На правах рукописи

Подтопельный Владислав Владимирович

МОДЕЛИ И МЕТОДИКА ОПРЕДЕЛЕНИЯ ПОСЛЕДОВАТЕЛЬНОСТЕЙ АТАКУЮЩИХ ВОЗДЕЙСТВИЙ НА СИСТЕМЫ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА ПРИ АУДИТЕ ИНФОРМАЦИОННОЙ БЕЗОПАСНОСТИ

2.3.6. Методы и системы защиты информации, информационная безопасность

Автореферат
диссертации на соискание ученой степени
кандидата технических наук

Санкт-Петербург – 2025

Работа выполнена в Федеральном государственном бюджетном образовательном учреждении высшего образования «Калининградский государственный технический университет» на кафедре информационной безопасности.

Научный руководитель: кандидат технических наук, доцент
Ветров Игорь Анатольевич

Официальные оппоненты: **Рытов Михаил Юрьевич**,
доктор технических наук, доцент,
Федеральное государственное бюджетное образовательное учреждение высшего образования «Брянский государственный технический университет»,
кафедра систем информационной безопасности,
заведующий кафедрой

Браницкий Александр Александрович,
кандидат технических наук, доцент,
Акционерное общество «Клаудран», программист

Ведущая организация: Федеральное государственное бюджетное учреждение науки «Санкт-Петербургский Федеральный исследовательский центр Российской академии наук» (СПб ФИЦ РАН), г. Санкт-Петербург

Защита состоится 10 декабря 2025 года в 15.00 на заседании объединенного диссертационного совета 99.2.038.03, созданного на базе Федерального государственного бюджетного образовательного учреждения высшего образования «Балтийский государственный технический университет «ВОЕНМЕХ» им. Д.Ф. Устинова», Федерального государственного автономного образовательного учреждения высшего образования «Санкт-Петербургский государственный университет аэрокосмического приборостроения», Федерального государственного бюджетного образовательного учреждения высшего образования «Санкт-Петербургский государственный университет телекоммуникаций им. проф. М.А. Бонч-Бруевича» по адресу: Санкт-Петербург, пр. Большевиков, д. 22, корп. 1, ауд. 554/1.

С диссертацией можно ознакомиться в библиотеке СПбГУТ по адресу Санкт-Петербург, пр. Большевиков, д. 22, корп. 1 и на сайте www.sut.ru.

Автореферат разослан 09 октября 2025 года.

Ученый секретарь
диссертационного совета 99.2.038.03,
канд. техн. наук, доцент А.Г. Владыко

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность темы исследования. Сегодня большое распространение получили атаки на системы искусственного интеллекта (ИИ). Системы ИИ могут быть интегрированы в другие информационные системы в качестве их подсистем. В таких случаях их можно назвать подсистемами искусственного интеллекта (ПИИ). Злоумышленники могут атаковать ПИИ, применяя различные методы компрометации. В большинстве случаев цели атак связаны с изменением параметров (диапазонов значений) механизма классификации в соответствии с требованиями злоумышленника без потери доверия к нему как к источнику данных. Для выявления и блокирования атак на ПИИ важно обладать большим объемом знаний о возможных действиях нападающего.

Можно выделить следующие проблемы машинного обучения (МО), напрямую влияющие на безопасность обработки данных в информационных системах (ИС):

1. Риск умышленного или случайного искажения обучающих и входных данных, что ставит под угрозу целостность и достоверность выводов модели.
2. Некорректный выбор признаков может привести к раскрытию конфиденциальных данных или сделать модель менее устойчивой к злонамеренным воздействиям.
3. Высокая вычислительная сложность может вынудить к компромиссам в области обеспечения безопасности ИИ.
4. Отсутствие у специалистов по МО глубоких знаний в области кибербезопасности.
5. Оптимизация моделей для высокой скорости работы часто достигается за счет упрощения архитектуры, что повышает риски успешных атак.

Приведенные проблемы позволяют реализовать несколько основных типовых наборов атак на системы ИИ:

1. Атаки «белого ящика», которые подразумевают полный доступ к модели машинного обучения, включая ее архитектуру, параметры и данные.
2. Атаки типа «черный ящик», которые подразумевают известность только входных и выходных данных модели. Однако с помощью различных методов атакующий может искажать выводы модели или даже получать некоторую информацию о ее внутреннем устройстве.
3. Атаки типа «серый ящик», предполагающие, что злоумышленник имеет частичные сведения о применяемой модели искусственного интеллекта.
4. Атаки, отравляющие данные. Этот тип атаки заключается во внесении изменений в обучающие данные модели.
5. Атаки, использующие уязвимости программной и аппаратной среды.

Следует отметить, что различные вычислительные модели в разной степени уязвимы к указанным атакам. В целом, уровень уязвимости определяется двумя факторами: известностью и распространенностью моделей (они могут быть типовыми или

индивидуальными), а также доступностью и распространённостью обучающих данных.

Не все из существующих методов моделирования одинаково применимы к задачам анализа атак подобного рода, поскольку уязвимости вычислительных моделей ИИ, а также их подсистем, осуществляющих сбор и обработку данных ИИ, достаточно специфичны. Атаки могут реализовываться на основе эксплуатации заданной неточности работы моделей ИИ, на основе манипуляции с исходными обучающими выборками и т.п. Таким образом, при аудите информационной безопасности возникает необходимость поиска возможных последовательностей атакующих воздействий, которые включают в свой состав эксплуатацию специфических особенностей (уязвимостей) новых интеллектуальных технологий, встраиваемых в современные информационные системы.

Степень разработанности темы. При разработке последовательности атакующих действий необходимо учитывать вероятность возникновения новых сценариев угроз. Важно также принимать во внимание возможность изменения направления атакующих действий. Вектор атаки может состоять из нескольких частей. Понимание этого особенно актуально при аудите систем, включающих элементы искусственного интеллекта.

Анализом современных типов атак на ПИИ и способов защиты от них занимаются многие исследователи. Показательными в этом случае являются работы Д.Е. Намиота и группы исследователей во главе с И. В. Котенко. Они отмечают специфические особенности реализации атак на системы ИИ, указывая ряд требований для систем защиты, основанных на методах предварительной обработки входных данных с использованием вычислительных моделей.

