DOI: 10.5281/zenodo.15321663
Link: https://zenodo.org/records/15321663

ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ В КИБЕРБЕЗОПАСНОСТИ: РОЛЬ МАШИННОГО ОБУЧЕНИЯ В ОБЕСПЕЧЕНИИ ЗАЩИТЫ ДАННЫХ

Аблизова Гулзаҳирям Алимовна
Старший преподаватель кафедры «Современные информационные технологи»
Узбекский государственный университет мировых языков
gulzahiraolim@gmail.com
+998903559466

***Аннотация**. В условиях стремительного развития цифровых технологий обеспечение кибербезопасности приобретает первостепенное значение для защиты информации в государственном, корпоративном и личном пространстве. В статье рассматриваются теоретико-прикладные аспекты использования технологий искусственного интеллекта, в частности методов машинного обучения, в целях повышения эффективности киберзащиты. Раскрываются механизмы применения алгоритмов анализа аномалий, классификации угроз и предиктивного моделирования атак. Особое внимание уделено реальным примерам внедрения интеллектуальных систем в практику информационной безопасности, а также анализу потенциальных рисков, связанных с использованием ИИ. Обоснована необходимость стратегической интеграции машинного обучения в архитектуру цифровой безопасности и подчеркнуты перспективы дальнейших исследований в данной области.*

***Ключевые слова:** искусственный интеллект, кибербезопасность, машинное обучение, анализ угроз, предиктивные алгоритмы, информационная безопасность, цифровая защита, интеллектуальные технологии.*

Введение

Современное цифровое общество характеризуется высокой степенью зависимости от информационных технологий, что одновременно повышает уязвимость ключевых инфраструктур к различным видам киберугроз. В условиях стремительного роста объёмов данных, усложнения архитектуры сетевых взаимодействий и расширения числа атакующих векторов классические подходы к обеспечению информационной безопасности становятся недостаточно эффективными. Это обуславливает необходимость внедрения инновационных решений, способных адаптивно и проактивно реагировать на возникающие угрозы.

Одним из наиболее перспективных направлений в сфере кибербезопасности является использование искусственного интеллекта (ИИ), а в особенности — методов машинного обучения. Эти технологии позволяют автоматизировать процессы выявления аномалий, прогнозирования атак, анализа поведения пользователей и классификации угроз в реальном времени. Благодаря способности систем ИИ самообучаться и улучшать точность моделей на основе поступающих данных, возрастает эффективность защиты критических информационных активов.

Актуальность темы определяется тем, что киберугрозы становятся всё более изощрёнными и масштабными, а потенциальный ущерб от атак — экономически и стратегически значимым. В этой связи изучение возможностей искусственного интеллекта в контексте цифровой безопасности представляет научный и практический интерес.

Цель данной статьи — проанализировать роль машинного обучения в обеспечении кибербезопасности, рассмотреть существующие методы, области их применения, а также потенциальные риски и перспективы использования интеллектуальных технологий в защите данных.

Методология

Теоретические основы искусственного интеллекта и кибербезопасности

Искусственный интеллект (ИИ) представляет собой совокупность технологических решений, направленных на моделирование когнитивных функций человека, таких как обучение, рассуждение, самоанализ и принятие решений. В основе современных ИИ-систем лежат алгоритмы машинного обучения, позволяющие автоматизировать процессы обработки больших массивов данных, выявлять закономерности и принимать решения без явного программного вмешательства. Машинное обучение, в свою очередь, подразделяется на несколько направлений — контролируемое, неконтролируемое и обучение с подкреплением, каждое из которых используется в зависимости от специфики задач в области безопасности.

Кибербезопасность как научная и прикладная дисциплина охватывает широкий спектр мер, направленных на защиту информационных ресурсов от внешнего и внутреннего воздействия. Она включает в себя разработку политик информационной безопасности, технические средства защиты, системы мониторинга и реагирования на инциденты. При этом с усложнением цифровой инфраструктуры традиционные средства защиты становятся всё менее эффективными против сложных и адаптивных угроз.

Интеграция ИИ в сферу кибербезопасности основывается на способности интеллектуальных систем анализировать огромное количество данных в режиме реального времени, выявлять аномальные паттерны в поведении пользователей и сетевых процессов, а также оперативно прогнозировать потенциальные угрозы. В отличие от статических систем защиты, ИИ способен адаптироваться к изменяющимся условиям и выявлять ранее неизвестные векторы атак, что существенно повышает уровень устойчивости информационной среды.

