

УДК 004.056.5

DOI: [10.26102/2310-6018/2026.61.10.005](https://doi.org/10.26102/2310-6018/2026.61.10.005)

Тренды и вызовы: методы искусственного интеллекта в кибербезопасности

М.Б. Гузаиров, С.С. Валеев✉, Н.В. Кондратьева

Уфимский университет науки и технологий, Уфа, Российская Федерация

Резюме. В статье проводится анализ современного состояния методов искусственного интеллекта в области кибербезопасности. Обсуждается принцип интеграции искусственного интеллекта в контур безопасности на основе концепции «человек в контуре» (Human-in-the-loop). Отмечено, что развитие программных мультиагентных систем позволяет создавать распределенные экосистемы информационной защиты. Анализируются методы глубокого обучения (Deep Learning), которые являются технологическим фундаментом современного искусственного интеллекта в кибербезопасности благодаря способности автоматически извлекать иерархические признаки из сырых данных. Показано, что для повышения эффективности, надежности и интерпретируемости активно применяются гибридные архитектуры, сочетающие сильные стороны различных нейросетевых парадигм. Также искусственный интеллект выступает активным компонентом управления операционной устойчивостью, то есть способностью организации выполнять свои ключевые функции в условиях воздействия кибератак или сбоя некоторых подсистем инфраструктуры. Выявлены ключевые тренды и вызовы развития искусственного интеллекта в сфере кибербезопасности. Среди перспективных направлений развития выделены объяснимый искусственный интеллект, федеративное обучение, большие языковые модели, защита от состязательного искусственного интеллекта. К основным трудностям внедрения относятся масштабируемость в реальном времени, эволюция атак и операционная сложность.

Ключевые слова: искусственный интеллект, кибербезопасность, операционная устойчивость, глубокое обучение, обнаружение аномалий.

Для цитирования: Гузаиров М.Б., Валеев С.С., Кондратьева Н.В. Тренды и вызовы: методы искусственного интеллекта в кибербезопасности. *Моделирование, оптимизация и информационные технологии*. 2026;14(10). URL: <https://moitvvt.ru/ru/journal/article?id=2578>
DOI: 10.26102/2310-6018/2026.61.10.005

Trends and challenges: artificial intelligence methods in cybersecurity

M.B. Guzairov, S.S. Valeev✉, N.V. Kondratyeva

Ufa University of Science and Technology, Ufa, the Russian Federation

Abstract. This article analyzes the current state of artificial intelligence in cybersecurity. It discusses the principle of integrating artificial intelligence into the security environment based on the "human-in-the-loop" concept. It is noted that the development of software multi-agent systems enables the creation of distributed information security ecosystems. Deep learning methods are analyzed, which form the technological foundation of modern artificial intelligence in cybersecurity due to their ability to automatically extract hierarchical features from raw data. It is shown that hybrid architectures combining the strengths of various neural network paradigms are actively used to improve efficiency, reliability, and interpretability. Artificial intelligence also serves as an active component in operational resilience management, that is, the ability of an organization to perform its key functions in the face of cyberattacks or the failure of certain infrastructure subsystems. Key trends and challenges in the development of artificial intelligence in cybersecurity are identified. Promising areas of development include

explainable artificial intelligence, federated learning, large-scale language models, and protection against adversarial artificial intelligence. Key implementation challenges include real-time scalability, attack evolution, and operational complexity.

Keywords: artificial intelligence, cybersecurity, operational resilience, deep learning, anomaly detection.

For citation: Guzairov M.B., Valeev S.S., Kondratyeva N.V. Trends and challenges: artificial intelligence methods in cybersecurity. *Modeling, Optimization and Information Technology*. 2026;14(10). (In Russ.). URL: <https://moitvvt.ru/ru/journal/article?id=2578> DOI: 10.26102/2310-6018/2026.61.10.005

Введение

Современные системы кибербезопасности (КБ) переходят от применения реактивной модели при обнаружении угроз («обнаружил – отреагировал») к применению проактивной модели реагирования [1, 2]. Традиционно используемые сигнатурные методы обнаружения угроз не всегда могут эффективно решать свои задачи в условиях динамически меняющегося ландшафта атак [3, 4]. В качестве эффективного решения используются методы искусственного интеллекта (ИИ), позволяющие на их основе обрабатывать большие объемы входящих данных в режиме реального времени, а также выявлять скрытые закономерности [5, 6].

Основным принципом интеграции ИИ в контур безопасности является использование концепции «человек в контуре» (Human-in-the-loop) [1]. В этом случае предполагается разделение ответственности между ИИ и человеком: алгоритмические системы берут на себя рутинные операции по мониторингу, корреляции событий и первичной классификации инцидентов, тем самым, освобождая время и силы экспертов-аналитиков для эффективного принятия стратегических решений и тщательного расследования сложных случаев КБ. Кроме того, развитие программных мультиагентных систем позволяет создавать распределенные экосистемы информационной защиты, где множество автономных программных ИИ-агентов координируют свои действия для обеспечения эшелонированной системы защиты [7, 8].

При анализе источников применена методика PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) [9, 10], адаптированная для области КБ.