В сфере обеспечения информационной безопасности существует множество подходов к выявлению и анализу вредоносных воздействий. Однако эти методы не всегда принимают во внимание специфику реальных процессов обеспечения информационной безопасности систем искусственного интеллекта. Эта специфика связана с моделированием вредоносных воздействий, с использованием методик описания атак, а также с неполнотой сведений об атаке или атакуемой системе. Для выявления инцидентов безопасности применяются различные методы анализа, такие как статистический анализ, вейвлет-анализ, кластерный анализ, фрактальный анализ, метод опорных векторов, генетические алгоритмы, иммунные системы, нейронные сети, деревья решений, байесовские сети, разные экспертные системы. Следует отметить, что статистические методы без дополнительной обработки результатов могут привести к тому, что модель описания проблемной области окажется слишком привязанной к конкретному временному промежутку или слишком обобщённой. Преимущества систем, базирующихся на группе статистических методов, заключаются в следующем: они обеспечивают понятность результатов, легко адаптируются к изменениям, эффективно выяв-

ляют модифицированные атаки. Вейвлет-анализ представляет собой мощный инструмент, позволяющий эффективно анализировать большие объёмы данных, выделяя наиболее значимые компоненты и сглаживая несущественные шумы. Фрактальный анализ позволяет обнаруживать самоподобные паттерны в данных, которые могут указывать на циклические или масштабируемые атаки, проявляющиеся в разные моменты времени. Кластерный анализ помогает сегментировать данные на группы, выделяя кластеры, соответствующие нормальной работе, и кластеры-аномалии, требующие расследования. Нейронные сети часто используются для обнаружения аномалий, но они требуют значительных вычислительных ресурсов и времени на обучение.

В работах Артеменкова С. Л., Алхимова В. И., Баранова С. Н., Беляевой О. Б., Кубарева А. В., Лапсаря А. П., Думина П. Н. рассматриваются подходы к применению марковских моделей в задачах оценки и прогнозирования состояния сложных технических объектов, включая защиту от компьютерных инцидентов. Работа Марковой О. С. посвящена рассмотрению вектора атаки в зависимости от психологических склонностей нарушителя с учётом применения теории вероятности. Особенностью этого исследования является совмещение аспектов психологии нарушителя, которые автор пытается параметризировать, и аспектов процесса атаки. Работа Чечулина А.А. посвящена построению атакующих последовательностей на компьютерные сети на основе «деревьев атак». Носаль И. А. в своей работе обосновывает меры защиты информации с применением марковских методов. В других исследованиях предпринимаются попытки анализа этапов сетевых атак с учётом уязвимостей. Однако все эти работы, используя различные подходы, не учитывают специфику систем искусственного интеллекта при моделировании атак.

Цель диссертационного исследования заключается в повышении степени полноты описания атак на системы искусственного интеллекта при аудите их информационной безопасности.

В интересах решения сформулированной научной задачи и достижения цели диссертационного исследования решались задачи, указанные ниже.

Решаемая научно-техническая задача: разработка моделей и методики определения последовательностей атакующих воздействий на системы искусственного интеллекта для повышения качества процессов аудита ИБ ПИИ.

Научная задача заключается в разработке на основе методов марковских процессов принятия решений (МППР) моделей и методики построения и анализа атак на системы ИИ с учетом современной специфики формирования сценариев атак и рекомендаций в области ИБ при аудите информационных систем. Полнота при моделировании атак на ПИИ с использованием МППР, в данном случае, представляется как способность модели учитывать все состояния атаки и переходы между ними, возникающие при воздействии атакующего на систему ИИ.

Для достижения данной цели в диссертационной работе поставлены и решены следующие частные задачи:

1. Произведен анализ существующих подходов поиска и оценки событий безопасности в информационных системах с элементами ИИ.

2. Определены возможности использования доступных наборов данных для разработки моделей формирования последовательностей атакующих воздействий на системы ИИ, определены параметры используемых при моделировании данных.

3. Разработаны модели построения и анализа атак на ПИИ при определении мер противодействия атакам.

4. Разработаны методика и алгоритм, позволяющие моделировать атакующие воздействия, используемые в процессе аудита ИБ ПИИ, изучать их динамику, использовать для составления сценариев атак.

5. Разработана архитектура программного решения, которое позволяет автоматизировать процессы моделирования.

Объектом исследования являются атаки, направленные на эксплуатацию уязвимостей моделей и архитектур подсистем искусственного интеллекта в контексте общей архитектуры корпоративной информационной системы, процессы построения и анализа атакующих последовательностей для повышения качества аудита защищенности систем искусственного интеллекта.

Предметом исследования выступают вычислительные модели искусственного интеллекта, методики и алгоритмы моделирования атак на подсистемы ИИ.

Научная новизна результатов исследования заключается в следующем (все результаты, выносимые на защиту, являются новыми):

1. С применением предложенного аппарата (моделей, методики, алгоритма, программного решения) на основе марковских процессов принятия решений появится возможность с помощью математических методов обосновать построение сценария атаки как последовательности событий безопасности с учетом специфики систем ИИ. Это позволит улучшить совместную работу смежных систем контроля и поиска событий безопасности (за счет введения методов анализа на основе МППР), повысить качество аудита ИБ.

2. Предлагается использовать совмещение нескольких методов анализа признаков событий безопасности, связанных с компрометацией ПИИ, ориентируясь на вычислительные методы, основанные на МППР. Комплекс марковских моделей, с учетом формализации типовых атакующих последовательностей (техник и тактик), позволяет выявлять и учитывать ранее необнаруженные этапы развития вредоносного воздействия на информационные системы с ПИИ. При анализе учитываются: топология сети предприятия, архитектурные особенности подсистем ИИ, методики описания действий атакующего, применяемые при аудите, и другие факторы.

3. Разработаны модели, методика и алгоритм определения атакующих последовательностей для сценариев атак на ПИИ, формируемых в процессе аудита ИБ ПИИ.

Теоретическая и практическая значимость результатов исследования. Разработанные модели представляют собой научно-методическую основу для определения последовательностей атакующих воздействий на системы ИИ и обоснования формируемых сценариев атак, используемых при проведении аудита ИБ, что в конечном итоге позволит повысить качество защиты систем ИИ. Практическая реализация позволяет построить прогноз проявления событий безопасности как этапов атак. При этом повышается точность и полнота построения сценария атаки, что позволяет на практике эффективно применять разработанный подход к формированию модели угроз.

Методология и методы исследования. В качестве математических положений, используемых в диссертации, применены: МППР для определения последовательностей атакующих воздействий на ПИИ, механизмы машинного обучения как вспомогательные элементы для регистрации и анализа событий с применением аналитико-статистических методов.

Положения, выносимые на защиту:

1. Модели определения последовательности и анализа атакующих воздействий на ПИИ.
2. Алгоритм определения последовательности действий и состояний при атаке на ПИИ.
3. Методика определения последовательностей атакующих воздействий на ПИИ.
4. Архитектура и программные компоненты системы построения последовательности атакующих воздействий на ПИИ на основе разработанных моделей МППР.

