УДК 004.056:004.8

СОЦИАЛЬНАЯ ИНЖЕНЕРИЯ И ТЕХНОЛОГИИ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА КАК СОВРЕМЕННЫЕ УГРОЗЫ ИНФОРМАЦИОННОЙ БЕЗОПАСНОСТИ.

Несмеянов С. Д., Жумабаев А. Н., Әділ М. С.

Настоящая работа посвящена исследованию роли искусственного интеллекта как ключевого элемента трансформации современных систем информационной безопасности. Актуальность исследования обусловлена экспоненциальным ростом киберугроз, усложнением их архитектуры и снижением эффективности традиционных сигнатурных методов защиты. В статье анализируются основные направления применения интеллектуальных алгоритмов, включая поведенческий анализ пользователей, обнаружение аномалий и автоматизацию реагирования на инциденты. Рассматриваются как защитные, так и атакующие сценарии использования ИИ, включая адаптивные вредоносные программы и алгоритмы обхода систем безопасности. Обоснована необходимость перехода к проактивной модели киберзащиты, основанной на принципах предиктивной аналитики и концепции Zero Trust. Результатом исследования является формирование комплексного подхода к интеграции ИИ в архитектуру информационной безопасности.

Ключевые слова: искусственный интеллект, информационная безопасность, машинное обучение, киберугрозы, аномалия, Zero Trust, автоматизация, кибератаки, поведенческий анализ.

Введение

В условиях стремительной цифровизации экономики, государственного управления и социальной сферы информационные технологии становятся фундаментальной основой функционирования современных социотехнических систем. Масштабное внедрение облачных вычислений, интернета вещей (IoT), распределённых систем хранения данных и цифровых платформ приводит к экспоненциальному росту объемов обрабатываемой информации и усложнению архитектуры информационных инфраструктур.

Одновременно с этим происходит существенное изменение ландшафта киберугроз. Современные атаки характеризуются высокой степенью автоматизации, скрытности и адаптивности, что делает их труднообнаружимыми при использовании традиционных средств защиты. Наблюдается устойчивая тенденция к увеличению количества инцидентов

информационной безопасности, связанных как с внешними злоумышленниками, так и с внутренними нарушителями. Особую опасность представляют атаки нулевого дня (zero-day), а также многоэтапные целевые атаки, направленные на длительное скрытое присутствие в информационной системе и постепенное извлечение конфиденциальной информации.

Традиционные подходы к обеспечению информационной безопасности, основанные на сигнатурном анализе и статических механизмах фильтрации, демонстрируют ограниченную эффективность в условиях динамически изменяющейся среды. Неспособность оперативно адаптироваться к новым типам угроз обусловливает необходимость перехода к более гибким и интеллектуальным методам защиты, способным функционировать в условиях неопределённости и высокой вариативности атакующих сценариев.

В этой связи возрастает роль технологий искусственного интеллекта, позволяющих анализировать большие массивы данных, выявлять скрытые закономерности и прогнозировать угрозы. Их интеграция способствует формированию проактивной модели информационной безопасности.

Вместе с тем развитие искусственного интеллекта имеет двойственный характер. Наряду с расширением возможностей систем защиты, данные технологии активно используются злоумышленниками для повышения эффективности кибератак. Генеративные модели, методы автоматизированного анализа данных и адаптивные алгоритмы позволяют создавать более сложные и труднообнаружимые сценарии атак, что приводит к усложнению процессов обеспечения безопасности.

Таким образом, современная информационная безопасность формируется в условиях постоянного противостояния интеллектуальных систем защиты и атакующих алгоритмов, что требует переосмысления существующих подходов и разработки новых концепций обеспечения устойчивости информационных систем.

1. Искусственный интеллект в обеспечении информационной безопасности

Искусственный интеллект (ИИ) представляет собой совокупность методов, алгоритмов и моделей, направленных на имитацию интеллектуальной деятельности человека. В контексте информационной безопасности ИИ используется для анализа огромных массивов данных, выявления закономерностей и прогнозирования потенциальных угроз. Основная цель применения ИИ - повышение эффективности систем защиты и сокращение времени реакции на инциденты.

Одним из ключевых направлений является обнаружение аномалий в сетевом трафике. Алгоритмы машинного обучения создают модели нормального поведения сетевой инфраструктуры, выявляя отклонения, которые

могут указывать на потенциальные атаки. Данный подход позволяет обнаруживать ранее неизвестные угрозы, включая многоэтапные и целенаправленные атаки на корпоративные системы, которые традиционные сигнатурные методы не способны распознать.