С теоретической точки зрения, ключевое преимущество ИИ в кибербезопасности заключается в способности к генерализации и обучению на неструктурированных и частично размеченных данных, что особенно важно в условиях нехватки полных данных о новых видах угроз. Также важным аспектом является возможность масштабирования ИИ-моделей на различные уровни архитектуры информационной безопасности: от защиты конечных устройств до обеспечения устойчивости распределённых облачных систем.

Синергия между технологиями искусственного интеллекта и средствами кибербезопасности открывает новые горизонты в формировании проактивной модели защиты, где ключевую роль играют не только алгоритмы, но и способность системы к самообучению и адаптации в условиях постоянно меняющейся киберсреды.

Методы машинного обучения в кибербезопасности

Применение методов машинного обучения (МЛ) в кибербезопасности представляет собой один из наиболее динамично развивающихся сегментов в сфере интеллектуальных технологий. Эти методы позволяют анализировать большие объёмы разнородных данных, выявлять скрытые угрозы, адаптироваться к новым формам атак и обеспечивать защиту в условиях ограниченной предопределённой информации о потенциальных нарушениях. Особенность МЛ состоит в способности алгоритмов учиться на примерах, находить закономерности и принимать решения без необходимости прямого программного указания всех возможных сценариев.

Обнаружение вредоносного программного обеспечения (malware detection)

Алгоритмы обучения с учителем (supervised learning), такие как решающие деревья, метод опорных векторов и градиентный бустинг, используются для классификации файлов как безопасных или вредоносных на основе метаданных, кода, поведения в среде исполнения и сетевой активности. Эффективность таких моделей зависит от объёма и качества обучающей выборки, при этом они способны выявлять не только известные, но и модифицированные варианты вредоносных программ.

Выявление аномалий в сетевом трафике

Методы неконтролируемого обучения (unsupervised learning), включая кластеризацию (например, алгоритмы k-средних, DBSCAN) и понижение размерности (t-SNE, PCA), позволяют определять отклонения от нормального поведения в сетевой активности, что часто свидетельствует о попытках несанкционированного доступа или проведения атак типа «отказ в обслуживании» (DoS/DDoS). Такие подходы не требуют заранее размеченных данных и эффективно работают в условиях изменяющейся цифровой среды.

Предиктивное моделирование кибератак

Системы с элементами обучения с подкреплением (reinforcement learning) и гибридные нейросетевые модели позволяют выстраивать поведенческие профили пользователей, устройств и приложений, предсказывая вероятность наступления инцидентов информационной безопасности. В таких моделях может использоваться обучение на последовательностях (например, LSTM-сети) для выявления угроз на основе временных зависимостей в логах событий.

Автоматизация процессов реагирования на инциденты

Методы МЛ применяются для разработки интеллектуальных систем принятия решений, которые могут самостоятельно интерпретировать сигналы системы обнаружения атак и выполнять предварительные действия по локализации и нейтрализации угроз. Это особенно важно в условиях, когда время реагирования критично, а количество поступающих событий превышает возможности человеческого анализа.

Фильтрация фишинговых атак и спама

Классификационные модели, обученные на основе текстового и поведенческого анализа, применяются для распознавания фишинга, вредоносных ссылок и поддельных доменов. Использование наивного байесовского классификатора, логистической регрессии и нейросетевых

архитектур обеспечивает высокую точность в фильтрации электронных писем и веб-ресурсов.

РЕЗУЛЬТАТЫ

Обеспечение информационной безопасности в современных условиях требует оперативной обработки масштабных объёмов разнородных данных, поступающих с различных узлов информационных систем. Учитывая возрастающую интенсивность и сложность киберугроз, традиционные методы защиты уже не справляются с поставленными задачами в полной мере. Это актуализирует необходимость внедрения прогрессивных подходов, среди которых особое место занимают интеллектуальные технологии.

Использование инструментов искусственного интеллекта (ИИ) и машинного обучения (ML) позволяет выполнять анализ потоков данных в реальном времени, выявлять аномалии в поведении компонентов системы и оперативно реагировать на потенциальные угрозы. Одним из ключевых достоинств таких решений является возможность адаптации к новым типам атак без необходимости ручного обновления сигнатурных баз.

Алгоритмы машинного обучения демонстрируют высокую эффективность при обнаружении нетипичных поведенческих шаблонов, что особенно важно для предотвращения целевых атак и нарушений, возникающих в обход стандартных систем мониторинга. Кроме того, интеллектуальные технологии способствуют автоматизации процессов обеспечения соответствия нормативным требованиям, тем самым повышая уровень доверия к цифровым продуктам и минимизируя влияние человеческого фактора.