Степень достоверности результатов. Обоснованность и достоверность представленных в диссертационной работе научных положений обеспечивается за счет тщательного анализа состояния исследований в заданной области, подтверждается согласованностью результатов, полученных при компьютерной реализации, успешной апробацией основных теоретических положений диссертации на ряде научных конференций всероссийского и международного уровня, а также публикацией основных положений, раскрывающих результаты работы, в ведущих рецензируемых научных изданиях.

Соответствие диссертации научной специальности. Представленные результаты соответствуют специальности 2.3.6. Методы и системы защиты информации, информационная безопасность.

Апробация результатов. Научные результаты, полученные в диссертации, внедрены в научно-исследовательскую работу, образовательный процесс и практику деятельности в ФГБОУ ВО «Калининградский государственный технический университет» (г. Калининград), ФГАОУ ВО «Балтийский федеральный университет им. И. Канта» (г. Калининград), ООО «Центр защиты информации» (г. Калининград), ООО «Радиоэлектронные системы» (г. Екатеринбург).

Основные положения и результаты докладывались и обсуждались на следующих конференциях:

1. Международной научно-практической конференции VIII Международного Балтийского морского форума (Калининград, 2020 г).
2. III Международной научной конференции «Экосистемы без границ - 2022» IX Международного Балтийского морского форума (Калининград, 2021 г).
3. X национальной научной конференции с международным участием «Морская техника и технологии. Безопасность морской индустрии» в рамках X Международного Балтийского морского форума (Калининград, 2022 г).
4. XVIII Всероссийской научно-практической конференции «Информационная безопасность цифровой экономики» в рамках форума информационной безопасности «Сибирь-Дальний Восток-2022» (Хабаровск, 2022).
5. XI Национальной научной конференции с международным участием «Морская техника и технологии. Безопасность морской индустрии», в рамках XI Балтийского морского форума (Калининград, 2023 г).
6. XIX Всероссийской научно-практической конференции «Информационная безопасность цифровой экономики» в рамках форума информационной безопасности «Сибирь-Дальний восток-2023» (Улан-Удэ, 2023).
7. Всероссийской научно-технической конференции «Актуальные проблемы радиоэлектроники и телекоммуникаций» (Самара, 2024).
8. V Всероссийской научно-практической конференции «Социотехнические и гуманитарные аспекты информационной безопасности» (Пятигорск, 2024).

Публикации. Основные результаты диссертации изложены в 17-ти публикациях, в том числе, в 5-ти статьях, опубликованных в ведущих рецензируемых журналах, входящих в перечень ВАК, в материалах четырех международных конференций. Получено 7 свидетельств о государственной регистрации программ для ЭВМ.

Личный вклад соискателя. Все выносимые на защиту результаты получены лично автором. Лично разработаны модели определения последовательности атакующих воздействий на ПИИ, алгоритм и методика формирования последовательности атакующих воздействий. Результаты моделирования могут использоваться при аудите ИБ и для обоснования повышения степени защиты ПИИ.

Структура и объем работы. Диссертационная работа изложена на 250 машинописных страницах, включает 4 главы, 50 рисунков, 20 таблиц и список литературы (136 наименований).

СОДЕРЖАНИЕ РАБОТЫ

Во введении обоснована актуальность темы диссертации, сформулированы цели исследования, решаемая научно-техническая задача и положения, выносимые на защиту, отражена суть и новизна основных научных результатов.

В первой главе диссертации проведен анализ целей, задач и возможностей моделирования атакующих воздействий с использованием марковских процессов принятия решений. Рассмотрены особенности проведения аудита систем ИИ с учетом известных способов (методик) описания атакующих воздействий (в том числе MITRE ATLAS и Методики оценки угроз безопасности информации (ФСТЭК РФ) (далее – Методика ФСТЭК)). Проанализированы современные способы атак на подсистемы ИИ, определена их специфика. Отмечается, что в общем случае опасность уязвимости зависит от известности и распространённости моделей, обучающих данных и наличия уязвимостей в компонентах ПИИ. Также выделяются наборы объединённых в блоки компонентов ПИИ, которые способны функционировать в режимах обучения и эксплуатации в составе ИС (при этом они реализуют логику работы ИИ (уровень логики ИИ)): блок вычислительной модели (включает в свой состав собственно вычислительную модель, а также интерфейс доступа к ней в режиме её эксплуатации), блок датасетов. При атакующих воздействиях на них минимизируется потребность в прямом взаимодействии с технической вычислительной инфраструктурой. К двум блокам, которые актуальны для злоумышленника при атаках, направленных на изменение логики работы ИИ, следует добавить, как отдельный блок, набор компонентов технического уровня ИС, связанный с работой ПИИ. Он обеспечивает вычисления. На основе анализа известных способов моделирования атак, с учетом специфики ПИИ, предлагается в качестве метода моделирования использовать марковские процессы принятия решений, что позволяет на основе результатов моделирования формировать наиболее вероятные сценарии атак на ПИИ, используемые при аудите. Сформулирована задача исследования. Она заключается, прежде всего, в разработке: (1) моделей построения и анализа атак на системы ИИ, то есть моделей, предназначенных для определения последовательности и анализа атакующих воздействий на ПИИ; (2) алгоритма определения последовательности действий и состояний при атаке на ПИИ; (3) методики определения последовательностей атакующих воздействий на ПИИ; (4) архитектуры программного решения и программных компонентов системы построения последовательности атакующих воздействий на ПИИ на основе разработанных моделей МППР.

Во второй главе определяется специфика применения методов моделирования атак на ПИИ с использованием МППР. Описание последовательностей состояний и действий атаки основывается на функциях ценности (полезности) с учетом специфики классификации действий атакующей стороны (в том числе, с использованием тактик и техник методик описания атак (MITRE ATLAS или Методики ФСТЭК)). На основе параметров вероятностей переходов и начальных состояний формируется граф состояний

и действий атаки. Для этого необходимо ассоциировать тактики (этапы атаки) и состояния, которые будут являться вершинами графа состояний и действий. Отмечается, что состояния и переходы между ними должны быть сопоставлены этапам атаки, которые соответствуют степени компрометации и показывают развитие атаки. Таким образом, достигается упорядоченность в последовательности состояний и действий. Для этого в модель интегрируются техники и тактики популярных баз знаний и методик, предлагаются способы их обобщения.

Рассматривается два представления о переходах между состояниями при моделировании атак. Первое представление предполагает: передвижения злоумышленника с помощью действий через уязвимости рассматриваются в режиме имитации развертывания атаки без априорного предположения о логичности действий злоумышленника (режим on-line). При этом переходы будут многочисленны, допускаются возвраты в ранее пройденные состояния. Второе представление (режим off-line), актуальное при плановом аудите, позволяет определить наиболее эффективный путь атаки, исходя из того, что при атаке не повторяются уже ранее достигнутые этапы (не учитывается алогичность действий злоумышленника), так как возвратные состояния отдаляют злоумышленника от поставленной цели.