Критерии включения публикаций:

- публикации за 2019–2026 гг. в рецензируемых журналах (Q1–Q2) и трудах конференций (A*–A);
- исследования, содержащие экспериментальную оценку методов ИИ для обнаружения атак или аномалий;
- работы, в которых сравниваются как минимум две архитектуры нейронных сетей.

Источники публикаций – открытые депозитарии статей профильных журналов. Процедура отбора:

- первичный поиск – 124 статьи;
- скрининг по заголовкам и аннотациям – 41 статья;
- полнотекстовый анализ – 97 статей;
- итоговый пул – 41 статья, отвечающая заданным критериям.

Кластерный анализ публикаций (с использованием LDA – Latent Dirichlet Allocation) выявил 6 основных тематических кластеров:

- CNN для анализа изображений вредоносного ПО;
- LSTM/GRU для временных рядов сетевого трафика;

- GNN для графовых структур (коммуникации, зависимости);
- гибридные модели (CNN+LSTM, CNN+GNN и др.);
- объяснимый ИИ (XAI) в КБ;
- федеративное и состязательное обучение.

Как показал анализ, лишь 12 % работ содержат прямое сравнение гибридных архитектур между собой на единых датасетах, а 70 % исследований не учитывают временные затраты на обучение и инференс.

Как отмечалось ранее, современный ландшафт киберугроз характеризуется ростом числа сложных многостадийных атак, использующих методы обхода сигнатурных средств защиты [10, 11]. Традиционные подходы, основанные на правилах и сигнатурах, демонстрируют низкую эффективность при обнаружении zero-day уязвимостей и полиморфных вредоносных программ [12]. В ответ на эти вызовы активно внедряются методы ИИ, способные выявлять скрытые закономерности в больших потоках данных [13, 14]. Однако, несмотря на обилие публикаций, практическая реализация ИИ-систем сталкивается с рядом нерешенных проблем:

- Отсутствие универсальной архитектуры – ни один из существующих методов (CNN, LSTM, GNN, трансформеры) не является оптимальным для всех типов атак [15].
- Сложность интерпретируемости – модели «черного ящика» не позволяют аналитикам обоснованно принимать решения [16].

Таблица 1 – Классификация типов гибридизации
Table 1 – Classification of hybridization types

Тип гибридизации	Примеры реализаций	Описание	Преимущества	Недостатки
Последовательная (конвейерная)	Выход одной сети подается на вход другой	Первый этап – извлечение признаков, второй – классификация CNN (извлечение) + LSTM (классификация), AE (сжатие) + SVM	Простота калибровки, модульность	Накопление ошибок, отсутствие обратной связи
Параллельная (ансамблевая)	Несколько моделей работают независимо, решения агрегируются (голосование, усреднение)	CNN + LSTM + GNN с мета-классификатором (Random Forest)	Робастность, устойчивость к шуму	Высокие вычислительные затраты, сложность интерпретации
Иерархическая (вложенная)	Модели разных уровней обрабатывают разные аспекты данных (например, пакеты → потоки → сессии)	Иерархический LSTM (пакетный уровень, потоковый уровень)	Учет многоуровневой структуры атак	Сложность обучения, требует больших данных
Гибридная с обратной связью (циклическая)	Результаты классификации используются для дообучения компонентов в реальном времени (Human-in-the-loop)	CNN + LSTM + модуль экспертных правил	Адаптивность к новым атакам	Высокая сложность реализации

– Высокая вычислительная сложность – обработка трафика в реальном времени требует значительных ресурсов [17].

– Изменение сложности атак – злоумышленники активно используют ИИ для генерации обходных векторов [18].

Целью является систематический анализ методов ИИ в КБ, разработка формальной классификации гибридных архитектур, выполнение сравнительной экспериментальной оценки их эффективности и формулирование практических рекомендаций по выбору и внедрению. В рамках данного исследования выполнены:

– классификация гибридных пайплайнов по типу интеграции;
– количественное сравнение производительности на репрезентативных датасетах.

На основе анализа литературы мы предлагаем классификацию гибридных систем по «способу интеграции компонентов» (Таблица 1).

Рассмотрим далее особенности основных методов ИИ, используемых в современных системах обеспечения КБ.

Материалы и методы

Глубокое обучение как основа анализа больших потоков данных. Методы глубокого обучения (Deep Learning) являются технологическим фундаментом современного ИИ в КБ благодаря способности автоматически извлекать иерархические признаки из сырых, неструктурированных данных, что недоступно в рамках классических алгоритмов машинного обучения [19, 20]. Результаты анализа современных методов ИИ, используемых в КБ, представлены в виде онтологии (Рисунок 1).



Рисунок 1 – Архитектуры нейронных сетей и их применение в КБ
Figure 1 – Neural network architectures and their applications in cybersecurity

К основным проблемам внедрения методов глубокого обучения в КБ можно отнести:

– Недостаточные объемы размеченных данных, используемых при обучении нейронных сетей, т. е. формирование качественных датасетов об атаках затруднено из-

за их случайного появления и требований соблюдения политики конфиденциальности [21, 22].