На сегодняшний день ИИ активно применяется в различных направлениях киберзащиты, включая поведенческий анализ, фильтрацию вредоносного контента, управление доступом, обработку событий безопасности и предиктивное моделирование атак. Эти подходы формируют основу интеллектуальной архитектуры информационной безопасности нового поколения, ориентированной на устойчивость и адаптивность в условиях высокоизменчивой цифровой среды[106] [1].

Рис 1. Применение технологии ИИ в безопасной разработке ПО[107].

[106] Ковцур М.М., Кириллов Д.И., Михайлова А.В., Потемкин П.А. Разработка методики внедрения машинного обучения для повышения информационной безопасности web-приложения // Техника средств связи. 2020 №4 (152). С. 74-86.

[107] Ангапов Василий Данилович, Бобров Андрей Владимирович, Тимонин Вадим Андреевич, Вишняков Александр Сергеевич ИСПОЛЬЗОВАНИЕ ТЕХНОЛОГИЙ МАШИННОГО ОБУЧЕНИЯ В ЗАЩИТЕ ИНФОРМАЦИОННЫХ СИСТЕМ // Наука, техника и образование. 2023. №4 (92).

Применение искусственного интеллекта в сфере кибербезопасности открывает новые возможности для предиктивного анализа угроз. С помощью машинного обучения (ML) становится возможным анализировать большие массивы данных и поведенческие модели, характерные для действий злоумышленников, что позволяет заранее выявлять потенциально опасные ситуации. Такие системы способны обнаруживать ранее неизвестные векторы атак, которые могли бы остаться вне зоны внимания при использовании классических средств анализа.

Особую значимость приобретает способность интеллектуальных систем к непрерывному мониторингу информационной среды предприятия. Постоянный анализ входящего трафика, поведения пользователей и изменений в конфигурациях позволяет в реальном времени фиксировать отклонения от нормы и оперативно реагировать на инциденты. Это особенно актуально в условиях, когда характер киберугроз постоянно эволюционирует, и от служб информационной безопасности требуется гибкость и высокая скорость принятия решений.

Кроме того, технологии машинного обучения находят применение в автоматизации процессов тестирования на устойчивость, включая поиск уязвимостей и проведение виртуальных атак для проверки защищённости систем. Такие алгоритмы значительно ускоряют процедуры аудита ИБ, снижая нагрузку на специалистов и повышая точность оценки текущего состояния системы защиты.

Для иллюстрации практического применения ML в задаче идентификации угроз можно использовать алгоритмическую схему, в которой отображается последовательность этапов обработки данных, классификации событий и выработки защитных реакций. Эта схема отражает логику функционирования современных интеллектуальных систем информационной безопасности[108] [2].

Технологии машинного обучения обладают значительным потенциалом в сфере информационной безопасности, открывая принципиально новые возможности для анализа, интерпретации и реагирования на угрозы. Их преимущество заключается не только в способности обрабатывать огромные массивы разнородных данных, но и в умении выявлять закономерности и отклонения, неочевидные для человека. Обучение моделей на базе обширных исторических данных позволяет точно распознавать подозрительную активность и оперативно инициировать меры противодействия, что способствует снижению вероятности реализации атак и минимизации потенциального ущерба для цифровой инфраструктуры[109].

Одним из ключевых отличий алгоритмов машинного обучения от традиционных методов является способность идентифицировать угрозы, ранее не встречавшиеся в практике. В то время как классические системы защиты

[108] Артюшкина Е.С., Андирякова О.О., Тюрина Д.А. Использование методов машинного обучения при анализе сетевого трафика и вредоносного программного обеспечения // Индустриальная экономика. 2023 №4. С. 12-15.
[109] Ожиганова М.И., Куртаметов Э.С. Применение машинного обучения в защите веб-приложений // NBI-technologies. 2020 №2. С. 16-20.

опираются преимущественно на известные сигнатуры и шаблоны, ML-модели выявляют нетипичное поведение, которое может свидетельствовать о новой, ранее незарегистрированной форме атаки. Это обеспечивает гибкость и адаптивность систем киберзащиты, критически важную в условиях постоянно эволюционирующей угрозной среды.