Существенным преимуществом применения ИИ является его способность к самообучению и адаптации. Интеллектуальные алгоритмы постоянно корректируют свои модели поведения на основе новых данных, что снижает количество ложных срабатываний и повышает точность выявления угроз. Это особенно важно для динамичных корпоративных сетей, где новые уязвимости и методы атак появляются ежедневно.

Современные информационные системы генерируют огромные объёмы данных (Big Data): сетевые пакеты, журналы действий пользователей, логи приложений, сообщения об ошибках. Анализ вручную такой информации практически невозможен, в то время как алгоритмы ИИ способны выявлять скрытые закономерности и сложные паттерны поведения, которые могут свидетельствовать о подготовке к атаке или активной компрометации.

ИИ также позволяет интегрировать данные из множества источников: облачных сервисов, IoT-устройств, внешних аналитических платформ и локальных баз данных. Это создаёт единый проактивный слой защиты, повышающий гибкость и адаптивность системы безопасности. Внедрение таких решений позволяет не только реагировать на атаки, но и прогнозировать потенциальные угрозы, формируя превентивный подход к защите корпоративных данных.

2. Поведенческий анализ и обнаружение угроз

Поведенческий анализ пользователей является одним из наиболее перспективных направлений ИИ в кибербезопасности. Он основан на построении цифрового профиля каждого сотрудника, отражающего его типичное поведение в информационной системе: время входа, последовательность действий, используемые ресурсы, интенсивность работы с данными и даже частоту ошибок.

Алгоритмы машинного обучения сравнивают текущие действия пользователей с их историческим поведением. Любое отклонение - например, доступ к данным вне обычного рабочего времени или скачивание большого объёма файлов - фиксируется как потенциальная угроза. Это позволяет выявлять инсайдерские угрозы даже при наличии легитимных учетных данных, которые злоумышленник использует для обхода традиционных систем защиты.

Применение рекуррентных нейронных сетей (RNN) и LSTM-моделей позволяет анализировать временные зависимости действий пользователей. Например, последовательность запросов к разным базам данных может выглядеть безобидно по отдельности, но в контексте времени и порядка

выполнения действий выявляется как потенциальный паттерн подготовки к компрометации.

Совмещение этих методов с UEBA (User and Entity Behavior Analytics) обеспечивает автоматическую реакцию на подозрительные действия: ограничение доступа, уведомление администраторов и запуск внутреннего расследования. В крупных распределённых организациях, где человеческий фактор остаётся главной уязвимостью, такой подход снижает риск внутренних угроз и повышает устойчивость инфраструктуры.

Примеры из практики показывают эффективность ИИ: крупные финансовые компании используют поведенческий анализ для предотвращения мошенничества с клиентскими счетами, а промышленные предприятия - для защиты критической инфраструктуры, предотвращая утечку данных и саботаж.

3. Автоматизация процессов реагирования

Современные системы информационной безопасности постепенно отходят от традиционной схемы реагирования на инциденты, где человек вручную анализировал события, и переходят к концепции SOAR, которая объединяет данные из IDS, SIEM, EDR и других источников, создавая единый интеллектуальный слой реагирования. Искусственный интеллект позволяет не просто классифицировать инциденты по критичности и типу, но и прогнозировать их развитие на основе исторических данных и текущего поведения пользователей и систем, создавая тем самым проактивную модель защиты, способную предотвращать атаки на ранних стадиях. Автоматизация в этом контексте позволяет мгновенно блокировать скомпрометированные аккаунты, ограничивать доступ к критическим ресурсам, уведомлять администраторов о подозрительных действиях и запускать внутренние расследования без необходимости ожидания ручного вмешательства, что критично при распространении сложных атак типа ransomware или многоэтапных APT-кампаний.

ИИ не ограничивается лишь реактивными действиями, он анализирует многомерные массивы данных, выявляет скрытые паттерны поведения пользователей, аномалии сетевого трафика и потенциальные признаки вторжений. На практике это позволяет предсказывать цели атак, вероятность успешного проникновения и масштабы возможного ущерба. Корпорации, внедрившие SOAR с ИИ, отмечают значительное снижение числа успешных инцидентов, сокращение времени расследования с нескольких часов до минут и уменьшение нагрузки на специалистов по безопасности, что позволяет им сосредоточиться на стратегическом планировании, анализе сложных инцидентов и оптимизации процессов защиты. Использование алгоритмов машинного обучения в рамках SOAR также обеспечивает адаптивность системы: она способна сама корректировать правила реагирования, учитывая

новые типы атак и изменяющееся поведение пользователей, создавая динамическую и устойчивую инфраструктуру защиты.

