AMERICAN JOURNAL OF EDUCATION AND LEARNING
ISSN: 2996-5128 (online) | ResearchBib (IF) = 10.91 IMPACT FACTOR
Volume-4| Issue-7| 2026 Published: |30-07-2026|

ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ В КИБЕРБЕЗОПАСНОСТИ И ИХ ПРИМЕНЕНИЯ.

https://doi.org/10.5281/zenodo.21251958

Исрайилжанова Гулбахор Саминжановна
Академический лицей № 2
при Ферганском государственном техническом университете

Аннотация

В статье рассматривается роль искусственного интеллекта в современной кибербезопасности. Актуальность темы связана с ростом количества кибератак, усложнением методов злоумышленников и увеличением объёмов цифровых данных, которые необходимо анализировать в режиме реального времени. Искусственный интеллект используется как инструмент защиты информации: он помогает выявлять фишинговые письма, анализировать вредоносное программное обеспечение, обнаруживать аномалии в поведении пользователей, автоматизировать работу центров мониторинга безопасности и находить уязвимости в программном коде. Вместе с тем ИИ становится и новым инструментом злоумышленников: с его помощью создаются убедительные фишинговые сообщения, видео и голосовые подделки, автоматизируются атаки и повышается их масштаб. Цель статьи — проанализировать основные направления применения искусственного интеллекта в кибербезопасности, определить его преимущества, риски и перспективы развития.. Сделан вывод о том, что искусственный интеллект является одновременно мощным средством защиты и потенциальным инструментом атакующих, поэтому его применение требует комплексного подхода, технического контроля, обучения сотрудников и внедрения.

Ключевые слова

искусственный интеллект, кибербезопасность, фишинг, машинное обучение, SOC, вредоносное ПО.

Введение

В условиях цифровизации общества кибербезопасность становится одной из ключевых задач для организаций, государственных структур и отдельных пользователей. Современные информационные системы ежедневно обрабатывают огромные объёмы данных, а количество цифровых сервисов, онлайн-платформ, облачных решений и удалённых рабочих мест постоянно растёт. Вместе с этим увеличивается и число киберугроз: фишинговые атаки, вредоносные программы, утечки персональных данных, атаки на корпоративную инфраструктуру, компрометация учётных записей и социальная инженерия [7].

Традиционные методы защиты, основанные только на статических правилах и сигнатурах, уже не всегда способны своевременно обнаруживать

AMERICAN JOURNAL OF EDUCATION AND LEARNING
ISSN: 2996-5128 (online) | ResearchBib (IF) = 10.91 IMPACT FACTOR
Volume-4| Issue-7| 2026 Published: |30-07-2026|

новые угрозы. Злоумышленники быстро меняют тактику, используют автоматизацию, скрытые каналы связи, новые уязвимости и методы обхода средств защиты. Поэтому современная кибербезопасность требует технологий, которые могут анализировать события в реальном времени, выявлять нестандартное поведение и адаптироваться к новым типам атак.

Одной из таких технологий является искусственный интеллект. В широком смысле искусственный интеллект — это способность компьютерных систем выполнять задачи, которые обычно требуют человеческого интеллекта: анализировать информацию, распознавать закономерности, обучаться на данных, классифицировать объекты и предлагать решения. В кибербезопасности искусственный интеллект применяется для мониторинга сети, обнаружения атак, анализа вредоносного программного обеспечения, автоматизации SOC, выявления аномалий и исследования поведения пользователей [1].

Особое значение имеют машинное обучение и глубокое обучение. Машинное обучение позволяет алгоритмам находить закономерности в данных без прямого программирования каждого действия. Глубокое обучение, основанное на нейронных сетях, помогает распознавать сложные паттерны, например подозрительные действия пользователей, вредоносный код или признаки фишинговых сообщений. Генеративный искусственный интеллект способен создавать тексты, изображения, программный код и голос, что открывает новые возможности не только для защиты, но и для атак.

Актуальность темы заключается в двойственной природе искусственного интеллекта. С одной стороны, он повышает эффективность киберзащиты, ускоряет анализ событий и снижает нагрузку на специалистов. С другой стороны, те же технологии могут использоваться злоумышленниками для создания более убедительных и масштабных атак. Например, генеративные модели могут писать грамотные фишинговые письма, имитировать стиль общения руководителя, создавать deepfake-видео и генерировать вредоносные скрипты.

