Министерство образования и науки Российской Федерации

Государственное образовательное учреждение

высшего образования

**Московский авиационный институт**

**(национальный исследовательский университет) «МАИ»**

РЕФЕРАТ ПО ПРЕДМЕТУ «ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ»

**Применение технологий искусственного интеллекта в информационной безопасности**

Выполнил: студент группы М4В-301Б

Тихомиров А.С.

Руководитель: [Терехин Алексей Геннадиевич](https://mai.ru/education/schedule/ppc.php?guid=a9ec7c4c-1d99-11e0-9baf-1c6f65450efa)

Москва – 2021

Оглавление

[История создания 2](#_Toc61485033)

[Определение ИИ, машинного обучения и глубокого обучения в рамках кибербезопасности 2](#_Toc61485034)

[Круг современных проблем в сфере ИИ 3](#_Toc61485035)

[Применение технологий искусственного интеллекта в информационной безопасности 7](#_Toc61485036)

[Будущее кибербезопасности 9](#_Toc61485037)

# История создания

Идеи создания машин, обладающих сознанием, возникали еще в Древней Греции. В средние века и Новое время ученые создавали механизмы, заменяющие человеческий труд, например, в 17 веке Паскаль изобрел первую механическую цифровую вычислительную машину, в 19 веке Джозеф-Мари Жаккард создал программируемый ткацкий станок с инструкциями на перфокартах. В 1937 году Алан Тьрюнинг обнародовал свое изобретение – универсальную машину Тьюринга, в 1939 году в Нью-Йорке были представлены первый механический человек Electro с собакой Sparco.

Однако возможность разрабатывать программы, выполняющие сложные интеллектуальные задачи, появилась только после появления современных компьютеров после Второй мировой войны. В 1950-х годах ученые из различных областей стали задумываться о возможности создания искусственного мозга. Тогда исследования в области неврологии показали, что мозг представляет собой нейронную сеть, а А. Тьюнинг предположил, что любой вид вычислений можно представить в цифровом виде, и в 1951 году была создана первая нейронная сеть SNARC аспирантом Марвином Мински. К 1950 году А. Тьюринг разработал тест, определяющий уровень схожести действий машины с сознанием человека, впоследствии названный тестом Тьюринга. Название «искусственный интеллект» впервые было использовано на Дартмутской конференции в 1956 году, тогда же и появилась научная дисциплина «Исследование искусственного интеллекта».

Впоследствии было создано множество машин, понимающих речь человека, умеющих поддерживать беседы на заданные темы, роботов, играющих в настольные игры: знаменитый матч между компьютером и Каспаровым в шахматах закончился победой машины. Сейчас искусственный интеллект занимает важную позицию в развитии науки, особенно в рамках концепции Интернета вещей, ведь недостаточно только собирать данные, необходимо их обрабатывать, анализировать и действовать в тех случаях, когда человек этого сделать не может.

# Определение ИИ, машинного обучения и глубокого обучения в рамках кибербезопасности

В скором будущем искусственный интеллект на основе машинного обучения станет мощным инструментом обеспечения кибербезопасности. В этой сфере, как и в других отраслях, участие человека давно считается важным, незаменимым элементом. И хотя в настоящее время кибербезопасность по-прежнему во многом зависит от работы специалистов, в решении определенных задач машины постепенно начинают нас опережать.

Каждое технологическое улучшение позволяет более эффективно поддерживать роль человека в сфере безопасности. Данные разработки относятся к следующим основным областям:

*1.Искусственный интеллект (ИИ)* призван в полной мере наделить машину способностью человеческого разума к реагированию. Это основная дисциплина, охватывающая многие другие, включая машинное обучение и глубокое обучение.

*2.Машинное обучение* использует существующие модели поведения для принятия решений на основе имеющихся данных и выводов. При этом вмешательство человека по-прежнему требуется для внесения необходимых исправлений. На сегодняшний день машинное обучение, вероятно, является наиболее актуальной дисциплиной в области кибербезопасности на основе ИИ.

*3.Глубокое обучение* действует аналогично машинному – решения принимаются на основе имеющихся шаблонов, но предполагают самостоятельное внесение коррективов. В настоящее время глубокое обучение в области кибербезопасности относится к сфере машинного обучения, поэтому мы будем рассматривать в основном эту более широкую сферу.

# Круг современных проблем в сфере ИИ

*1.Ошибки конфигурирования, вызванные человеческим фактором*

С человеческим фактором связана значительная часть слабых мест кибербезопасности. Например, даже при наличии большой команды IT-специалистов правильное конфигурирование системы может быть невероятно трудной задачей. Компьютерная безопасность постоянно совершенствуется, и на сегодняшний день эта область стала более сложной, чем когда-либо. Адаптивные инструменты могут помочь в поиске и устранении проблем, возникающих при замене, модификации и обновлении сетевых систем.