При моделировании последовательности атакующих воздействий, представляемой в виде графа состояний и действий, принимается во внимание следующая информация о компонентах анализируемой системы ИИ: инфраструктурная (сетевая) идентификация объекта, описываемого в контексте состояний атаки; состояния атаки; идентификатор (ID) уязвимости (при этом сетевые узлы (компоненты ПИИ) могут иметь более одной уязвимости); действия злоумышленника (тип действий); CVSS-оценка уязвимости, другие параметры функции $\mathrm{R}(s, a, s')$.

Марковская модель атаки описывается кортежем (S, A, P, R, γ). В моделировании МППР учитывается специфика проблемной области ИБ ПИИ, что отражено в таблице 1. Параметры модели определяются специалистом ИБ. Допускается для определения параметров использовать системы автоматизированного сбора информации. Оптимальная стратегия нападения строится с использованием уравнения Беллмана.

При моделировании атак на системы ИИ (в частности, на их модели) необходимо учитывать специфику принятых методологий построения последовательностей атакующих воздействий. Методология MITRE ATLAS уже включает в свой состав действия, обозначающие не только перемещение между инфраструктурными компонентами, но и между состояниями, которые сопоставлены этапам компрометации вычислительной модели ИИ или подсистемы ИИ в целом. Соответственно, действия злоумышленника при атаке (с учетом уровня их детального описания) могут быть соотнесены с содержательной частью тактик MITRE ATLAS, а их осуществление означает развертывание состояния атаки. Так фиксируются действия на некотором этапе атаки. Таким же образом можно использовать альтернативные методологии описания атак, схожие по своим принципам организации с MITRE ATLAS.

Таблица 1 - Параметры модели определения последовательности атакующих воздействий на ПИИ

Параметр	Пояснение
Множество состояний (S)	Это множество этапов атаки на ПИИ. Каждое состояние представляет определенный этап компрометации системы. При детальном описании состояний используются тактики следующих методик и баз знаний: MITRE ATLAS, Методики ФСТЭК (частично).
Множество действий (A)	Это множество переходов, представляющее собой набор действий, направленных на продвижение атаки до достижения цели злоумышленника. При детальном рассмотрении атак действия соотносятся с техниками тактик методик и баз описаний атак (MITRE ATLAS, Методики ФСТЭК). Используются типы действий злоумышленника (легальные (*C*), нелегальные (*D*), действия возврата в пройденные состояния или сброса (*R*)).
Функция вознаграждения (R)	Это функция определения награды $R(s, a, s')$ за переход в определенное состояние атаки. С помощью нее можно охарактеризовать эффективность или сложность действий при эксплуатации уязвимостей. Модификация функции приведена ниже (4).
Вероятность переходов (P)	Это вероятность успешного перехода $P(s, a, s')$ между состояниями атаки. Отражает неопределённость эксплуатации уязвимостей.
Коэффициент дисконтирования (γ)	Это коэффициент дисконтирования ($0 \leq \gamma \leq 1$). Он необходим для регулирования скорости сходимости.
Функция ценности $V^*(s)$	Это оптимальная ожидаемая совокупная награда, которую можно получить, начиная из состояния s и следуя оптимальной стратегии. Является целевой функцией для оптимизации стратегии атаки.
Оптимальная политика $\pi^*(s)$	Это функция, определяющая оптимальное действие для каждого состояния. Реализует стратегию атаки на тактическом уровне. Ее стандартное определение в МППР следующее: $\pi^*(s) = \underset{a}{\operatorname{argmax}} \sum_{s'} P(s' \mid s, a,) \ [\mathrm{R}(s, a, s') + \gamma V^*(s')]$.

В соответствии с матрицей MITRE ATLAS и Методикой ФСТЭК РФ (частично) формируется обобщенный перечень тактик и техник реализации угроз, направленных на системы машинного обучения. Подобный подход, учитывая использование МППР, позволяет рассматривать атаки как последовательности атакующих воздействий и серии сменяющихся состояний, фиксирующих приближение нарушителя к цели. При

этом возможно отследить специфику сопряжённости состояний, определить наиболее опасные и наименее опасные последовательности (определяются по величине $V_{i+1}^{*}(s)$). Функция ценности (полезности) каждого действия с учетом его типа (*D, C, R*) может определяться отдельно. Действия рассматриваются в контексте следующих функций (1-3):

$$V_{i+1}^{*}(S, a = \mathrm{D}) = \sum_{s' \in S} P(s, \mathrm{D}, s')\ [\mathrm{R}\ (s, D, s') + \gamma V_i^{*}(s')]. \quad (1)$$

$$V_{i+1}^{*}(S, a = \mathrm{C}) = \sum_{s' \in S} P(s, \mathrm{C}, s')\ [\mathrm{R}\ (s, C, s') + \gamma V_i^{*}(s')]. \quad (2)$$

$$V_{i+1}^{*}(S, a = \mathrm{R}) = \sum_{s' \in Sprev} P(s, \mathrm{R}, s')\ [\mathrm{R}\ (s, R, s') + \gamma V_i^{*}(s')]. \quad (3)$$

В представленных формулах для сохранения единообразия записи используется $V_{i+1}^{*}(s, a)$, чтобы явно указать связь с итерационным процессом вычисления ценности состояний (оценку полезности действий допускается обозначать как $Q_{i+1}(s, a)$). При определении параметров функции вознаграждения в работе предлагается перевод метрик оценки систем ИИ (в частности, нейросетей) в метрики CVSS при учете того, что CVSS предназначается для оценки уязвимостей и их влияния на безопасность, тогда как AUC и другие метрики относятся к производительности моделей классификации. Для вычисления вознаграждения за переход из одного состояния в другое в модели используется основная формула (4):

$$\mathrm{R}(s, a, s, n) = \Big(w_{cvss}\ \cdot CVSS(s, s') + w_i \cdot I(s, s') + w_y \cdot Y(s, s') + w_a \cdot A(s') + w_l \cdot L(s, a)\Big) \cdot D(n) - \delta, \quad (4)$$

где: R(*s*, *a*, *s′*, *n*) — общее вознаграждение за переход из состояния s в состояние *s′* при выполнении действия *a* и использовании уязвимости *n* раз; *CVSS*(*s*,*s′*) — оценка уязвимости по метрикам CVSS для перехода из состояния *s* в состояние *s′*; *I*(*s*,*s′*) — индикатор возможности взаимодействия между узлами сети (0 или 1); *Y*(*s*,*s′*) — индикатор возможности сопряжения уязвимостей (0 или 1); *A*(*s′*) — индикатор приближения к целевому состоянию по методике (MITRE ATLAS или ФСТЭК); *L*(*s*,*a*) — индикатор легальности действия (0 или 1); *n* — количество раз, когда уязвимость использовалась (количество переходов); *D*(*n*) — коэффициент затухания; w_{cvss} – величина важности параметра CVSS; w_i – величина важности параметра *I*(*s*,*s′*); w_y – величина важности параметра *Y*(*s*,*s′*); w_a– величина важности параметра *A*(*s′*); w_l– величина важности параметра *L*(*s*,*a*); δ - штраф за шаг.