Для достижения цели поставлены следующие задачи:

1. Раскрыть сущность искусственного интеллекта в контексте кибербезопасности.
2. Рассмотреть основные сферы применения ИИ для защиты информации.
3. Проанализировать угрозы, возникающие при использовании ИИ злоумышленниками.
4. Описать методы обнаружения фишинга, deepfake-атак, аномалий и вредоносного ПО.
5. Определить перспективы развития искусственного интеллекта в кибербезопасности.

Методы

Метод анализа позволил рассмотреть искусственный интеллект как комплексную технологию, которая включает машинное обучение, глубокое

AMERICAN JOURNAL OF EDUCATION AND LEARNING
ISSN: 2996-5128 (online) | ResearchBib (IF) = 10.91 IMPACT FACTOR
Volume-4| Issue-7| 2026 Published: |30-07-2026|

обучение, обработку естественного языка, генеративные модели и поведенческую аналитику. Метод сравнения применялся для сопоставления положительных и отрицательных сторон использования ИИ: как инструмента защиты и как средства усиления кибератак. Метод систематизации позволил распределить материал по основным направлениям: обнаружение фишинга, анализ deepfake, выявление аномалий, UEBA, поиск уязвимостей, анализ вредоносного программного обеспечения, инструменты ИИ и риски применения.

Исследование имеет обзорно-аналитический характер. Статья не содержит собственного эксперимента, но обобщает и раскрывает ключевые положения в форме научного текста. Такой подход позволяет показать логику развития темы и представить материал не в виде отдельных слайдов, а как цельное исследование.

Результаты

Роль искусственного интеллекта в современной кибербезопасности . Искусственный интеллект становится необходимым элементом современной системы защиты информации. Это связано с тем, что объёмы данных, которые должны анализироваться специалистами по безопасности, значительно превышают возможности ручной обработки. Системы SIEM, EDR, NDR, почтовые шлюзы, облачные сервисы и корпоративные приложения ежедневно создают большое количество событий и логов. Без автоматизированного анализа специалистам сложно быстро определить, какие события являются обычными, а какие могут указывать на атаку. ИИ позволяет ускорить обработку таких данных. Он способен анализировать сетевой трафик, действия пользователей, активность устройств, электронную почту, доменные имена, URL-адреса, файлы и процессы. Благодаря этому система может быстрее выявлять признаки фишинга, заражения вредоносным ПО, компрометации аккаунтов или подозрительных подключений. Основными направлениями применения ИИ в кибербезопасности являются мониторинг сети в реальном времени, обнаружение атак и вторжений, анализ вредоносного программного обеспечения, автоматизация работы SOC, анализ поведения пользователей и устройств, поиск уязвимостей в программном коде, а также обнаружение фишинговых писем и deepfake-контента [1,2].

Главное преимущество ИИ заключается в скорости. Если человек анализирует инцидент вручную, ему требуется время на сбор данных, проверку логов, сопоставление событий и подготовку решения. Искусственный интеллект может выполнять эти действия значительно быстрее, особенно если он интегрирован с системами мониторинга и реагирования. Поэтому ИИ становится важным инструментом для снижения времени обнаружения и реакции на угрозы.

Искусственный интеллект как инструмент злоумышленников.

Несмотря на пользу для защиты, искусственный интеллект используется и в атакующих целях. Злоумышленники применяют ИИ для автоматизации

AMERICAN JOURNAL OF EDUCATION AND LEARNING
ISSN: 2996-5128 (online) | ResearchBib (IF) = 10.91 IMPACT FACTOR
Volume-4| Issue-7| 2026 Published: |30-07-2026|

подготовки атак, создания фишинговых писем, генерации вредоносных скриптов, поиска уязвимых систем и проведения социальной инженерии [2; 7].

Одним из наиболее опасных направлений является AI-generated phishing — фишинг, созданный с помощью искусственного интеллекта. Раньше фишинговые письма часто можно было распознать по ошибкам, странным формулировкам и шаблонному стилю. Современные языковые модели позволяют создавать грамматически правильные, убедительные и персонализированные сообщения. Такие письма могут быть написаны на нужном языке, учитывать должность жертвы, стиль корпоративной переписки и информацию из открытых источников. ИИ также помогает злоумышленникам проводить OSINT-анализ. Они могут собирать данные из социальных сетей, корпоративных сайтов, открытых баз и утечек, чтобы сделать атаку более персонализированной. В результате фишинговое письмо выглядит не как массовая рассылка, а как индивидуальное сообщение от знакомой организации, руководителя или коллеги.