Представим, что поверх старой локальной среды необходимо установить новую интернет-инфраструктуру, например систему облачных вычислений. В целях безопасности корпоративных систем команде IT-специалистов необходимо обеспечить их совместимость. Оценка надежности конфигурации вручную может стать очень трудоемким процессом, так как работникам IT-службы придется совмещать работу с бесконечными обновлениями и повседневные задачи. При наличии интеллектуальных адаптивных средств автоматизации специалисты могут оперативно получать советы по решению обнаруживаемых проблем. На основе таких средств можно даже создать систему для автоматической настройки необходимых параметров.

*2.Эффективность ручного труда при воспроизведении повторяющихся действий*

Эффективность ручного труда – еще одна проблема кибербезопасности. Процесс, выполняемый вручную, невозможно каждый раз воспроизводить в точности одинаково, особенно в такой динамичной среде, какой является современный ландшафт кибербезопасности. Индивидуальная настройка множества корпоративных [конечных устройств](https://www.kaspersky.ru/enterprise-security/threat-management-defense-solution) – одна из самых трудоемких задач. После первоначальной подготовки устройств IT-специалистам зачастую приходится снова возвращаться к ним, чтобы исправить конфигурацию или обновить настройки, которые нельзя изменить удаленно.

Не стоит также забывать, что характер угроз постоянно меняется. Если за реагирование на них отвечают люди, скорость их действий может быть снижена при столкновении с неожиданными проблемами. Система, основанная на ИИ и технологиях машинного обучения, может работать в тех же условиях с минимальной задержкой.

*3.Усталость от оповещений об угрозах*

Усталость от оповещений об угрозах может стать еще одной проблемой для организаций, не принимающих меры борьбы с ней. Чем более сложным становится многоуровневое построение систем безопасности, тем больше становится и поверхность атаки. Многие системы безопасности реагируют на известные проблемы потоком автоматических оповещений. В результате для того, чтобы найти решение и принять меры, IT-специалистам приходится анализировать их по отдельности.

Но из-за большого количества поступающих сигналов этот процесс становится очень трудоемким. В результате усталость от принятия решений становится повседневной проблемой для сотрудников служб кибербезопасности. Принятие проактивных мер по нейтрализации известных угроз и уязвимостей является оптимальным вариантом, однако многим командам не хватает времени и сотрудников, чтобы держать оборону на всех направлениях.

Иногда командам приходится сосредоточиться на наиболее острых проблемах, а второстепенные задачи отодвинуть на второй план. Использование ИИ для обеспечения кибербезопасности может помочь IT-специалистам эффективно справляться с большим количеством угроз. Противодействие каждой из них можно значительно упростить, если объединять однотипные угрозы вместе с помощью автоматической маркировки. Кроме того, некоторые проблемы может устранить сам алгоритм машинного обучения.

*4.Время реагирования на угрозу*

Время реагирования на угрозу – один из важнейших показателей эффективности службы кибербезопасности. Известно, что атаки очень быстро переходят от эксплуатации уязвимости к развертыванию. Раньше, прежде чем начать атаку, злоумышленникам приходилось вручную проверять все уязвимые места и обходными путями выводить из строя системы безопасности – иногда этот процесс мог занимать недели.

К сожалению, технологические инновации существуют не только в области киберзащиты. Сейчас автоматизация кибератак становится все более распространенным явлением. Такие угрозы, как недавно появившиеся [шифровальщики](https://www.kaspersky.com/resource-center/threats/ransomware-threats-an-in-depth-guide) LockBit, значительно сократили время, необходимое для вредоносного вторжения. Сегодня некоторые атаки успешно проводятся всего за полчаса.

Реакция человека может быть недостаточно быстрой, даже если тип атаки хорошо известен. Именно поэтому многие команды специалистов по безопасности чаще занимаются устранением последствий успешных атак, нежели предотвращают их. Отдельную опасность представляют необнаруженные атаки.

Технологии машинного обучения способны извлекать данные об атаках, группировать их и подготавливать для анализа. Они могут предоставлять специалистам по кибербезопасности отчеты, чтобы упростить обработку данных и принятие решений. Помимо отчетов, такой тип системы безопасности может также предложить рекомендуемые действия для ограничения дальнейшего ущерба и предотвращения дальнейших атак.

*5.Выявление и прогнозирование новых угроз*

Выявление и прогнозирование новых угроз – это еще один фактор, влияющий на время реагирования на кибератаки. Как отмечалось выше, задержка при реагировании возникает даже при угрозах известных типов. Новые виды атак, модели поведения и инструменты могут сбить специалистов с толку, в результате чего они будут реагировать еще медленнее. Хуже того, такие менее заметные угрозы, как кража данных, иногда могут остаться и вовсе необнаруженными. [Опрос, проведенный компанией Fugue в апреле 2020 года](https://www.fugue.co/press/releases/fugue-survey-finds-widespread-concern-over-cloud-security-risks-during-the-covid-19-crisis), показал, что примерно 84% IT-специалистов обеспокоены тем, что могут не знать об уже совершенном взломе их облачных систем.