– Проблема объяснимости результатов (Explainable AI-XAI), так как методы «черного ящика» осложняют работу аналитиков безопасности и требуются обоснования при разработке мер противодействия киберугроз. Следует отметить, что в настоящее время для решения проблемы активно ведутся разработки методов визуализации карт атак (*attention maps*) и локальных интерпретируемых моделей (LIME, SHAP) [23, 24].

– Достаточно высокая вычислительная сложность этих методов, тем самым, требования к вычислительным ресурсам необходимых для анализа высокоскоростных потоков данных остаются очень высокими [25, 26].

Гибридные архитектуры как метод повышения эффективности. Ни один из существующих методов ИИ не является универсальным решением для всех задач КБ. Для повышения эффективности, надежности и интерпретируемости активно развиваются гибридные архитектуры, сочетающие сильные стороны различных нейросетевых парадигм. На Рисунке 2 представлены результаты анализа гибридных ИИ и их применения в КБ в виде онтологии.

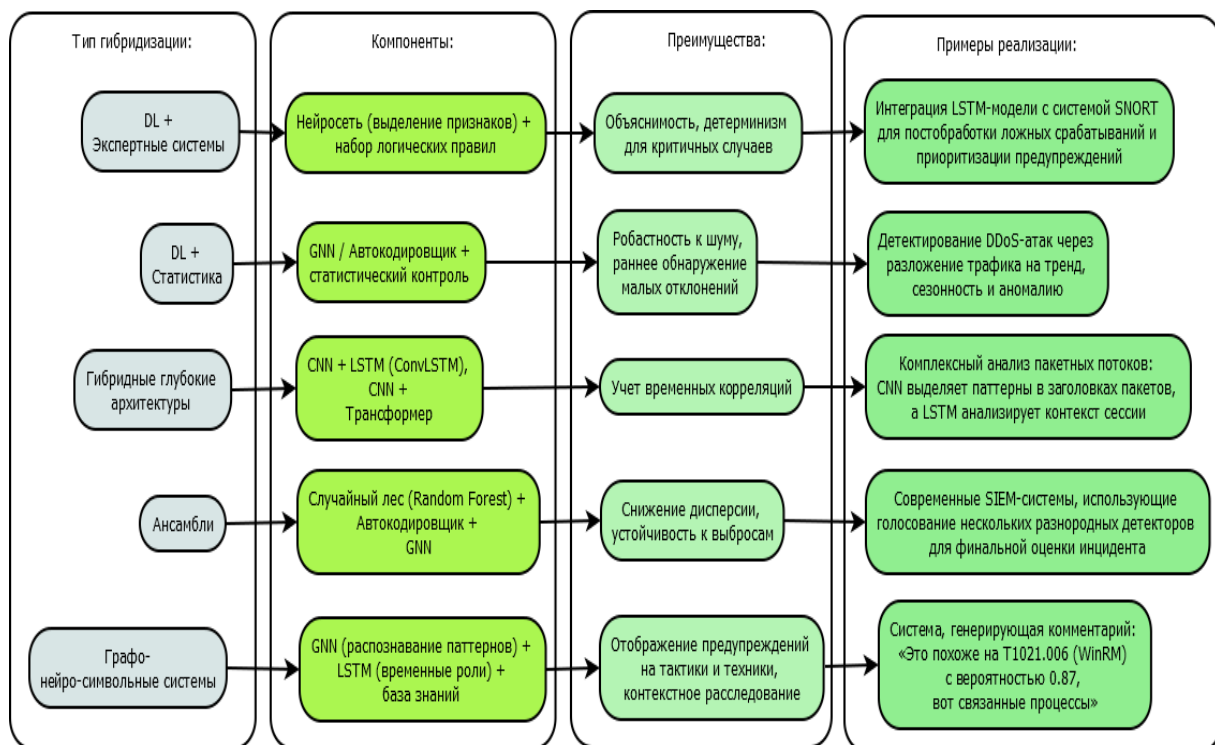


Рисунок 2 – Типы гибридизации и примеры реализации
Figure 2 – Types of hybridization and implementation examples

Ниже приведена обобщенная последовательность шагов при анализе угроз:

Этап 1. Первичная фильтрация. Сети CNN/LSTM анализируют поток данных и выполняет маркировку аномального события.

Этап 2. Классификация угрозы. Графовый классификатор определяет тип атаки.

Этап 3. Формирование рекомендаций. Экспертная система или база знаний (MITRE, ФСТЭК) предлагает план реагирования.

Этап 4. Обратная связь. Действия эксперта используются для дообучения всей системы (Human-in-the-loop) [27, 28].

Операционная устойчивость и адаптивность. Операционная устойчивость (Operational Resilience) – это способность организации выполнять свои ключевые функции даже в условиях воздействия кибератак или сбоя некоторых подсистем инфраструктуры. В этом контексте методы ИИ перестают быть просто инструментом КБ, а становятся активным компонентом управления операционной устойчивостью [29, 30]. На Рисунке 3 представлены результаты анализа роли ИИ в обеспечении операционной устойчивости организации.