Кроме того, ИИ используется для автоматизации полного цикла атаки. Он может помогать в разведке, сканировании, поиске уязвимостей, генерации эксплойтов, закреплении в системе и выводе данных. Это снижает порог входа для киберпреступников: даже злоумышленники с ограниченными техническими знаниями могут использовать готовые инструменты для создания более сложных атак.

Обнаружение фишинга с помощью искусственного интеллекта. Одним из важных защитных применений ИИ является обнаружение фишинговых сообщений. Современная антифишинговая система анализирует не только текст письма, но и множество дополнительных признаков: ссылки, домены, метаданные, репутацию отправителя, структуру URL и контекст переписки.

ИИ может выполнять NLP-анализ содержания письма. Он определяет признаки социальной инженерии, срочности, давления на пользователя, подозрительные формулировки, просьбы перейти по ссылке или передать конфиденциальные данные. Например, фразы вроде «срочно обновите пароль», «действие до конца дня», «ваша учётная запись будет заблокирована» могут быть признаками манипуляции.

Следующим этапом является анализ ссылок. Система проверяет структуру URL, наличие сокращённых ссылок, использование похожих доменов, SSL-сертификаты, редиректы и репутацию сайта. Также важна проверка домена: WHOIS-информация, возраст домена, DNS-записи, геолокация сервера, история домена и совпадение с чёрными списками.

Общая схема обнаружения фишинга может быть представлена следующим образом: письмо поступает в систему, затем проходит AI-анализ, проверку домена, анализ репутации и только после этого принимается решение — пропустить письмо, отправить его в карантин или заблокировать. Такой многоуровневый подход позволяет выявлять даже новые фишинговые

AMERICAN JOURNAL OF EDUCATION AND LEARNING
ISSN: 2996-5128 (online) | ResearchBib (IF) = 10.91 IMPACT FACTOR
Volume-4| Issue-7| 2026 Published: |30-07-2026|

кампании, которые ещё не попали в классические базы сигнатур. Особенно опасной является подмена голоса. Злоумышленник может имитировать голос руководителя компании и убедить сотрудника срочно перевести деньги, передать доступы или выполнить действие, нарушающее политику безопасности. Подделка видео также может использоваться в BEC-атаках, где создаётся иллюзия общения с реальным директором или партнёром. Система ищет артефакты синтеза речи: неестественные паузы, искажения частот, необычную интонацию и признаки искусственной генерации.

Биометрический анализ: проверка микродвижений лица, моргания, мимики и естественных морщин. Третий метод — анализ изображения: поиск несоответствий в освещении, тенях, текстуре кожи, глазах и синхронности движений губ. Однако технических средств недостаточно. Важным правилом остаётся организационная проверка. Если сотрудник получает подозрительный звонок или видеосообщение с просьбой выполнить финансовую операцию или передать данные, необходимо связаться с человеком по заранее известному номеру или использовать дополнительный канал подтверждения.

Процесс обнаружения аномалий состоит из нескольких этапов. Сначала собираются данные о нормальном поведении: время входа, используемые устройства, частота запросов, объём скачиваний, типичные приложения и местоположение. Затем модель обучается на этих данных и формирует поведенческий профиль. После этого начинается мониторинг в реальном времени. Если действия пользователя выходят за допустимый порог, система создаёт предупреждение.

Преимущество такого подхода заключается в том, что он помогает выявлять не только внешние атаки, но и внутренние угрозы. Например, сотрудник с легитимным доступом может начать выполнять подозрительные действия. Традиционные системы могут не заметить нарушение, потому что пользователь формально авторизован. **Автоматизированный поиск уязвимостей и эксплойтов.**

ИИ применяется и в анализе программного кода. Современные AI-сканеры помогают находить уязвимости в приложениях, библиотеках, зависимостях и контейнерах. Они могут выявлять SQL Injection, XSS, Command Injection, SSRF, Path Traversal, Buffer Overflow, жёстко прописанные пароли, секретные ключи и другие ошибки безопасности. Эти категории связаны с наиболее распространёнными рисками веб-приложений, которые систематизируются в материалах OWASP Top 10 [3].