Определяется специфика использования тактик (действий, техник) злоумышленника с учетом особенностей организации систем ИИ. Рассматривается возможность реализации тактик с учетом доступа к датасетам, вычислительной модели ПИИ, инфраструктурным элементам систем ИИ (или ПИИ).

В третьей главе приводятся модели, построенные на основе МППР (описываются набором состояний, переходов между ними, функциями ценности (полезности)), а также методика и алгоритм, используемые для формирования и анализа последовательности атакующих воздействий на ПИИ. При необходимости использования модели нарушителя ее специфика, как и доступные нарушителю способы взаимодействия с атакуемой системой, определяется в соответствии с представляемыми видами атакующих воздействий. При этом учитывается доступность информации о компонентах ПИИ и специфика доступности самих компонентов, а также местоположение нарушителя относительно ПИИ. При моделировании в качестве основного метода расчетов используется метод итераций по значениям для оценки состояний. Каждая функция ценности (полезности) обновляется на основе текущих значений и максимизации ожидаемого вознаграждения нарушителя с учетом вероятностей переходов между состояниями атаки. Вводятся уровни абстракции, которые определяются как степени детализации состояний и атакующих воздействий при атаке. Причина введения уровней абстракции – необходимость снижения сложности описания сценария атаки при наличии множества учитываемых параметров и компонентов. С учетом уровней абстракции формируются три типа моделей:

1. **Типовые (общие) модели.** Данные модели используются для общего анализа специфики атаки на системы ИИ. Они не предполагают ассоциирование состояний и атакующих действий с тактиками и техниками. Состояния моделей и переходы между ними приведены на рисунке 1.

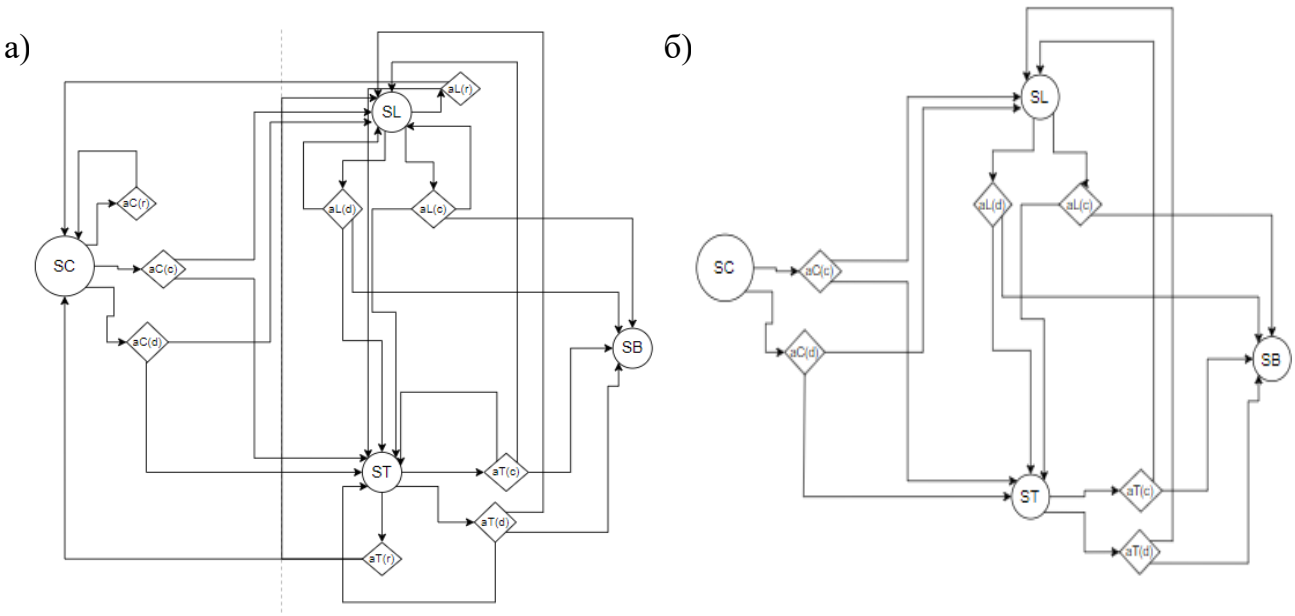

Рисунок 1 – Графы состояний и действий общих моделей для режимов on-line (а) и off-line (б)

2. **Упрощенные модели.** В данных моделях используются состояния и действия, которые ассоциированы с группами тактик MITRE ATLAS в режимах моделирования on-line и off-line.

3. **Подробные (полные) модели.** В данных моделях действия ассоциируются с техниками, а состояния прямо ассоциируются с тактиками MITRE ATLAS. В этом случае тактики рассматриваются отдельно, каждая с совокупностью всех принадлежащих ей техник.

При моделировании учитываются режимы работы ПИИ: режим обучения, режим эксплуатации. Формируются последовательности действий и состояний атаки с учетом стратегии вознаграждений. Формулируются допущения, которые необходимо учесть в процессе моделирования при решении задач аудита (аудит плановый, превентивный, при аудите учитывается возможность тестирования систем). Допущения связаны с предположением защищающейся стороны о действиях злоумышленника. Соответственно, требуется определить целевые состояния, важные для защищаемой системы, поскольку подразумевается, что к ним стремится злоумышленник.

Особенности процесса моделирования следующие:

1. Злоумышленник способен использовать как легитимные, так и нелегитимные способы взаимодействия с ИИ в процессе атаки, что влияет на величину вознаграждения в модели МППР.

2. Ядром каждой модели МППР являются функции ценности (полезности), на основе которых производится поиск лучших состояний и оптимальных стратегий (поиск стратегий производится по политике π^*). В итоге это позволяет определить наилучшие и наихудшие для злоумышленника последовательности.

3. После итераций по значениям используются формулы определения оптимальных действий, которые могут быть различными и включать целые множества *D, C, R*. В формулах допустимо использовать общее обозначение *S* вместо конкретных состояний (например, *SP*). Это позволяет применять формулы к любому состоянию системы ИИ.