В настоящее время разработано множество прикладных решений на основе технологий машинного обучения, предназначенных для реализации различных аспектов информационной безопасности. К их числу относятся:

– интеллектуальные сканеры уязвимостей, способные обнаруживать потенциальные слабые места в программном обеспечении;

– системы мониторинга, функционирующие в режиме реального времени и осуществляющие оценку поступающих данных по широкому спектру параметров;

– когнитивные аналитические платформы, которые собирают и интерпретируют информацию из открытых источников, включая новостные ленты и форумы, с целью раннего выявления новых угроз и формирования стратегий реагирования.

Эти инструменты становятся основой построения многоуровневой архитектуры киберзащиты, в которой каждый компонент выполняет свою функцию, опираясь на возможности интеллектуального анализа и автоматического принятия решений[110].

На практике внедрение интеллектуальных технологий в архитектуру киберзащиты реализуется посредством использования специализированных решений, основанных на алгоритмах машинного обучения и искусственного интеллекта. Ниже приведены некоторые из наиболее эффективных программных платформ, получивших широкое признание среди профессионального сообщества и успешно применяющихся в корпоративной среде для обеспечения устойчивости информационных систем:

IBM Watson for Cyber Security

Данная интеллектуальная система представляет собой комплексное решение, использующее возможности когнитивных вычислений и машинного обучения для выявления и нейтрализации киберугроз. Программный модуль анализирует многогранные источники информации — журналы событий, отчёты по информационной безопасности, публикации в открытых источниках — с целью оперативного распознавания и классификации новых угроз. Watson способен не только предсказывать потенциальные векторы атак, но и формировать рекомендации по их предотвращению.

FireEye Malware Protection System

Эта платформа ориентирована на проактивную защиту от вредоносного программного обеспечения, включая вирусы, трояны и другие формы вредоносного кода. Алгоритмы машинного обучения, встроенные в систему,

[110] Щербаков А.Е. Исследование применения искусственного интеллекта и машинного обучения в области кибербезопасности: техники обнаружения аномалий и предотвращения угроз // Вестник науки. 2023 №7 (64). С. 151-156.

позволяют распознавать поведенческие паттерны и сетевую активность, характерные для вредоносных объектов. Благодаря возможности анализа в режиме реального времени достигается высокая точность в блокировке аномальной активности ещё до того, как она нанесёт ущерб инфраструктуре.

Darktrace Enterprise Immune System

Инновационная система, реализующая концепцию цифрового иммунитета предприятия. Используя методы машинного обучения и нейросетевые алгоритмы, платформа моделирует «нормальное» поведение пользователей, устройств и процессов в организации. Любое отклонение от установленных поведенческих норм интерпретируется как потенциальная угроза, что позволяет выявлять атаки на ранних стадиях. Система функционирует непрерывно, самообучаясь и адаптируясь к изменениям в инфраструктуре, тем самым минимизируя вероятность ложных срабатываний.

К числу актуальных интеллектуальных решений, реализующих подходы машинного обучения в защите цифровой инфраструктуры, относятся и другие комплексные платформы, получившие широкое распространение на практике.

McAfee Advanced Threat Defense

Это решение ориентировано на противодействие сложным угрозам, включая вредоносные коды нового поколения. Система использует алгоритмы поведенческого анализа и машинного обучения для обнаружения вредоносной активности до того, как она способна повлиять на работоспособность корпоративной инфраструктуры. Благодаря способности к самообучению, система оперативно подстраивается под новые сценарии атак.

Cybereason Endpoint Protection Platform

Многофункциональная система, предназначенная для защиты конечных точек. В основе её работы лежит анализ тысяч сигнальных событий с последующей идентификацией подозрительных паттернов. Механизмы автоматического реагирования позволяют не только блокировать вредоносную активность, но и запускать сценарии восстановления компонентов системы, обеспечивая тем самым непрерывность бизнес-процессов.

Важно подчеркнуть, что приведённые решения представляют собой лишь малую часть доступных интеллектуальных инструментов, основанных на принципах машинного обучения. Все они демонстрируют эффективность в рамках определённых задач, и могут классифицироваться по следующим функциональным направлениям:

– системы предотвращения вторжений (IPS), основанные на ML, — идентифицируют и блокируют угрозы, ещё не зарегистрированные в сигнатурных базах;

– интеллектуальные средства управления доступом (IAM) — отслеживают поведение пользователей, фиксируют отклонения и предотвращают атаки на основе поведенческого анализа;

– системы поведенческого анализа и корреляции событий — выявляют аномалии, способные сигнализировать о попытках компрометации системных ресурсов.