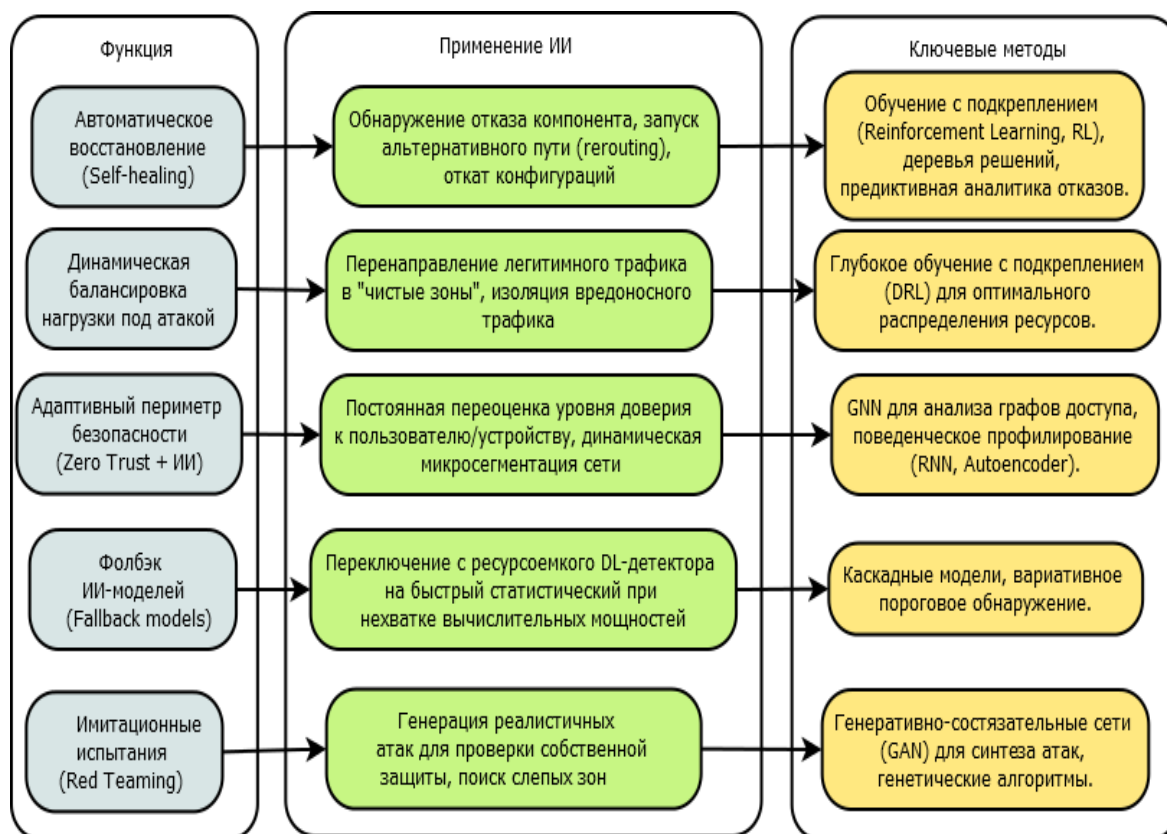


Рисунок 3 – Онтология применения методов ИИ для обеспечения устойчивости организации
Figure 3 – Ontology of AI applications for organizational resilience

Для оценки эффективности использования методов ИИ для обеспечения устойчивости организации используют ряд следующих метрик:

- метрика учета времени восстановления в случае воздействия атаки – RTO (Recovery Time Objective);
- метрика учета минимизации возможных потерь данных в случае использования прогноза предсказания возможной атаки на активы организации – RPO (Recovery Point Objective);
- метрика способности поддерживать точность на наборах данных о zero-day уязвимостях, отражающая адаптивность к новым видам атак [31, 32].

Направления и особенности применения ИИ в КБ. Анализ текущего состояния методов ИИ в КБ и перспектив развития ИИ позволяет определить основные направления применения ИИ в рассматриваемой области.

К основным перспективным направлениям можно отнести:

- Развитие технологий, позволяющих объяснять решения, полученные с помощью технологий ИИ (LIME, SHAP, карты внимания) [16, 33].

- Применение технологий федеративного обучения, позволяющих обучать ансамбли моделей ИИ на наборах данных из разных организаций без прямого обмена самими данными, что позволяет в ряде случаев решать задачи обеспечения конфиденциальности.

- Использование больших языковых моделей (LLM) для генерации отчетов о зарегистрированных инцидентах [34, 35].

- Разработка методов защиты от состязательного ИИ, т. е. обучение моделей на специально сгенерированных примерах для повышения робастности средств защиты [36, 37].

Особенности развития методов ИИ в КБ:

- Постоянная эволюция атак, совершаемых злоумышленниками, которые также могут использовать ИИ для создания новых методов обхода систем защиты [38].

- Повышение цены ошибок 2-го рода, т. е., например, пропуск атаки системой защиты информации может быть значительно дороже, чем реакция на ложную тревогу при реакции на ошибку 1-го рода [39].

- Сложность масштабируемости систем защиты в реальном времени. Например, обеспечение производительности систем защиты на основе методов ИИ требует повышения пропускной способности каналов связи и новых аппаратных средств.

- Отладка, обновление и сопровождение комплекса связанных между собой гибридных моделей ИИ требуют новых навыков и знаний у персонала, а также необходима разработка новых инструментальных средств [40].