4. Важно отметить, что поскольку состояние (тактика как этап) в последовательности атакующих воздействий интерпретируется как получение злоумышленником новых возможностей для осуществления дальнейших компрометирующих действий, то множество *A* будет включать в свой состав те действия, которые позволяется реализовывать в достигнутом состоянии атаки. В данном случае под действиями подразумеваются техники тактик, которые входят во множества действий, включающих классы: *D, C* или *R*.

5. Допустимо использовать различные методы для расчета вероятностей переходов, опираясь на характеристики уязвимостей, размеры вознаграждений и алгоритмы машинного обучения, которые применяются при анализе инцидентов безопасности.

Разработанные модели предполагают следующее:

1. **Типовые (общие) модели на основе МППР** предназначены для общего (начального) анализа специфики атаки в режиме on-line и off-line (5-7). Модели подобного типа позволяют исследовать общее состояние безопасности ПИИ (минимально детализированное) или одно из состояний более сложной модели. Эти модели не предназначены для детального описания тактик и методов атак на ПИИ. Однако они позволяют получить общее представление о том, как злоумышленник может взаимодействовать с системой. Используются следующие основные состояния: контролируемое взаимодействие (SC), легальное взаимодействие (SL); доверенное взаимодействие (ST). Также добавляется дополнительное состояние блокировки (SB), не предусматривающее действий. Набор функций ценности (полезности) для данных моделей (on-line, off-line) в обобщенном виде следующий:

$$V_{i+1}^{*}(s = ST) = \max_{a \in A} \sum_{s' \in S} P(ST, a, s') \ [\mathrm{R}(ST, a, s') + \gamma V_i^{*}(s')] \,. \quad (5)$$

$$V_{i+1}^{*}(s = SL) = \max_{a \in A} \sum_{s' \in S} P(SL, a, s') \ [\mathrm{R}(SL, a, s') + \gamma V_i^{*}(s')] \,. \quad (6)$$

$$V_{i+1}^{*}(s = SC) = \max_{a \in A} \sum_{s' \in S} P(SC, a, s') \ [\mathrm{R}(SC, a, s') + \gamma V_i^{*}(s')] \,. \quad (7)$$

2. **Модели с упрощенной классификацией тактик MITRE** предполагают объединение тактик на основе сходства функционального назначения техник (действий) в несколько классов состояний. В этом случае учитывается возможность ограничения использования тактик, исходя из специфики доступа: к датасетам, к вычислительной модели ПИИ, к инфраструктурным элементам систем ИИ (или ПИИ). Также при моделировании учитываются уровни доступности ПИИ для действий злоумышленника и контроля им состояний атакуемой ПИИ. Выделяются следующие уровни: уровень, где действия возможны без обязательного взаимодействия с моделью ИИ; уровень, где действия возможны, когда вычислительная модель доступна; уровень, где действия возможны, но требуется обязательное контролируемое злоумышленником взаимодействие с моделью ИИ.

Набор функций ценности (полезности) для моделей on-line и off-line с учетом специфики доступности действий и состояний (*S*Р (Разведка), *S*П (Подготовка), *S*Д (Доступ), *S*В (Выполнение), *S*Б (Блокировка)) в обобщенном виде следующий (8-12):

$$V_{i+1}^{*}(S\mathrm{Р}) = \mathrm{R}(S\mathrm{Р}) + \gamma max_{a\mathrm{Р}} E[\mathrm{R}(S\mathrm{Р}, a\mathrm{Р}, s') + V_i^{*}(s')]. \quad (8)$$

$$V_{i+1}^{*}(S\mathrm{П}) = \mathrm{R}(S\mathrm{П}) + \gamma max_{a\mathrm{П}} E[\mathrm{R}(S\mathrm{П}, a\mathrm{П}, s') + V_i^{*}(s')]. \quad (9)$$

$$V_{i+1}^{*}(S\mathrm{Д}) = \mathrm{R}(S\mathrm{Д}) + \gamma max_{a\mathrm{Д}} E[\mathrm{R}(S\mathrm{Д}, a\mathrm{Д}, s') + V_i^{*}(s')]. \quad (10)$$

$$V_{i+1}^{*}(S\mathrm{В}) = \mathrm{R}(S\mathrm{В}) + \gamma max_{a\mathrm{В}} E[\mathrm{R}(S\mathrm{В}, a\mathrm{В}, s') + V_i^{*}(s')]. \quad (11)$$

$$V_{i+1}^{*} = (s = S\mathrm{Б}) = \mathrm{R}(S\mathrm{Б}). \quad (12)$$

3. Подробная (полная) модель предполагает полное описание и использование тактик и техник MITRE ATLAS (представлена на рисунке 2 для режима on-line, для

представления в режиме off-line достаточно исключить обратные переходы). Для получения полного графа состояний атаки требуется, чтобы механизмы ПИИ позволяли использовать состояние S1 (Тактика 1). Модель позволяет определить наилучшую последовательность состояний атаки с точки зрения злоумышленника при всех возможных переходах. Основным механизмом регулирования использования тактик (состояний) и техник (действий) является функция вознаграждения. Допускается неточность описания инфраструктуры ПИИ. Соответственно, инфраструктурные ограничения необязательны.