Преимущество ИИ-сканеров заключается в том, что они не только находят проблему, но и помогают приоритизировать риски. Не каждая уязвимость имеет одинаковую опасность. Некоторые ошибки теоретически возможны, но почти не эксплуатируются, а другие уже используются злоумышленниками. ИИ может учитывать CVSS-оценку, наличие публичного эксплойта, контекст системы и активность атак.

AMERICAN JOURNAL OF EDUCATION AND LEARNING
ISSN: 2996-5128 (online) | ResearchBib (IF) = 10.91 IMPACT FACTOR
Volume-4| Issue-7| 2026 Published: |30-07-2026|

Также важным направлением является интеграция проверки безопасности в CI/CD. Это означает, что анализ кода проводится при каждом коммите, pull request или деплое. Такой подход позволяет находить уязвимости раньше, до выхода продукта в эксплуатацию. В результате безопасность становится частью процесса разработки, а не отдельным этапом после завершения проекта. ИИ также может использоваться для моделирования сценариев атак. Он помогает анализировать цепочку kill chain, оценивать, какие уязвимости могут быть использованы вместе, и определять реальный риск для организации.

Анализ вредоносного программного обеспечения. ИИ играет важную роль в анализе вредоносного ПО. Вредоносные программы постоянно изменяются, используют обфускацию, шифрование, скрытые каналы связи и методы обхода антивирусов. Поэтому традиционного сигнатурного анализа часто недостаточно.

ИИ может применяться на нескольких уровнях.

Первый уровень — статический анализ, то есть изучение кода без запуска файла. Система проверяет строки, макросы, PowerShell-команды, подозрительные API-вызовы, зашифрованные команды и сходство с известными семействами вредоносного ПО.

Второй уровень — динамический анализ. Файл запускается в изолированной среде, например в песочнице, где отслеживаются его действия: создание или изменение файлов, сетевые соединения, попытки закрепления в системе, отключение защитных средств, изменение реестра и запуск процессов.

Третий уровень — классификация угроз. ИИ может определять, к какому типу относится вредоносная программа: ransomware, trojan, infostealer, spyware, loader или dropper. Это помогает специалистам быстрее понять характер атаки и выбрать правильную стратегию реагирования.

Четвёртый уровень — sandbox-анализ с использованием ИИ. Система не просто фиксирует действия файла, но и оценивает их поведенческую опасность. Например, если программа пытается подключиться к C2-серверу, шифрует файлы, отключает защиту и создаёт механизмы закрепления, вероятность вредоносности становится высокой.

Обсуждение

Результаты анализа показывают, что искусственный интеллект радикально меняет подходы к кибербезопасности. Он позволяет перейти от реактивной модели защиты к более проактивной. Раньше многие системы обнаруживали угрозы только после появления известных сигнатур или явных признаков атаки. ИИ даёт возможность выявлять подозрительное поведение раньше, ещё на стадии подготовки или начала атаки.

Однако использование ИИ не решает все проблемы автоматически. Одним из ключевых рисков являются ошибочные решения. Ложные срабатывания могут приводить к блокировке легитимной активности, а ложные отрицания — к пропуску реальных угроз. Поэтому результаты

AMERICAN JOURNAL OF EDUCATION AND LEARNING
ISSN: 2996-5128 (online) | ResearchBib (IF) = 10.91 IMPACT FACTOR
Volume-4| Issue-7| 2026 Published: |30-07-2026|

работы ИИ должны проверяться специалистами, особенно в критически важных случаях.

Второй риск связан с галлюцинациями моделей. Генеративные модели могут создавать убедительные, но неправильные объяснения, рекомендации или сценарии реагирования. Если аналитик полностью доверяет такой системе без проверки, это может привести к неправильным действиям во время инцидента.

Третий риск — data poisoning, то есть отравление обучающих данных. Злоумышленники могут попытаться повлиять на данные, на которых обучается модель, чтобы она перестала замечать определённые типы атак или начала неправильно классифицировать события.

Четвёртый риск — prompt injection. Если в систему встроена языковая модель, злоумышленник может попытаться манипулировать ею через входные данные, заставляя игнорировать правила, раскрывать конфиденциальную информацию или выполнять нежелательные действия.