Кроме того, современные подходы включают возможность моделирования сценариев атак и проактивного тестирования уязвимостей инфраструктуры. Системы ИИ могут симулировать действия злоумышленников, выявляя слабые места и указывая на потенциальные точки проникновения, что позволяет организациям заранее корректировать политики безопасности, минимизировать риск утечек и снижать вероятность крупномасштабных инцидентов. Такой подход особенно востребован в финансовой, медицинской и промышленной сферах, где своевременное предотвращение угроз критично для сохранения бизнеса и безопасности критической инфраструктуры.

4. Применение ИИ в кибератаках

Искусственный интеллект используется злоумышленниками не меньше, чем защитниками, что создаёт новую парадигму киберугроз. Генеративные модели позволяют создавать фишинговые письма и сообщения в корпоративных мессенджерах, имитирующие стиль конкретных сотрудников, значительно увеличивая вероятность успешной компрометации корпоративных аккаунтов. Технологии deepfake дают возможность подделывать аудио и видео, имитируя голоса и лица руководителей для социальной инженерии и мошеннических операций. Методы adversarial machine learning позволяют изменять данные таким образом, чтобы они обходили системы обнаружения угроз, модифицируя тексты, изображения, документы и сетевой трафик до того момента, когда их невозможно распознать традиционными системами защиты.

Реальные примеры демонстрируют опасность подобных технологий. В 2023 году злоумышленники использовали deepfake-аудио для имитации голоса CFO крупной компании, чтобы инициировать перевод средств на мошеннический счёт, и только своевременная работа ИИ-системы предотвратила ущерб в миллионы долларов. В другом случае аналитика поведения сотрудников и аномалий трафика позволила заблокировать скрытую кампанию по фишингу, в ходе которой планировалось внедрение вредоносного ПО через корпоративные мессенджеры. Угроза использования ИИ в атаках распространяется не только на коммерческие организации, но и на государственные структуры, критическую инфраструктуру, транспортные сети, энергетику и медицинские учреждения, что требует постоянного совершенствования методов защиты и превентивного мониторинга

Современные исследования показывают, что кибератаки с использованием ИИ становятся всё более автономными, способными самостоятельно адаптироваться к защитным системам и изменять тактики по мере обнаружения препятствий. Это делает традиционные подходы к кибербезопасности недостаточными и подчеркивает необходимость внедрения многоуровневой,

интеллектуальной и динамически адаптирующейся защиты, интегрирующей поведенческий анализ, прогнозирование атак и автоматизацию реагирования

5. Проблемы и ограничения

Несмотря на очевидные преимущества, внедрение искусственного интеллекта в информационную безопасность сталкивается с рядом серьёзных ограничений, которые необходимо учитывать при планировании стратегий защиты. Одной из ключевых проблем остаётся высокая стоимость разработки, внедрения и эксплуатации ИИ-систем. Создание эффективной модели требует значительных ресурсов для сбора, очистки и аннотирования больших объёмов обучающих данных, а также высокопроизводительной вычислительной инфраструктуры для обработки этих данных в реальном времени. Кроме того, качество самой модели напрямую зависит от репрезентативности и полноты данных, и недостаток актуальных данных или их искажение может привести к снижению точности обнаружения угроз, увеличению числа ложных срабатываний и появлению новых уязвимостей.

Дополнительно ИИ-системы остаются подвержены манипуляциям и атакующим воздействиям, включая adversarial-атаки, когда злоумышленники намеренно изменяют входные данные так, чтобы модель приняла их за безопасные события. Такие атаки способны снижать эффективность системы, создавать ложные сигналы тревоги или полностью обходить защиту. Это делает критически важным постоянное тестирование моделей, разработку механизмов защиты от манипуляций, внедрение многослойной архитектуры и непрерывный мониторинг качества данных, а также использование алгоритмов объяснимого ИИ, позволяющих специалистам понимать, на основе каких признаков модель принимает решения.

Не менее значимой проблемой является человеческий фактор и организационные аспекты внедрения ИИ. Специалисты по безопасности должны быть обучены работе с новыми интеллектуальными системами, понимать ограничения алгоритмов и уметь корректировать их работу. Недостаток квалифицированного персонала, отсутствие стандартов интеграции и слабое взаимодействие с существующими инструментами безопасности могут снижать эффективность ИИ и даже создавать дополнительные риски. В целом, успешное внедрение ИИ требует комплексного подхода, включающего сочетание технологий, процессов и кадровых ресурсов, что делает его внедрение сложной, но стратегически важной задачей для организаций любого масштаба.