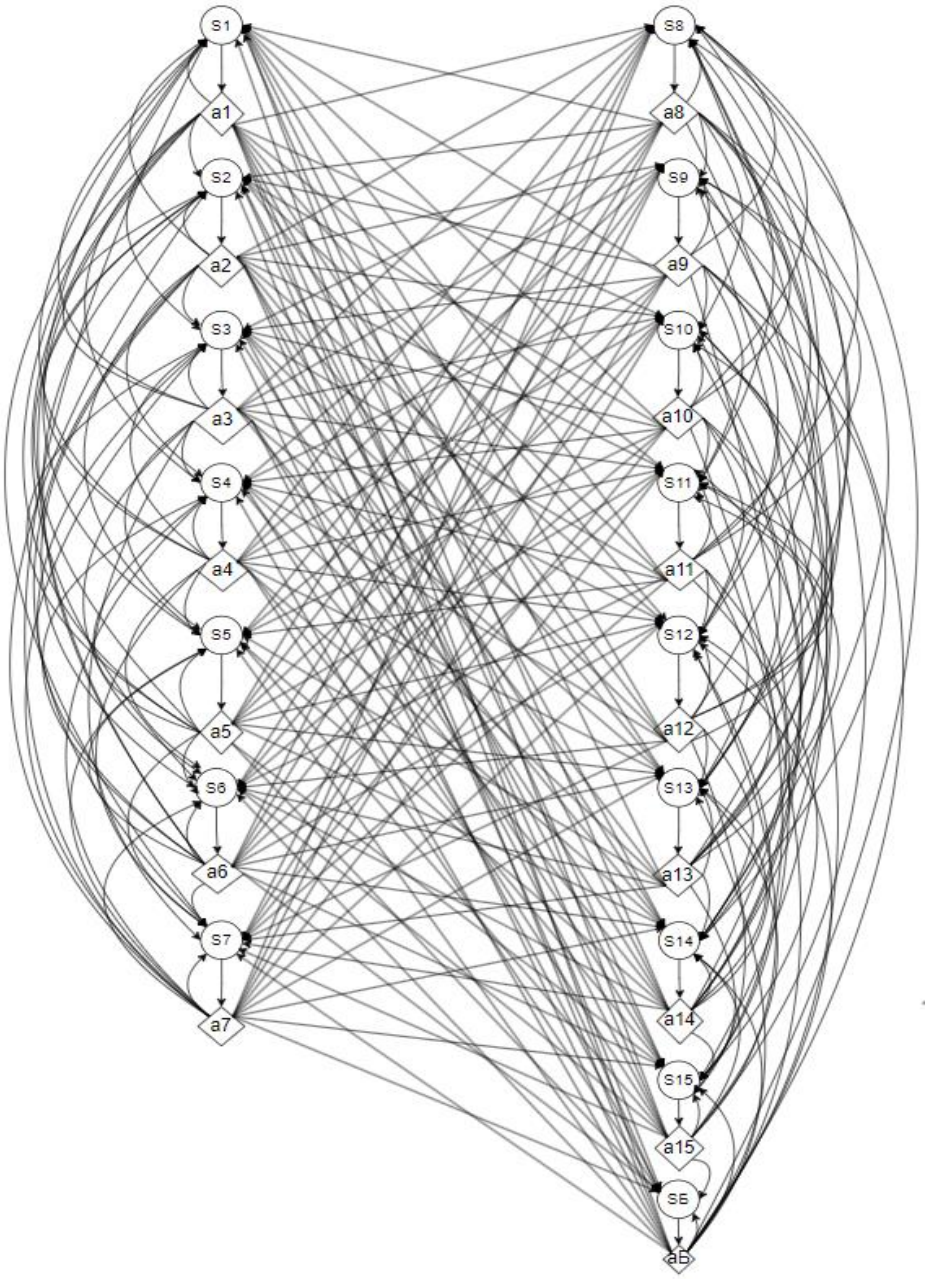

Рисунок 2 - Состояния и переходы подробных моделей МППР для атак с учетом актуальности MITRE ATLAS при использовании режима on-line

Определение значений для каждого состояния с использованием функций полезности подробных моделей на основе МППР для атак в режиме on-line и off-line на ПИИ приведено в выражениях ниже (13-14).

Для режима on-line обобщённая функция полезности следующая (в ней используется множество S_{all}={S1,S2,...,Sn,SБ}, то есть учитываются все состояния S_i в соответствии с числом актуальных тактик базы MITRE ATLAS (на период исследования в базе существует 15 тактик), а также добавляется состояние блокировки SБ, при этом вводятся состояния S_k={S1,S2,..., S_k}, число которых ограничено спецификой действия (перехода) R (сброс)):

$$V_{i+1}^{*}(s = S_k) = max\begin{Bmatrix}\sum_{s\prime \in S_{all}} P(S_k, D, s')[\mathrm{R}(S_k, D, s') + \gamma V_i^{*}(s')] \\ \sum_{s\prime \in S_{all}} P(S_k, C, s')[\mathrm{R}(S_k, C, s') + \gamma V_i^{*}(s')] \\ \sum_{s\prime \in S_k} P(S_k, R, s')[\mathrm{R}(S_k, R, s') + \gamma V_i^{*}(s')]\end{Bmatrix}. \quad (13)$$

Для режима off-line обобщённая функция полезности следующая (в ней используется $S_{current \to end}$ – множество состояний, начинающееся с текущего и включающее все последующие состояния, а также отдельное состояние SБ, при этом для S_k подразумевается k ∈{1,2,...,n, Б}):

$$V_{i+1}^{*}(s = S_k) = max\begin{Bmatrix}\sum_{s\prime \in S_{current \to end}} P(S_k, D, s')[\mathrm{R}(S_k, D, s') + \gamma V_i^{*}(s')] \\ \sum_{s\prime \in S_{current \to end}} P(S_k, C, s')[\mathrm{R}(S_k, C, s') + \gamma V_i^{*}(s')]\end{Bmatrix}. \quad (14)$$

Основные шаги алгоритма определения (формирования и анализа) последовательности атакующих воздействий на ПИИ (действий и состояний) с использованием моделей МППР при атаке на ПИИ включают следующее:

1. Определение набора действий и, следовательно, состояний модели. Следует учитывать следующее: набор действий и состояний определяется в соответствии с MITRE ATLAS (при необходимости, в сопряжении с Методикой ФСТЭК с учетом ее специфики) или в соответствии с альтернативой (модификацией) MITRE ATLAS; определяются условия, при которых переходы в следующие состояния последовательности возможны из предыдущих состояний.

2. Определение размера наград для каждого действия с учётом его стоимости (при необходимости). Следует учитывать следующее: каждое возможное действие сопряжено с уязвимостью, для которой высчитывается контекстная метрика CVSS; задается награда (штраф) за переход в состояние блокировки.

3. Определение допустимых переходов, вероятностей переходов между состояниями для каждого действия (вероятность перехода в следующее состояние, вероятность перехода в состояние блокировки). Определяется необходимый режим моделирования: off-line или on-line.

4. Проведение вычислений. Дополнительно требуется провести анализ зависимости оптимальной политики от стоимости (специфики) действий с последующим определением наиболее выгодных для злоумышленника переходов между состояниями.

Методика определения последовательностей атакующих воздействий на ПИИ (на системы ИИ) в процессе построения сценариев атак, используемых при проведении аудита ИБ показана на рисунке 3.