Поэтому ИИ в кибербезопасности должен использоваться не изолированно, а в составе комплексной системы защиты. Необходимы контроль качества данных, тестирование моделей, аудит решений, защита от манипуляций, разграничение доступа и постоянное участие специалистов. Такие подходы соответствуют общей логике управления рисками ИИ, где важными считаются надёжность, безопасность, прозрачность и ответственность при использовании интеллектуальных систем [1].

Перспективы развития ИИ в кибербезопасности связаны с появлением автономных SOC, AI-аналитиков и самообучающихся систем. Автономные SOC смогут самостоятельно обнаруживать, расследовать и нейтрализовать часть угроз. AI-аналитики будут помогать специалистам писать отчёты, проводить расследования, строить цепочку атаки и предлагать рекомендации. Самообучающиеся системы смогут адаптироваться к новым угрозам в реальном времени.

Однако даже при развитии автономных решений человек должен оставаться важным элементом защиты. Искусственный интеллект хорошо справляется с анализом больших объёмов данных, но стратегические решения, оценка контекста, этические вопросы и ответственность должны оставаться за специалистами.

Практические рекомендации

На основе рассмотренного материала можно выделить несколько рекомендаций для организаций.

- необходимо обучать сотрудников распознавать AI-угрозы. В программы повышения осведомлённости следует включать темы фишинга, deepfake, социальной инженерии и правил проверки подозрительных сообщений;
- важно внедрять процедуры проверки deepfake-контента. Подозрительные голосовые и видеозапросы, особенно связанные с

финансами, доступами или конфиденциальными данными, должны подтверждаться через дополнительный канал связи;

- организациям следует использовать AI-защиту: антифишинговые фильтры, системы поведенческой аналитики, инструменты анализа вредоносного ПО, AI-сканеры кода и решения для автоматизации SOC;

- необходимо применять многофакторную аутентификацию, особенно для привилегированных учётных записей. Это снижает риск компрометации паролей и несанкционированного доступа;

- важным принципом является Zero Trust — «никогда не доверяй, всегда проверяй». Этот подход предполагает проверку каждого пользователя, устройства и запроса независимо от того, находится ли он внутри корпоративной сети или за её пределами.

Заключение

Искусственный интеллект становится одним из важнейших факторов развития современной кибербезопасности. Он помогает анализировать большие объёмы данных, обнаруживать фишинг, выявлять аномалии, исследовать вредоносное программное обеспечение, искать уязвимости и автоматизировать работу специалистов по безопасности.

В то же время ИИ создаёт новые угрозы. Злоумышленники используют его для генерации фишинговых писем, создания deepfake-видео и голосовых подделок, автоматизации атак и повышения их масштаба. Поэтому искусственный интеллект нельзя рассматривать только как средство защиты. Он является технологией двойного назначения, эффективность и безопасность которой зависят от того, кто и как её применяет. Главная задача специалистов по кибербезопасности заключается в том, чтобы использовать возможности ИИ быстрее и эффективнее, чем это делают атакующие. Для этого необходимы современные инструменты защиты, обучение сотрудников, проверка подозрительных сообщений, многофакторная аутентификация, поведенческая аналитика и принцип Zero Trust.

Таким образом, искусственный интеллект в кибербезопасности является одновременно мощным инструментом защиты и новым вызовом для организаций. Его грамотное применение позволит повысить устойчивость информационных систем, ускорить реагирование на инциденты и снизить риски кибератак.

СПИСОК ЛИТЕРАТУРЫ:

1. NIST. Artificial Intelligence Risk Management Framework. National Institute of Standards and Technology, 2023.
2. MITRE ATT&CK. Knowledge Base of Adversary Tactics and Techniques.
3. OWASP. Top 10 Web Application Security Risks.
4. Russell S., Norvig P. Artificial Intelligence: A Modern Approach. 4th ed. Pearson, 2021.

5. Stallings W. Computer Security: Principles and Practice. Pearson, 2018.
6. ISO/IEC 27001:2022. Information security, cybersecurity and privacy protection — Information security management systems.
7. ENISA. Threat Landscape Reports. European Union Agency for Cybersecurity.
8. Материалы презентации «Искусственный интеллект в кибербезопасности».