Постоянное развитие технологий, стоящих на вооружении злоумышленников, и появление [атак нулевого дня](https://www.kaspersky.com/resource-center/definitions/zero-day-exploit) – это факторы, которые приходится всегда учитывать, строя защиту сетей. К счастью, методы кибератак обычно не изобретаются с нуля. Поскольку основой для них часто служат тактики, платформы и исходные коды прошлых атак, технологиям машинного обучения есть на чем базироваться при накоплении знаний.

Программа на основе машинного обучения поможет распознать атаку, выявив общие черты у новой угрозы и обнаруженных ранее. Машина, в отличие от человека, проведет такое сравнение быстро – что еще раз подчеркивает необходимость применения адаптивных моделей безопасности. Машинное обучение может облегчить прогнозирование новых угроз и сократить время реагирования за счет более эффективной работы с базой существующих угроз.

*6.Кадровый потенциал*

Проблема кадрового потенциала относится к числу систематических. С ней сталкиваются отделы IT и кибербезопасности множества компаний во всем мире. Иногда найти квалифицированных специалистов с необходимыми навыками может быть сложно.

Однако гораздо чаще проблема состоит в том, что наем сотрудников требует выделения немалых средств из бюджета организации. Содержание персонала требует не только оплаты повседневного труда, но и удовлетворения текущих потребностей в обучении и подтверждении квалификации. Профессионал в области кибербезопасности обязан идти в ногу со временем и быть в курсе постоянных инноваций, о которых мы упоминали выше.

Наличие инструментов на основе ИИ позволит сократить штат специалистов. Хотя им будет необходимо, постоянно повышая квалификацию, осваивать передовые достижения в области искусственного интеллекта и машинного обучения, компания сможет сэкономить время и деньги благодаря меньшей численности этих сотрудников.

*7.Адаптируемость*

В отличие от других аспектов проблема адаптируемости не так очевидна, однако может резко сказаться на возможностях службы безопасности. Специалистам может быть сложно привести свои навыки в соответствие с конкретными требованиями компаний.

Если сотрудники не знакомы с определенными методами работы, инструментами и системами, эффективность всей команды может оказаться невысокой. Даже такая, казалось бы, простая процедура, как принятие командой новых политик безопасности, может затянуться. Такова природа человека, мы не можем мгновенно освоить новые виды деятельности. На это нужно время. Однако с помощью правильных наборов данных можно превратить хорошо обученные алгоритмы в решения, соответствующие необходимым требованиям.

Применение технологий искусственного интеллекта в информационной безопасности

Для того чтобы справиться с растущим объёмом атак, производители систем защиты тоже начинают активно внедрять технологии искусственного интеллекта, машинного и глубокого обучения (ML / DL) для обнаружения, прогнозирования киберугроз, реагирования на них в режиме реального времени. Технологии искусственного интеллекта предоставляют возможность создавать решения существенно более высокой эффективности, позволяющие идентифицировать кибератаки с высокой скоростью, выбирать оптимальный ответ на инциденты безопасности, в автоматическом режиме проводить оценку актуальности и последствий инцидентов, в реальном времени вырабатывать пропорциональный ответ.  
Классификация продуктов с технологиями искусственного интеллекта по сценариям применения

Перечислим основные типы:

*1.EDR (Endpoint Detection and Response)* — платформы обнаружения атак на рабочих станциях, серверах, любых компьютерных устройствах (конечных точках) и оперативного реагирования на них. С помощью технологий ИИ продукты данной категории могут обнаруживать неизвестные вредоносные программы, автоматически классифицировать угрозы и самостоятельно реагировать на них, передавая данные в центр управления. ИИ принимает решения на основе общей базы знаний, накопленной путём сбора данных со множества устройств. Некоторые продукты данного типа используют технологии ИИ для разметки данных на конечных точках и дальнейшего контроля их перемещения, чтобы выявлять внутренние угрозы.

*2.NDR (Network Detection and Response)* — устройства и аналитические платформы, которые обнаруживают атаки на сетевом уровне и позволяют оперативно на них реагировать. Используя накопленную статистику и базу знаний об угрозах, продукты данного типа выявляют с помощью технологий ИИ угрозы в сетевом трафике и могут автоматически на них реагировать надлежащим образом, изменяя конфигурацию сетевых устройств и шлюзов. Часть продуктов данного типа специализируется на защите облачных провайдеров и их инфраструктуры. Дополнительный сценарий использования ИИ в сетевой защите — это анализ почтового трафика на предмет фишинга.