ЗАКЛЮЧЕНИЕ

Преимущества применения машинного обучения в информационной безопасности проявляются в нескольких ключевых аспектах. Во-первых, происходит масштабная автоматизация рутинных процессов — таких как анализ логов, фильтрация событий и предварительная классификация угроз. Это высвобождает ресурсы ИБ-специалистов для решения более сложных задач. Во-вторых, интеллектуальные системы обладают способностью к адаптации — обучаясь на новых входных данных, они своевременно подстраиваются под изменяющийся ландшафт угроз. В-третьих, ML-алгоритмы предоставляют структурированную информацию и рекомендации, способствуя более обоснованному принятию решений в вопросах стратегии защиты.

Тем не менее, использование машинного обучения в системах информационной безопасности сопряжено с рядом вызовов. Прежде всего, для формирования качественных моделей требуется большое количество достоверных и разметанных данных, доступ к которым может быть ограничен. Недостаток обучающего материала или его несбалансированность снижает точность прогнозирования и может привести к ложноположительным или ложноотрицательным срабатываниям.

Кроме того, в настоящее время остаётся нерешённой проблема интерпретируемости решений, принимаемых ИИ-системами. Алгоритмы часто функционируют как «чёрные ящики», генерируя высокоточные, но труднопроверяемые результаты. Это ограничивает возможности специалистов по ИБ в понимании механизмов обнаружения угроз и снижает уровень доверия к автоматизированным рекомендациям. Для преодоления данной проблемы появляются системы, способные транслировать действия искусственного интеллекта в форму, понятную человеку. Одним из таких решений выступает программный продукт AVSOFT ATHENA, обеспечивающий объяснимость принимаемых решений в контексте кибербезопасности.

СПИСОК ИСПОЛЬЗОВАННОЙ ЛИТЕРАТУРЫ

1. Ковцур М.М., Кириллов Д.И., Михайлова А.В., Потемкин П.А. Разработка методики внедрения машинного обучения для повышения информационной безопасности web-приложения // Техника средств связи. 2020 №4 (152). С. 74-86.
2. Ангапов Василий Данилович, Бобров Андрей Владимирович, Тимонин Вадим Андреевич, Вишняков Александр Сергеевич ИСПОЛЬЗОВАНИЕ ТЕХНОЛОГИЙ МАШИННОГО ОБУЧЕНИЯ В ЗАЩИТЕ ИНФОРМАЦИОННЫХ СИСТЕМ // Наука, техника и образование. 2023. №4 (92).
3. Артюшкина Е.С., Андирякова О.О., Тюрина Д.А. Использование методов машинного обучения при анализе сетевого трафика и вредоносного программного обеспечения // Индустриальная экономика. 2023 №4. С. 12-15.
4. Ожиганова М.И., Куртаметов Э.С. Применение машинного обучения в защите веб-приложений // NBI- technologies. 2020 №2. С. 16-20.

5. Щербаков А.Е. Исследование применения искусственного интеллекта и машинного обучения в области кибербезопасности: техники обнаружения аномалий и предотвращения угроз // Вестник науки. 2023 №7 (64). С. 151-156.

6. Худхейр А. Р., Заргарян Е. В., Заргарян Ю. А. Модели машинного обучения и глубокого обучения для электронной информационной безопасности в мобильных сетях // Известия ЮФУ. Технические науки. — 2022. — № 3 (227). — С. 211–222.

7. Козин И. С., Рощин А. А. Метод обеспечения безопасности информации при её обработке в информационной системе на основе машинного обучения // Техника средств связи. — 2019. — № 4 (148). — С. 70–82.

8. Власенко А. В., Дзьобан П. И., Жук Р. В. Обзор инструментов машинного обучения и их применения в области кибербезопасности // Прикаспийский журнал: управление и высокие технологии. — 2020. — № 1 (49). — С. 144–155.

9. Гетьман А. И., Горюнов М. Н., Мацкевич А. Г., Рыболовлев Д. А. Сравнение системы обнаружения вторжений на основе машинного обучения с сигнатурными средствами защиты информации // Труды ИСП РАН. — 2022. — № 5. — С. 111–126.

10. Базылев Н. С., Кораблёв А. Ю. Искусственный интеллект в обеспечении информационной безопасности: возможности и перспективы // Информационные технологии и вычислительные системы. — 2021. — № 3. — С. 102–110.

11. Кобринец С. Ю. Интеллектуальные методы в защите информационных систем: от теории к практике // Информационная безопасность. — 2022. — № 6. — С. 18–26.

12. Нечаев В. П., Смирнов А. В. Применение методов машинного обучения в построении адаптивных систем киберзащиты // Журнал прикладной информатики. — 2023. — Т. 18, № 1. — С. 145–152.