов Д.А. Теория защищенности информации в эргасистемах: монография. – М.: РГУП, 2022. – 276 с. /
Lovtsov D.A. Teoriya zashchishchennosti informacii v ergasistemah: monografiya. Moscow: RGUP Publ., 2022, 276 p.
12. Малюк А.А. Основы политики безопасности критических систем информационной инфраструктуры. – М.: Горячая линия-Телеком, 2019. – 314 с. /
Malyuk A.A. Osnovy politiki bezopasnosti kriticheskih sistem informacionnoj infrastruktury. Moscow: Goryachaya liniya-Telekom Publ., 2019, 314 p. (In Russ.).
13. Stafford S.P. Data, information, knowledge, and wisdom // Knowledge management, organizational intelligence and learning, and complexity. 2009. Т. 3. С. 179.
14. Sarker I.H. Data science and analytics: an overview from data-driven smart computing, decision-making and applications perspective // SN Computer Science. 2021. Т. 2, № 5. С. 377.
15. Ekundayo F., Atoyebi I., et al. Predictive Analytics for Cyber Threat Intelligence in Fintech Using Big Data and Machine Learning // International Journal of Research Publication and Reviews. 2024. Т. 5, № 11. С. 5934–5948.