6. Перспективы развития

Перспективы развития искусственного интеллекта в информационной безопасности связаны с постепенным переходом от традиционных реактивных систем к полностью автономным и проактивным решениям. Такие системы

будут способны не только обнаруживать уже совершённые атаки, но и прогнозировать потенциальные угрозы, анализируя поведение пользователей, сетевой трафик и паттерны взаимодействия между системами. Важнейшим направлением является внедрение концепции Zero Trust с использованием ИИ, где каждый запрос, действие и взаимодействие проверяются автоматически, а решения о доступе принимаются на основе аналитики данных в реальном времени, минимизируя вероятность успешного вторжения и повышая уровень контроля над критическими ресурсами.

Одновременно ведутся исследования по защите самих интеллектуальных систем от манипуляций, adversarial-атак и других видов вмешательства, что позволит обеспечить целостность моделей и их устойчивость к внешнему воздействию. Методы explainable AI становятся всё более востребованными, так как они дают специалистам возможность понимать логику принятия решений системой, оценивать качество работы алгоритмов и корректировать их при необходимости, что повышает доверие к интеллектуальным системам и соответствует нормативным требованиям.

Долгосрочная перспектива включает интеграцию ИИ с облачными платформами, аналитикой больших данных и другими технологиями киберзащиты, создавая высокоадаптивную и устойчивую инфраструктуру, способную предотвращать атаки ещё до того, как они реализуются. Развитие предиктивной аналитики и машинного обучения позволит создавать системы, которые динамически адаптируются к изменениям среды угроз, оценивают потенциальный ущерб и автоматизируют меры защиты, повышая общую безопасность корпоративных и государственных информационных систем.

Заключение

Искусственный интеллект становится фундаментальным инструментом информационной безопасности, обеспечивая повышение точности обнаружения угроз, автоматизацию процессов реагирования и прогнозирование инцидентов. Он позволяет создавать проактивные и адаптивные системы защиты, способные выявлять сложные многоэтапные атаки, анализировать поведение пользователей и предотвращать потенциальные угрозы до того, как они нанесут ущерб. Одновременно ИИ открывает новые возможности для злоумышленников, что требует постоянного совершенствования технологий защиты и комплексного подхода к кибербезопасности.

Эффективная информационная безопасность будущего будет строиться вокруг интеграции ИИ с существующими инструментами, облачными сервисами и аналитикой больших данных, создавая динамическую, самонастраивающуюся инфраструктуру. Такая система способна не только обнаруживать угрозы и реагировать на них в реальном времени, но и прогнозировать направления атак, оценивать риски, предотвращать утечки

информации и снижать вероятность человеческой ошибки. Кроме того, развитие методов explainable AI и прозрачных алгоритмов позволит специалистам анализировать работу систем, корректировать модели и обеспечивать доверие со стороны регуляторов и пользователей.

В конечном итоге искусственный интеллект в информационной безопасности становится стратегическим ресурсом организаций, обеспечивая сохранность данных, устойчивость корпоративных и государственных информационных систем, снижение финансовых и репутационных рисков и повышение общей киберустойчивости. Совмещение автоматизации, прогнозирования и адаптивного поведения ИИ создаёт основу для построения безопасной и надёжной инфраструктуры, способной отвечать на вызовы современного кибермира и развивающихся технологий, обеспечивая стабильность функционирования критически важных процессов и защиту информации на всех уровнях организации.

Список использованной литературы

1. Акилов Е.К., Есмаханова Л.Н. Искусственный интеллект в мире кибербезопасности // Журнал «Механика и технологические исследования». - Таразский региональный университет имени М. Х. Дулати, Казахстан. - 2024.

2. Аширбаева А.С. Роль искусственного интеллекта в кибербезопасности: возможности и риски // Modern Scientific Technology. -Нархоз Университет, Алматы, Казахстан. - 2024.

3. Балтабек А., Погорелов В., Разаке А., Кальпеева Ж. Искусственный интеллект в кибербезопасности: повышение эффективности обнаружения угроз и реагирования // Computing & Engineering. - Университет Сатпаева, Казахстан. - 2024.

4. Мукашева Г.Е., Бейсебаева Ж.Е. Защита информации с использованием искусственного интеллекта // Data Science. - Университет Алихана Бокейхана, Казахстан. - 2025.

5. Бийкенов Н.А. Некоторые проблемы кибербезопасности в Республике Казахстан // Вестник Института законодательства и правовой информации Республики Казахстан. - 2014.

