**УДК 004.056 + 004.8**

**Гордеев Данил Александрович,**

*Аспирант направления подготовки «Математическое и программное обеспечение вычислительных машин, комплексов и компьютерных сетей»,*

*dan.gor220@yandex.ru*

*Московский финансово-юридический университет МФЮА,*

*г. Москва*

*научный руководитель:* ***Амелькин Сергей Анатольевич,*** *к. т.н., доцент,*

*старший научный сотрудник,*

*amelkin@ist.education*

*руководитель Исследовательского центра системного анализа Института программных систем имени А. К. Айламазяна РАН,*

*г. Переславль-Залесский*

# РОЛЬ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА В ИНФОРМАЦИОННОЙ БЕЗОПАСНОСТИ

***Аннотация:***

*Раскрыта актуальность и роль применения искусственного интеллекта в обеспечении информационной безопасности. Проведен анализ научных исследований в данной области. Описана эволюция применения искусственного интеллекта в обеспечении информационной безопасности, выявлены основные методы и подходы. Проанализированы существующие примеры успешного использования искусственного интеллекта в области*

*информационной безопасности. Рассмотрены текущие проблемы применения технологии искусственного интеллекта и пути их решения. Определены перспективные тенденции развития искусственного интеллекта в области информационной безопасности.*

***Ключевые слова:*** *искусственный интеллект, информационная безопасность, киберугроза, кибербезопасность, глубокое обучение, машинное обучение, нейронная сеть.*

Непрерывное развитие информационных технологий, сопряженное с увеличивающейся сложностью киберугроз, требует эффективных мер по обеспечению информационной безопасности. Традиционные подходы, основанные на правилах функционирования систем и сигнатурном анализе, демонстрируют недостаточную эффективность перед быстро эволюционирующими угрозами [1]. Внедрение машинного обучения и нейронных сетей представляет собой парадигмальный сдвиг в области информационной безопасности, обеспечивая адаптивную и интеллектуальную защиту. Алгоритмы машинного обучения способны анализировать огромные объемы данных, выявлять нестандартные поведенческие шаблоны и автоматически адаптироваться к новым угрозам. Применение нейронных сетей открывает возможности для создания более сложных и интеллектуальных систем, способных не только распознавать известные угрозы, но и прогнозировать потенциально новые сценарии атак. Это помогает предупреждать инциденты до их возникновения и создавать более устойчивые стратегии обеспечения безопасности [1, 2].

Целью статьи является анализ и оценка роли искусственного интеллекта (далее – ИИ) в обеспечении информационной безопасности, выявление его потенциала и ограничений. Для достижения цели необходимо проанализировать эволюцию развития ИИ в информационной безопасности, рассмотреть основные подходы и методы применения ИИ в предотвращении киберугроз, привести примеры успешного использования

в данной области, а также проанализировать проблемы и этические соображения, связанные с его применением.

На ранних этапах развития, системы ИИ в области информационной безопасности применялись в основном для обнаружения аномалий в сетевом трафике и активности пользователей. С течением времени и с углублением понимания требований безопасности, эволюция систем ИИ привела к интеграции методов машинного обучения и глубокого обучения. Применение этих технологий позволило более точно анализировать и классифицировать аномалии, а также адаптироваться к новым угрозам без необходимости человеческого вмешательства [3].

Ключевым элементом эволюции ИИ в области информационной безопасности является способность к непрерывному обучению. Современные системы ИИ в обеспечении информационной безопасности осуществляют прогнозирование потенциальных угроз, основываясь на анализе больших объемов данных. Применение алгоритмов глубокого обучения позволило эффективно выявлять сложные шаблоны в поведении злоумышленников, что значительно улучшило проактивность в предотвращении атак. Системы автоматической адаптации позволили быстро реагировать на новые угрозы, обновляя и корректируя свои модели в режиме реального времени [2].

Обнаружение угроз – одна из областей, в которой ИИ проявляет свою наибольшую эффективность в обеспечении информационной безопасности. Традиционные методы часто отстают от быстро меняющихся тактик злоумышленников. В отличие от них, алгоритмы машинного обучения способны анализировать обширные наборы данных, выявляя шаблоны, свидетельствующие о потенциальных угрозах и аномалиях. Например, методы опорных векторов (SVM), бутстрэпа (bootstrap) или метод случайного леса (random forest) позволяют анализировать статистику атак, выявлять уязвимости информационной безопасности и строить онтологию