Сравнительный экспериментальный анализ. В качестве наборов данных для анализа эффективности рассмотренных алгоритмов использовались следующие датасеты:

- CIC-IDS-2017 – включает атаки DDoS, брутфорс, веб-атаки;

- UNSW-NB15 – включает особенности эксплуатации, черви, бэкдоры.

Оценка качества базировалась на следующих метриках: точность (Accuracy), полнота (Recall), F1-мера. В качестве базовых сетей были использованы: сверточная нейронная сеть (CNN-1D), рекуррентная нейронная сеть (LSTM с 2 слоями), графовая нейронная сеть (GCN – Graph Convolutional Network). В качестве гибридных нейронных сетей использовались: CNN+LSTM (последовательная схема), CNN+GNN (параллельная схема), CNN+LSTM+GNN (иерархическая схема).

Результаты и обсуждение

Результаты численного моделирования представлены в Таблицах 2 и 3.

Анализ результатов моделирования показал следующее:

- Гибридная модель CNN+LSTM+GNN превосходит отдельные архитектуры по точности и полноте.

- Прирост в точности достигается за счет учета пространственных (CNN), временных (LSTM) и структурных (GNN) зависимостей, что критично для многостадийных атак.

Для гибридной модели на основе методики SHAP (SHapley Additive exPlanations) проведен анализ, который показал, что наиболее значимыми признаками являются:

- длительность соединения;

- количество переданных байтов в обратном направлении;

- вариативность времени между пакетами.

Это коррелирует с известными практиками MITRE ATT&CK (например, T1046 – Network Service Scanning).

Таблица 2 – Сравнительные результаты на CIC-IDS-2017
Table 2 – Comparative results on CIC-IDS-2017

Модель	Точность	Полнота	Инференс (мс/1k)	F1 - мера
CNN	0,912	0,887	0,899	28
LSTM	0,925	0,901	0,913	52
GCN	0,907	0,876	0,891	41
CNN+LSTM	0,951	0,934	0,942	67
CNN+GNN	0,938	0,919	0,928	59
CNN+LSTM+GNN	0,969	0,958	0,963	84

Таблица 3 – Сравнительные результаты на UNSW-NB15
Table 3 – Comparative results on UNSW-NB15

Модель	Точность	Полнота	Инференс (мс/1k)	F1 - мера
CNN	0,879	0,853	0,866	25
LSTM	0,891	0,868	0,879	48
GCN	0,862	0,831	0,846	39
CNN+LSTM	0,923	0,904	0,913	63
CNN+GNN	0,914	0,895	0,904	55
CNN+LSTM+GNN	0,947	0,931	0,939	79

На основе экспериментальных данных и литературного анализа можно сформулировать следующие рекомендации (Таблица 4).

Таблица 4 – Рекомендации по выбору архитектуры в зависимости от сценария
Table 4 – Recommendations for choosing an architecture depending on the scenario

Сценарий	Рекомендуемая архитектура	Обоснование
Обнаружение DDoS и флуда	CNN (1D) + LSTM (базовый)	Быстрый инференс, высокая точность на однотипных атаках
Выявление многостадийных атак	CNN+LSTM+GNN (иерархическая)	Учет временных и структурных связей, устойчивость к маскировке
Обнаружение аномалий в IoT-сетях	GCN (граф устройств) + автоэнкодер	Эффективность при разреженных данных
Реагирование с участием эксперта	Циклическая гибридная с обратной связью (Human-in-the-loop)	Возможность дообучения, интерпретируемость

Проведен систематический анализ методов ИИ в кибербезопасности, предложена новая классификация гибридных архитектур и выполнена их экспериментальная сравнительная оценка. Показано, что иерархическая комбинация CNN+LSTM+GNN обеспечивает наилучший баланс точности и полноты. Разработанный прототип включает модуль объяснимости (SHAP), что повышает доверие экспертов.

Ключевые вызовы и направления развития:

– Эволюция атак с использованием ИИ. Злоумышленники применяют генеративно-сопоставительные сети (GAN) для создания обходных сигнатур. Рекомендация: внедрение сопоставительного обучения в пайплайн защиты.

– Масштабируемость в реальном времени. Использование квантования, дистилляции знаний и аппаратных ускорителей (GPU/TPU).

– Объяснимость. Интеграция ХАИ-модулей (LIME, SHAP, карты внимания) обязательна для сертифицированных систем (например, ФСТЭК). Предложенная модель включает SHAP по умолчанию.

– Федеративное обучение. Позволяет обучать модели на данных нескольких организаций без раскрытия конфиденциальной информации. Рекомендуются для межкорпоративных систем защиты.

– Защита от состязательных примеров. Использование проекции градиентов и устойчивых функций потерь.

Предполагается, что дальнейшие исследования будут направлены на адаптацию предложенной модели для облачных сред и внедрение механизмов самоадаптации на основе обучения с подкреплением.