*3.UEBA (User and Entity Behavior Analytics)* — системы поведенческого анализа пользователей и информационных сущностей. Они обнаруживают случаи необычного поведения и используют их для детектирования внутренних и внешних угроз.     Основной сценарий применения ИИ-технологий в продуктах типа UEBA — это автоматическое выявление аномалий в поведенческих моделях (отклонение от нормы или соответствие шаблону (паттерну) угрозы) для пользователей и различных сущностей информационных систем. Выявленные аномалии классифицируются с помощью ИИ как различные угрозы и риски для бизнеса. Аномальное поведение может выявляться в целях мониторинга и управления доступом, обнаружения мошенничества среди клиентов или сотрудников (антифрод), защиты конфиденциальных данных, проверки соблюдения тех или иных регламентов и нормативных актов.

*4.TIP (Threat Intelligence Platform)* —платформы раннего детектирования угроз и реагирования на них, действующие на основе большого количества различных данных (Data Lake) и индикаторов компрометации (IoC).Применение ИИ позволяет повысить эффективность выявления неизвестных угроз на ранних этапах; сценарий очень схож с работой SIEM-систем, но нацелен на внешние источники данных и внешние угрозы.

*5.SIEM (Security Information and Event Management)*  — решения, которые осуществляют мониторинг информационных систем, в режиме реального времени анализируют события безопасности, поступающие от сетевых устройств, средств защиты информации, ИТ-сервисов, инфраструктуры систем и приложений, и помогают обнаружить инциденты ИБ.  В системах такого класса накапливается огромное количество данных из различных источников, а применение технологий ИИ даёт возможность выявления аномалий эвристическими методами и сокращения ложных срабатываний при изменении паттернов и моделей данных. Применение ИИ в SIEM-системах позволяет достигнуть очень высокого уровня автоматизации.

*6.SOAR (Security Orchestration and Automated Response)* — системы, позволяющие выявлять угрозы информационной безопасности и автоматизировать реагирование на инциденты. В решениях данного типа, в отличие от SIEM-систем, ИИ помогает не только проводить анализ, но и автоматически реагировать надлежащим образом на выявленные угрозы.

*7.Средства защиты приложений (Application Security)* —  системы, позволяющие определять угрозы безопасности прикладных приложений, управлять дальнейшим циклом мониторинга и устранения таких угроз. Основной сценарий применения технологий ИИ в системах защиты прикладных приложений — автоматический сбор информации об уязвимостях, атаках и заражениях, доступной в открытых источниках, и основанная на его результатах автоматизация защитных действий: сканирования на уязвимости, изменения правил защиты для веб-приложений, выявления угроз и изменения рисковой модели.

*8.Антифрод (Antifraud)* — системы, позволяющие выявлять угрозы в бизнес-процессах и предотвращать мошеннические операции в режиме реального времени. В системах защиты от мошенничества технологии ИИ применяются для определения отклонений от установленных бизнес-процессов, тем самым помогая быстро реагировать на возможное финансовое преступление или уязвимость процессов. Применение ИИ в таких системах особенно актуально, так как позволяет быстро адаптироваться к изменению логики и различных метрик бизнес-процессов, а также использовать лучшие практики в индустрии.

# Будущее кибербезопасности

Несмотря на бурные обсуждения будущего этой сферы безопасности, все же существуют ограничения, о которых следует упомянуть.

*1.Для машинного обучения необходимы наборы данных, однако в некоторых случаях их сбор и использование могут противоречить*[*законам о конфиденциальности данных*](https://www.kaspersky.ru/resource-center/threats/internet-and-individual-privacy-protection)*.* Программным системам, обучающим алгоритмы, требуется множество точек данных для построения точных моделей, что плохо сочетается с «правом на забвение». Наличие идентифицирующей человека информации в некоторых данных может являться нарушением, поэтому необходимо предусмотреть возможные решения этой проблемы. Одно из них – системы, которые делают доступ к исходным данным после обучения практически невозможным. Анонимизация точек данных также рассматривается как возможный выход, но этот метод необходимо изучить глубже, чтобы избежать искажения логики программ.

*2.Отрасли нужно больше экспертов по обеспечению кибербезопасности на основе искусственного интеллекта и машинного обучения.* Эффективность средств сетевой безопасности, основанных на технологиях машинного обучения, значительно повысится при наличии сотрудников, способных обслуживать и настраивать их по мере необходимости. Однако предложение квалифицированных специалистов на мировом рынке гораздо меньше спроса на них.

*3.Команды специалистов останутся неотъемлемой частью отделов кибербезопасности.* Жизненно важное значение для принятия решений по-прежнему будут иметь критическое мышление и творческий подход. Как уже упоминалось выше, ни технологии машинного обучения, ни ИИ пока не обладают этими качествами. Поэтому они должны быть инструментом в руках вашей команды специалистов по кибербезопасности.