информационной безопасности в информационной системе [4]. Широко применим метод деревьев решений, используемый для анализа и прогнозирования событий, выявления угроз и принятия мер по обеспечению безопасности информационных систем. Однако данные методы написаны людьми и анализируют известные угрозы на основе исторических данных, в то время как методы обучения без «учителя», такие как кластеризация и выявление аномалий, оказываются неоценимыми в обнаружении новых угроз, которые могут не соответствовать предварительно заданным шаблонам злоумышленника. Появление глубокого обучения дополнительно повысило возможности ИИ в обнаружении угроз. Глубокие нейронные сети, особенно сверточные нейронные сети (convolutional neural network, CNN) и рекуррентные нейронные сети (recurrent neural network, RNN), отличаются в извлечении сложных характеристик и временных зависимостей из разнообразных источников данных, улучшая точность и эффективность выявления угроз. Генеративные состязательные нейронные сети (generative adversarial network, GAN) также становятся ценным инструментом для создания реалистичных наборов данных для обучения и имитации различных сценариев киберугроз. Взаимодействие архитектур глубокого обучения улучшает адаптивность и точность систем информационной безопасности [5].

Применение ИИ для оценки уязвимостей представляет проактивный подход к безопасности (comodo proactive security), направленный на выявление и устранение потенциальных угроз и атак до их проявления. Алгоритмы машинного обучения позволяют анализировать код, конфигурацию сети и архитектуру системы, чтобы выявлять уязвимости, которые могли бы ускользнуть от ручной проверки. Техники обработки естественного языка (natural language processing, NLP), взаимодействуя со сканерами уязвимостей, обеспечивают контекстное понимание сведений о безопасности, что облегчает более точную приоритизацию уязвимостей на

основе потенциального воздействия и эксплуатируемости, упрощая процесс устранения угроз [6].

Эффективное и своевременное реагирование на киберугрозы представляет собой ключевой этап в минимизации возможных негативных последствий. Алгоритмы ИИ проводят анализ поступающих потоков данных в реальном времени, оперативно выявляя подозрительную активность и значительно ускоряя процессы реагирования. Автоматизированный отклик на угрозы, поддерживаемый ИИ, обеспечивает более гибкое и согласованное реагирование, а системы реагирования, управляемые ИИ, способны учиться на основе исторических событий, постоянно совершенствуя свои способности в принятии решений. Итеративный процесс обучения систем реагирования повышает адаптивность средств и систем обеспечения информационной безопасности, улучшая автоматизированный отклик на постоянно меняющиеся угрозы [7].

Незаменимой частью обеспечения информационной безопасности является аутентификация и ИИ играет важную роль в совершенствовании её механизмов. Традиционные методы аутентификации на основе пароля, подвержены различным атакам, включая фишинг и попытки перебора. Применение адаптивной аутентификации с внедрением машинного обучения оценивает поведение пользователя, позволяя системе динамически настраивать требования к аутентификации на основе факторов риска. Совокупность аутентификации и ИИ создает более гибкие и эффективные механизмы защиты, способные адаптироваться к постоянно меняющимся угрозам и предоставлять более высокий уровень безопасности для конечных пользователей [6].

Объяснимый искусственный интеллект (explainable artificial intelligence, XAI) становится ключевым направлением в области информационной безопасности. Основной задачей XAI является не только предоставление точных результатов, но и объяснений, понятных для

человека. Эффективный XAI не только повышает доверие к системам ИИ, но также обеспечивает возможность более широкого внедрения этих технологий в общество [8].

Рассмотрим существующие системы, применяющие технологию ИИ в обеспечении информационной безопасности [9-11].

Примером успешной интеграции ИИ служит выдающаяся когнитивная система IBM Watson, которая производит анализ обширных массивов данных, включая блоги по безопасности, научные статьи и форумы, с целью выявления новых потенциальных угроз и уязвимостей. Система предоставляет оперативные исследования аналитикам безопасности данных, обеспечивая им возможность принятия решений в реальном времени.

Инновационной технологией в сфере информационной безопасности также является система Darktrace, успешно применяющая методы самообучения и машинного обучения для создания уникальной модели, известной как «Enterprise Immune System» или иммунная система предприятия. Система проводит анализ сетевого трафика, выявляя аномалии на основе поведения устройств и пользователей. Одним из значимых преимуществ подхода Darktrace является способность обнаруживать новые угрозы, не соответствующие предварительно установленным шаблонам.

Платформа IBM QRadar, спроектированная для управления информационной безопасностью и мониторинга событий, внедряет синтез данных о безопасности из разнообразных источников, включая события, угрозы и служебные данные. QRadar применяет алгоритмы машинного обучения для выявления аномалий и обнаружения потенциальных угроз.