Заключение

Постоянная интеграция методов ИИ в системы информационной безопасности является современным перспективным направлением повышения эффективности средств защиты информации. Современные методы на основе глубокого обучения обеспечивают достаточно высокий уровень эффективности противодействия при обнаружении угроз для случаев известных паттернов, так и для случаев проявления аномалий. Рассмотренные гибридные архитектуры ИИ позволяют решать сложные задачи обеспечения точности, скорости и интерпретируемости получаемых результатов. Следует отметить, что конечной целью применения технологий ИИ в КБ является обеспечение системной операционной устойчивости, где методы ИИ функционируют как незаменимая часть замкнутого контура управления безопасностью организации. Проведенный анализ показывает, что на практике наиболее эффективными являются комплексные гибридные пайплайны, объединяющие сверточные нейронные сети для анализа контента, динамические сети на основе LSTM/GNN сетей для учета динамики процессов, а интерфейсы на базе LLM-моделей для повышения эффективности взаимодействия с аналитиком-специалистом.

СПИСОК ИСТОЧНИКОВ / REFERENCES

1. Olawade D.B., Plabon Sh.B., Ojo A., Ogunbona M.A., Makanjuola B.D., Olasilola O.R. Human in the loop artificial intelligence in healthcare: applications, outcomes, and implementation challenges. *International Journal of Medical Informatics*. 2026;213:106362. <https://doi.org/10.1016/j.ijmedinf.2026.106362>
2. Alqudhaibi A., Albarrak M., Aloose A., et al. Proactive cybersecurity in industry 4.0: a survey of cybersecurity threat prediction approaches in manufacturing systems. *International Journal of Information Security*. 2026;25(1):14. <https://doi.org/10.1007/s10207-025-01188-9>
3. Валеев С.С., Кондратьева Н.В., Мельников А.В. Архитектура предприятия и архитектура нулевого доверия. *Вестник УрФО. Безопасность в информационной сфере*. 2023;2(2):49–53. <https://doi.org/10.14529/secur230204>
Valeev S.S., Kondratyeva N.V., Melnikov A.V. Enterprise architecture and zero trust architecture. *Journal of the Ural Federal District. Information security*. 2023;2(2):49–53. (In Russ.). <https://doi.org/10.14529/secur230204>
4. Duran F., Conrad S.H., Conrad G.N., Duggan D.P., Held E.B. Building a system for insider security. *IEEE Security & Privacy*. 2009;7(6):30–38. <https://doi.org/10.1109/MSP.2009.111>
5. Alladi T., Chamola V., Zeadally Sh. Industrial control systems: Cyberattack trends and countermeasures. *Computer Communications*. 2020;155(22):1–9. <https://doi.org/10.1016/j.comcom.2020.03.007>

6. Валеев С.С., Кондратьева Н.В., Гузайров М.Б., Мельников А.В. Этапы реинжиниринга информационной системы предприятия в рамках технологии нулевого доверия. *Вестник Российского нового университета. Серия: Сложные системы: модели, анализ и управление*. 2023;(3):136–143.
Valeev S.S., Kondratyeva N.V., Guzairov M.B., Melnikov A.V. Stages of reengineering the information system of the enterprise within the framework of zero trust technology. *Vestnik of the Russian New University. Series: Complex Systems: models, analysis and management*. 2023;(3):136–143. (In Russ.).
7. Валеев С.С., Кондратьева Н.В., Гузайров М.Б., Исмагилова А.С. Иерархическая динамическая система управления информационной безопасностью информационной системы предприятия. *Инженерный вестник Дона*. 2023;(11). URL: <https://ivdon.ru/ru/magazine/archive/n11y2023/8802>
Valeev S.S., Kondratyeva N.V., Guzairov M.B., Ismagilova A.S. Hierarchical Dynamic Information Security Management System for an Enterprise Information System. *Engineering Journal of Don*. 2023;(11). (In Russ.). URL: <https://ivdon.ru/en/magazine/archive/n11y2023/8802>
8. Валеев С.С., Кондратьева Н.В., Гузайров М.Б., Исмагилова А.С. Многоагентные системы как технологическая база реализации концепции нулевого доверия. *Вестник Российского нового университета. Серия: Сложные системы: модели, анализ и управление*. 2024;(3):116–123.
Valeev S.S., Kondratyeva N.V., Guzairov M.B., Ismagilova A.S. Multi-agent systems as a technological basis for the implementation of zero trust concept. *Vestnik of the Russian New University. Series: Complex Systems: models, analysis and management*. 2024;(3):116–123. (In Russ.).
9. Page M.J., McKenzie J.E., Bossuyt P.M., et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *The BMJ*. 2021;372:n71. <https://doi.org/10.1136/bmj.n71>
10. Heckman K.E., Stech F.J., Thomas R.K., Schmoker B., Tsow A.W. Intrusions, Deception, and Campaigns. In: *Cyber Denial, Deception and Counter Deception: A Framework for Supporting Active Cyber Defense*. Cham: Springer; 2015. P. 31–52. https://doi.org/10.1007/978-3-319-25133-2_3
11. Ding D., Han Q.-L., Xiang Y., Ge X., Zhang X.-M. A survey on security control and attack detection for industrial cyber-physical systems. *Neurocomputing*. 2018;275:1674–1683. <https://doi.org/10.1016/j.neucom.2017.10.009>
12. Валеев С.С., Кондратьева Н.В. Особенности проектирования систем безопасности на базе архитектуры нулевого доверия. *Инженерный вестник Дона*. 2023;(8). URL: <https://ivdon.ru/ru/magazine/archive/n8y2023/8627>
Valeev S.S., Kondratyeva N.V. Features of designing security systems based on zero trust architecture. *Engineering Journal of Don*. 2023;(8). (In Russ.). URL: <https://ivdon.ru/en/magazine/archive/n8y2023/8627>
13. Yan J., Qi Y., Rao Q. Detecting malware with an ensemble method based on deep neural network. *Security and Communication Networks*. 2018;2018:7247095. <https://doi.org/10.1155/2018/7247095>
14. Venkatraman S., Alazab M., Ravi V. A hybrid deep learning image-based analysis for effective malware detection. *Journal of Information Security and Applications*. 2019;47:377–389. <https://doi.org/10.1016/j.jisa.2019.06.006>
15. Zhang J., Qin Zh., Yin H., Ou L., Xiao Sh., Hu Y. Malware variant detection using opcode image recognition with small training sets. In: *2016 25th International Conference on Computer Communication and Networks (ICCCN), 01–04 August 2016, Waikoloa, HI, USA*. IEEE; 2016. <https://doi.org/10.1109/ICCCN.2016.7568542>