6. Мустафина Т.В., Төребаев О.А. Возможности и риски искусственного интеллекта в трансформации современного Казахстана // Journal of Philosophy, Culture and Political Science. - Евразийский национальный университет, Казахстан. - 2025.

ӘЛЕУМЕТТІК ИНЖЕНЕРИЯ ЖӘНЕ ЖАСАНДЫ ИНТЕЛЛЕКТ ТЕХНОЛОГИЯЛАРЫ АҚПАРАТТЫҚ ҚАУІПСІЗДІКТІҢ ЗАМАНАУИ ҚАУІПТЕРІ РЕТІНДЕ

Несмеянов С. Д., Жумабаев А. Н., Әділ М. С.

Бұл жұмыс ақпараттық қауіпсіздіктің заманауи жүйелерін трансформациялаудың негізгі элементі ретінде жасанды интеллекттің рөлін зерттеуге арналған. Зерттеудің өзектілігі киберқатерлердің экспоненциалды өсуімен, олардың архитектурасының күрделенуімен және дәстүрлі сигнатуралық қорғау әдістерінің тиімділігінің төмендеуімен анықталады. Мақалада интеллектуалды алгоритмдерді қолданудың негізгі бағыттары, оның ішінде пайдаланушылардың мінез-құлқын талдау, аномалияларды анықтау және оқиғаларға ден қоюды автоматтандыру қарастырылады. ЖИ қолданудың қорғаныштық және шабуыл жасаушы сценарийлері, соның ішінде бейімделгіш зиянды бағдарламалар мен қауіпсіздік жүйелерін айналып өту алгоритмдері талданады. Болжамды аналитика қағидаттарына және Zero Trust тұжырымдамасына негізделген проактивті киберқорғау үлгісіне көшудің қажеттілігі негізделген. Зерттеу нәтижесі – ЖИ-ді ақпараттық қауіпсіздік архитектурасына біріктірудің кешенді тәсілін қалыптастыру.

Кілт сөздері: жасанды интеллект, ақпараттық қауіпсіздік, машиналық оқыту, киберқатерлер, аномалия, Zero Trust, автоматтандыру, кибершабуылдар, мінез-құлықты талдау.

SOCIAL ENGINEERING AND ARTIFICIAL INTELLIGENCE TECHNOLOGIES AS MODERN THREATS TO INFORMATION SECURITY

Nesmeyanov S. D., Zhumabayev A. N., Adil M. S.

This paper is devoted to the study of the role of artificial intelligence as a key element in the transformation of modern information security systems. The relevance of the research is driven by the exponential growth of cyber threats, the increasing complexity of their architecture, and the declining effectiveness of traditional signature-based protection methods. The article analyzes the main areas of application of intelligent algorithms, including user behavior analysis, anomaly detection, and incident response automation. Both defensive and offensive use cases of AI are examined, including adaptive malware and algorithms designed to bypass security systems. The necessity of transitioning to a proactive cyber defense model based on predictive analytics and the Zero Trust concept is substantiated. The result of the research is the formation of a comprehensive approach to integrating AI into information security architecture.

Keywords: artificial intelligence, information security, machine learning, cyber threats, anomaly, Zero Trust, automation, cyber attacks, behavioral analysis.

REFERENCES

1. Akilov, Ye.K., Yesmakhanova, L.N. (2024). Artificial Intelligence in the World of Cybersecurity. ***Journal of Mechanics and Technological Research***. M. Kh. Dulaty Taraz Regional University, Kazakhstan.

2. Ashirbayeva, A.S. (2024). The Role of Artificial Intelligence in Cybersecurity: Opportunities and Risks. ***Modern Scientific Technology***. Narxoz University, Almaty, Kazakhstan.

3. Baltabek, A., Pogorelov, V., Razake, A., Kalpeyeva, Zh. (2024). Artificial Intelligence in Cybersecurity: Enhancing Threat Detection and Response Efficiency. ***Computing & Engineering***. Satbayev University, Kazakhstan.

4. Mukasheva, G.Ye., Beisebayeva, Zh.Ye. (2025). Information Protection Using Artificial Intelligence. ***Data Science***. Alikhan Bokeikhan University, Kazakhstan.

5. Biykenov, N.A. (2014). Some Problems of Cybersecurity in the Republic of Kazakhstan. ***Bulletin of the Institute of Legislation and Legal Information of the Republic of Kazakhstan***.

6. Mustafina, T.V., Torebayev, O.A. (2025). Opportunities and Risks of Artificial Intelligence in the Transformation of Modern Kazakhstan. ***Journal of Philosophy, Culture and Political Science***. Eurasian National University, Kazakhstan.