Symantec Endpoint Protection (SEP), в свою очередь, интегрирует передовые технологии машинного обучения для эффективного выявления и

блокировки вредоносных программ. SEP производит анализ поведения программ и файлов, идентифицируя потенциальные угрозы, включая те, которые представляют собой новые варианты вредоносного программного обеспечения.

Решение Microsoft Defender ATP представляет собой высокотехнологичную систему обеспечения безопасности на уровне конечных точек. Используя искусственный интеллект для анализа поведения пользователей и устройств, а также для выявления аномалий, платформа эффективно предотвращает угрозы безопасности.

Рассмотренные платформы, основанные на технологиях ИИ, представляют собой внушительный набор инструментов для противостояния современным киберугрозам. Использование когнитивного компьютера IBM Watson, системы Darktrace, платформы IBM QRadar, Symantec Endpoint Protection и Microsoft Defender ATP демонстрирует разностороннюю функциональность и способность эффективно выявлять, анализировать и противостоять угрозам информационной безопасности.

Внедрение ИИ в обеспечение информационной безопасности сопряжено с рядом преимуществ, но также несет существенные риски. Одной из ключевых проблем является угроза враждебных атак на модели машинного обучения. Злоумышленники могут пытаться воздействовать на системы ИИ, предоставляя им ложные данные с целью нарушения их функциональности [12]. Для предотвращения подобных инцидентов необходимо внедрение мер, направленных на обеспечение безопасности процессов обучения моделей ИИ, а также регулярную проверку их надежности.

Еще одним важным аспектом является интерпретируемость и объяснимость систем информационной безопасности, управляемых ИИ. С увеличением сложности этих систем становится все труднее понимать логику принимаемых ими решений. Гарантирование прозрачности в работе

моделей ИИ имеет ключевое значение для создания доверия в информационном сообществе и обеспечения ответственности за принимаемые ими решения. В этом смысле обеспечение возможности понимания принципов функционирования ИИ и его применения в сфере информационной безопасности является критически важным [8, 12].

Обработка огромных объемов данных в системах ИИ вызывает вопросы в области конфиденциальности. Поиск баланса между необходимостью полного анализа данных и уважением конфиденциальности пользователей остается актуальной проблемой. Применение таких технологий, как федеративное обучение и гомоморфное шифрование, играет ключевую роль в решения вопроса обеспечения конфиденциальности. Эти подходы позволяют проводить анализ данных, не раскрывая саму информацию, что способствует достижению сбалансированного сочетания между обеспечением безопасности данных и защитой частной информации пользователей [12].

Интеграция ИИ с другими технологиями, такими как интернет вещей (IoT) и блокчейн, обещает формирование более устойчивых и безопасных цифровых экосистем. Применение ИИ для выявления аномалий в устройствах IoT и механизмы аутентификации на основе блокчейна – это лишь несколько примеров взаимодействия, которые существенно улучшат информационную безопасность [13]. Внедрение ИИ в разработку информационной безопасности, вероятно, приведет к созданию более сложных платформ анализа угроз.

С ростом роли ИИ в информационной безопасности выдвигаются этические вопросы. Этика использования ИИ подразумевает устранение предвзятости в алгоритмах, обеспечение прозрачности в процессах принятия решений и защиту конфиденциальности пользователей. Сообществу необходимо активно участвовать в разработке и придерживаться этических принципов, чтобы предотвратить

непредвиденные последствия и недопустимое использование технологий ИИ [12].

В заключении можно отметить, что интеграция искусственного интеллекта в обеспечение информационной безопасности представляет значительный шаг в развитии цифровой защиты. Результаты проведенной работы позволяют оценить роль ИИ в обеспечении информационной безопасности. Сочетание машинного обучения, глубокого обучения и других технологий ИИ предоставляет специалистам по информационной безопасности мощные инструменты для борьбы с современными киберугрозами. Несмотря на перспективы, описанные в работе, необходимо осознанно подходить к процессу интеграции ИИ в области информационной безопасности, учитывая атаки злоумышленников, проблемы интерпретации и вопросы конфиденциальности.

С эволюцией искусственного интеллекта его роль в укреплении цифровой устойчивости также будет эволюционировать. Этот непрерывный путь требует постоянного исследования, этического осознания и активного подхода к решению возникающих проблем. Принятие трансформационного потенциала искусственного интеллекта в обеспечении информационной безопасности позволит успешно справляться с постоянно меняющимися угрозами и строить более безопасное цифровое будущее, которое уже сегодня обещает многое, включая продвижение объяснимого искусственного интеллекта, интеграцию с квантовой криптографией и постоянное развитие ИИ-ориентированных платформ информационной безопасности.