16. Zhou G., Tian W., Buyya R., Xue R., Song L. Deep reinforcement learning-based methods for resource scheduling in cloud computing: a review and future directions. *Artificial Intelligence Review*. 2024;57(5):124. <https://doi.org/10.1007/s10462-024-10756-9>
17. Zohrevand Z., Glässer U. *Should I raise the red flag? A comprehensive survey of anomaly scoring methods toward mitigating false alarms*. arXiv. URL: <https://arxiv.org/abs/1904.06646> [Accessed 5th June 2026].
18. Park J., Kim D.S., Lim H. Privacy-preserving reinforcement learning using homomorphic encryption in cloud computing infrastructures. *IEEE Access*. 2020;8:203564–203579. <https://doi.org/10.1109/ACCESS.2020.3036899>
19. Thakur P., Kansal V., Rishiwal V. Hybrid deep learning approach based on LSTM and CNN for malware detection. *Wireless Personal Communications*. 2024;136(3):1879–1901. <https://doi.org/10.1007/s11277-024-11366-y>
20. Akandwanaho S.M., Kooblal M. Intelligent Malware Detection Using a Neural Network Ensemble Based on a Hybrid Search Mechanism. *The African Journal of Information and Communication*. 2019;(24). <https://doi.org/10.23962/10539/28660>
21. Yakura H., Shinozaki Sh., Nishimura R., Oyama Y., Sakuma J. Malware analysis of imaged binary samples by convolutional neural network with attention mechanism. In: *CODASPY '18: Proceedings of the Eighth ACM Conference on Data and Application Security and Privacy, 19–21 March 2018, Tempe, AZ, USA*. New York: ACM; 2018. P. 127–134. <https://doi.org/10.1145/3176258.3176335>
22. Yoo S., Kim S., Kang B.B. The image game: exploit kit detection based on recursive convolutional neural networks. *IEEE Access*. 2020;8:18808–18821. <https://doi.org/10.1109/ACCESS.2020.2967746>
23. Yazı A.F., Çatak F.Ö., Gül E. Classification of metamorphic malware with deep learning (LSTM). In: *2019 27th Signal Processing and Communications Applications Conference (SIU), 24–26 April 2019, Sivas, Turkey*. IEEE; 2019. <https://doi.org/10.1109/SIU.2019.8806571>
24. Yue S., Wang T. *Imbalanced malware images classification: a CNN based approach*. arXiv. URL: <https://arxiv.org/abs/1708.08042> [Accessed 5th June 2026].
25. Wang T., Xu N. Malware variants detection based on opcode image recognition in small training set. In: *Proceedings of the 2017 IEEE 2nd International Conference on Cloud Computing and Big Data Analysis (ICCCBDA), 28–30 April 2017, Chengdu, China*. IEEE; 2017. P. 328–332. <https://doi.org/10.1109/ICCCBDA.2017.7951933>
26. Khan R., Maynard P., McLaughlin K., Lavery D., Sezer S. Threat analysis of BlackEnergy malware for synchrophasor based real-time control and monitoring in smart grid. In: *ICS-CSR '16: Proceedings of the 4th International Symposium for ICS & SCADA Cyber Security Research 2016, 23–25 August 2016, Belfast, UK*. Swindon: BCS Learning & Development Ltd.; 2016. <https://doi.org/10.14236/ewic/ICS2016.7>
27. Crossman A., Plummer A.R., Sekharudu Ch., Warriar D., Yekrangian M. Auspex: Building threat modeling tradecraft into an artificial intelligence-based copilot. In: *2025 IEEE Conference on Artificial Intelligence (CAI), 05–07 May 2025, Santa Clara, CA, USA*. IEEE; 2025. P. 1160–1167. <https://doi.org/10.1109/CAI64502.2025.00201>
28. Bhardwaj A., Godika S., Loonker Y. *MALCDF: A Distributed Multi-Agent LLM Framework for Real-Time Cyber*. arXiv. URL: <https://arxiv.org/abs/2512.14846> [Accessed 8th June 2026].
29. Georgiadou A., Mouzakis S., Askounis D. Assessing MITRE ATT&CK Risk Using a Cyber-Security Culture Framework. *Sensors*. 2021;21(9):3267. <https://doi.org/10.3390/s21093267>

30. Friedberg I., Skopik F., Settanni G., Fiedler R. Combating advanced persistent threats: From network event correlation to incident detection. *Computers & Security*. 2014;48:35–57. <https://doi.org/10.1016/j.cose.2014.09.006>
31. Hundman K., Constantinou V., Laporte Ch., Colwell I., Soderstrom T. *Detecting spacecraft anomalies using LSTMs and nonparametric dynamic thresholding*. arXiv. URL: <https://arxiv.org/abs/1802.04431> [Accessed 8th June 2026].
32. Laheri R., Sukumar H.K.K., Kola Ch., Makin Y. Self-healing infrastructure: leveraging reinforcement learning for autonomous cloud recovery and enhanced resilience. *Journal of Information Systems Engineering and Management*. 2025;10(49s):352–357. <https://doi.org/10.52783/jisem.v10i49s.9888>
33. Mehdi Syed A.A., Anazagasty E. AI-driven infrastructure automation: leveraging AI and ML for self-healing and auto-scaling cloud environments. *International Journal of Artificial Intelligence, Data Science, and Machine Learning*. 2024;5(1):32–43. <https://doi.org/10.63282/3050-9262.IJAIDSML-V5I1P104>
34. Thein Th., Myo M.M., Parvin S., Gawanmeh A. Reinforcement learning based methodology for energy-efficient resource allocation in cloud data centers. *Journal of King Saud University – Computer and Information Sciences*. 2020;32(10):1127–1139. <https://doi.org/10.1016/j.jksuci.2018.11.005>
35. Mangalampalli S., Karri G.R., Kumar M., et al. DRLBTSA: Deep reinforcement learning based task-scheduling algorithm in cloud computing. *Multimedia Tools and Applications*. 2024;83(3):8359–8387. <https://doi.org/10.1007/s11042-023-16008-2>
36. Sharma A., Singh U.K. Modelling of smart risk assessment approach for cloud computing environment using AI and supervised machine learning algorithms. *Global Transitions Proceedings*. 2022;3(1):243–250. <https://doi.org/10.1016/j.gltp.2022.03.030>
37. Qu G., Wu H., Li R., Jiao P. DMRO: A deep meta reinforcement learning-based task offloading framework for edge-cloud computing. *IEEE Transactions on Network and Service Management*. 2021;18(3):3448–3459. <https://doi.org/10.1109/TNSM.2021.3087258>
38. Sheela M.S., Jayakanth J.J., Ramathilagam A., Gracewell J. Secure wireless sensor network transmission using reinforcement learning and homomorphic encryption. *International Journal of Data Science and Analytics*. 2025;20(3):2851–2870. <https://doi.org/10.1007/s41060-024-00633-7>
39. Joshi V., Patel P., Chandwani N., Bhatia J., Kumhar M. Intelligent auto-scaling in cloud infrastructure using machine learning and reinforcement learning. In: *Advances in Data-Driven Computing and Intelligent Systems: Selected Papers from ADCIS 2024: Volume 2, 20–21 September 2024, Goa, India*. Singapore: Springer; 2024. P. 217–239. https://doi.org/10.1007/978-981-96-4536-7_17
40. Jha N.N. Accelerating Cloud Outage Recovery Through Adaptive AI: A Reinforcement Learning Approach. *European Journal of Computer Science and Information Technology*. 2025;13(20):1–10. <https://doi.org/10.37745/ejcsit.2013/vol13n26110>

ИНФОРМАЦИЯ ОБ АВТОРАХ / INFORMATION ABOUT THE AUTHORS

Гузаиров Мурат Бакеевич, доктор технических наук, профессор, профессор кафедры управления информационной безопасностью, Уфимский университет науки и технологий, Уфа, Российская Федерация.
e-mail: mbguzairov@gmail.com

Murat B. Guzairov, Doctor of Engineering Sciences, Professor, Professor at the Department of Information Security Management, Ufa University of Science and Technology, Ufa, the Russian Federation.

Валеев Сагит Сабитович, доктор технических наук, профессор, профессор кафедры управления информационной безопасностью, Уфимский университет науки и технологий, Уфа, Российская Федерация.
e-mail: vss2000@mail.ru

Sagit S. Valeev, Doctor of Engineering Sciences, Professor, Professor at the Department of Information Security Management, Ufa University of Science and Technology, Ufa, the Russian Federation.

Кондратьева Наталья Владимировна, кандидат технических наук, доцент, доцент кафедры информатики, Уфимский университет науки и технологий, Уфа, Российская Федерация.
e-mail: knv24@mail.ru

Natalya V. Kondratyeva, Candidate of Engineering Sciences, Docent, Associate Professor at the Department of Computer Science, Ufa University of Science and Technology, Ufa, the Russian Federation.

Статья поступила в редакцию 13.07.2026; одобрена после рецензирования 14.09.2026; принята к публикации 12.10.2026.

The article was submitted 13.07.2026; approved after reviewing 14.09.2026; accepted for publication 12.10.2026.