## СПИСОК ЛИТЕРАТУРЫ

1. Ламонина, Л. В. К вопросу о применении искусственного интеллекта в обеспечении информационной безопасности / Л. В. Ламонина, О. Б. Смирнова // Электронный научно-методический журнал Омского ГАУ. – 2023. – № 3(34).

2. Брюханов, В. А. Применяемость искусственного интеллекта для решения задач информационной безопасности / В. А. Брюханов, В. В. Грызунов, А. Д. Гарбар // Информационные технологии и системы: управление, экономика, транспорт, право. – 2023. – № 1(45). – С. 35-45.
3. Силкина, О. Ю. Тенденции в развитии искусственного интеллекта / О. Ю. Силкина, Р. С. Зарипова // Информационные технологии в строительных, социальных и экономических системах. – 2020. – № 3(21). – С. 63-65.
4. Зыков, Д. А. Исследование статистических методов для повышения защищенности информационных систем / Д. А. Зыков, В. В. Комашинский // Обработка, передача и защита информации в компьютерных системах 22 : Сборник докладов Второй Международной научной конференции, Санкт-Петербург, 11–15 апреля 2022 года. – Санкт-Петербург: Санкт-Петербургский государственный университет аэрокосмического приборостроения, 2022. – С. 236-241.
5. Клюев, С. Г. Проблемы обучения глубоких нейронных сетей для обнаружения угроз нарушения безопасности в сетях с динамической топологией / С. Г. Клюев, Е. Е. Трунов // Моделирование, оптимизация и информационные технологии. – 2021. – Т. 9, № 1(32).
6. Скрыпников, А. В. Использование методов машинного обучения при решении задач информационной безопасности / А. В. Скрыпников, В. В. Денисенко, И. А. Саранов // Вестник Воронежского института ФСИН России. – 2020. – № 4. – С. 69-73.
7. Решение задач информационной безопасности с использованием искусственного интеллекта / А. В. Скрыпников, В. В. Денисенко, Е. Г. Хитров [и др.] // Современные наукоемкие технологии. – 2021. – № 6-2. – С. 277-281.
8. Аверкин, А. Н. Объяснимый искусственный интеллект как часть искусственного интеллекта третьего поколения / А. Н. Аверкин // Речевые технологии. – 2023. – № 1. – С. 4-10.

9. Бочаров, М. И. Системы машинного обучения в безопасности / М. И. Бочаров, В. А. Лыков // Открытая наука 2021 : Сборник материалов научной конференции с международным участием, Москва, 22 апреля 2021 года. – Москва: Издательство «Aegitas», 2021. – С. 101-106.
10. Калининский, Д. С. Сравнительный анализ антивирусных решений для обеспечения безопасности / Д. С. Калининский // Международный журнал гуманитарных и естественных наук. – 2023. – № 7-1(82). – С. 203-207.
11. Власова, А. В. Обзор программных решений SIEM-технологий. Краткая характеристика программных продуктов SIEM / А. В. Власова, В. А. Дударев, Т. И. Новикова // Современные научные исследования: теория, методология, практика : Сборник научных статей по материалам IX Международной научно-практической конференции, Уфа, 06 декабря 2022 года. Том Часть 4. – Уфа: Общество с ограниченной ответственностью "Научно-издательский центр "Вестник науки", 2022. – С. 144-149.
12. Сычев, Д. И. Искусственный интеллект и кибербезопасность: будущие тенденции и вызовы / Д. И. Сычев // Международный журнал информационных технологий и энергоэффективности. – 2023. – Т. 8, № 5-2(31). – С. 9-14.
13. Белова, Е. И. Перспективы применения искусственного интеллекта в сфере информационной безопасности / Е. И. Белова // Развитие современной науки и технологий в условиях трансформационных процессов : сборник материалов XII Международной научно-практической конференции, Москва, 02 июня 2023 года. – Санкт-Петербург: Печатный цех, 2023. – С. 16-18.

**Gordeev Danil Alexandrovich**

**Amelkin Sergey Anatolyevich**

# THE ROLE OF ARTIFICIAL INTELLIGENCE IN INFORMATION SECURITY

***Annotation:***

*The relevance and role of applying artificial intelligence in ensuring information security have been disclosed. An analysis of scientific research in this field has been conducted. The evolution of the application of artificial intelligence in information security has been described, highlighting key methods and approaches. Existing examples of successful artificial intelligence use in information security have been analyzed. Current issues in the application of artificial intelligence technology and ways to address them have been examined. Promising trends in the development of artificial intelligence in the field of information security have been identified.*

***Keywords:*** *Artificial intelligence, information security, cyber threat, cybersecurity, deep learning, machine learning, neural network.*