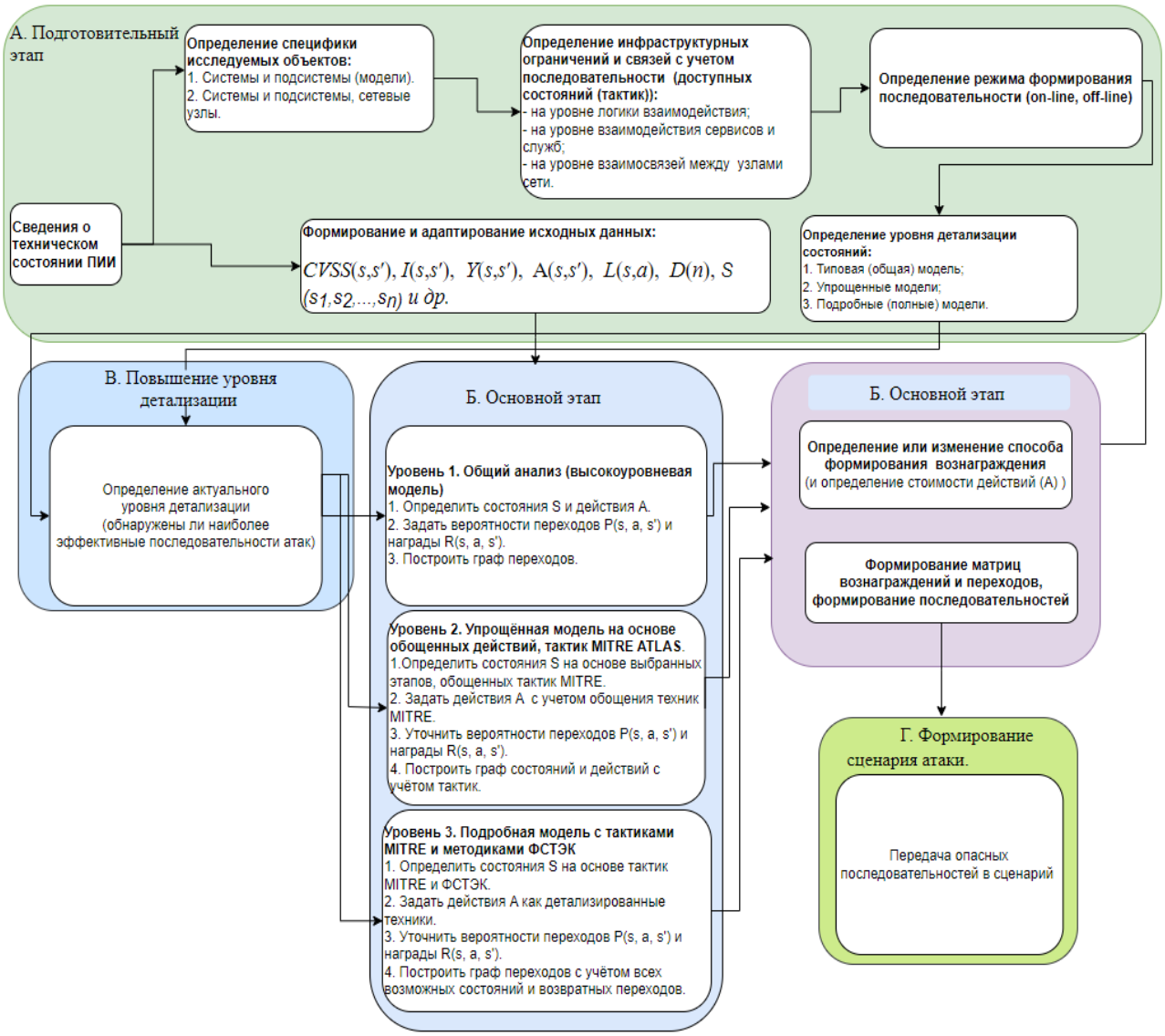

Рисунок 3 – Методика определения последовательностей атакующих воздействий на системы ИИ (ПИИ) (с возможностью выявления наилучшей для злоумышленника последовательности)

Методика включает следующие обобщенные этапы с соответствующими особенностями реализации:

А. Подготовительный этап. Важно отметить, что на данном этапе производится определение исходных данных и входных параметров (приведены в формуле вознаграждения), при этом используется техническое описание ПИИ для формирования связей состояний на техническом и логическом уровнях, производится определение способа моделирования.

Б. Основной этап (применяется к выбранному режиму и уровню моделирования) предполагает реализацию алгоритма определения последовательности состояний и действий (учитывается доступность действий). Соответственно, формируется список опасных последовательностей.

В. Дополнительный этап используется в том случае, когда необходим выбор последовательного прохождения уровней детализации атакующих воздействий при моделировании. После этого предполагается повтор этапа Б.

Г. Итоговый этап предполагает формирование сценария атаки.

В четвертой главе приводится экспериментальная оценка разработанных моделей. Полнота при моделировании атак на ПИИ с использованием МППР в данном случае представляется как способность модели учитывать все допустимые состояния атаки и переходы между ними. При испытании моделей полнота описания набора действий и состояний достигается следующим образом: рассматриваются все возможные состояния системы с учетом методики описания, максимально полное описание состояний достигается при использовании тактик MITRE ATLAS; учитываются типы действий (нелегальное взаимодействие (*D*), легальное взаимодействие (*C*)), а также все возможные действия атакующего с учетом методики описания сценария атаки (действия как техники из MITRE ATLAS или Методики ФСТЭК); учитываются обратные действия атакующего, возвращения злоумышленника в предыдущие состояния атаки (действия сброса (*R*)) в режимах on-line; при поиске наилучшей последовательности атакующих воздействий перебираются все варианты состояний и переходов между ними. Упорядоченность состояний и действий, основанная на известных базах знаний и методиках описания атак, позволяет формировать сценарий нападения как набор последовательных этапов развития атаки (например, от разведки до тактики-состояния «выполнение»). Обновление описания состояний и действий допускается при изменении баз знаний и методик описания атак.

В отличие от основанных на иных методах моделей, которые не принимают во внимание динамические особенности атак (например, действия сброса (R)) и не классифицируют действия по типам (включая их соответствие тактикам MITRE ATLAS), предлагаемые модели на основе МППР позволяют задействовать и учитывать значительно большее число действий и состояний. Это увеличивает меру сложности моделирования, а, следовательно, понижает скорость расчетов. Однако при процедурах аудита ИБ ПИИ, реализуемых до начала инцидента безопасности, скорость расчетов менее важна, чем полнота описания сценария возможной атаки. При рассмотрении полноты (15) учитывается следующее: $|A|$ – мощность множества действий всех типов для

состояний модели (используются наборы нелегальных (D), легальных (C), возвратных (R) действий); $|A'|$ – мощность множества действий без учета возвратных действий; $|A''|$ – мощность множества действий без учета специфики их типов; также учитываются все существующие связи между выделенными компонентами системы ИИ (ПИИ).

$$|A| \geq |A'| \geq |A''|. \quad (15)$$

Снижение меры сложности моделей достигается путём учёта следующего: технологических ограничений функционирования систем ИИ; уровней детализации описания атакующих воздействий (от описания логики работы искусственного интеллекта (важно при атаках на логику работы вычислительной модели искусственного интеллекта) до описания с учётом каждого узла и его уязвимостей).

Следует отметить, что количество действий увеличивается с приростом количества элементов системы ИИ. На рисунке 4 демонстрируется изменение количества действий в зависимости от количества узлов (компонентов), присутствующих в инфраструктуре системы ИИ.

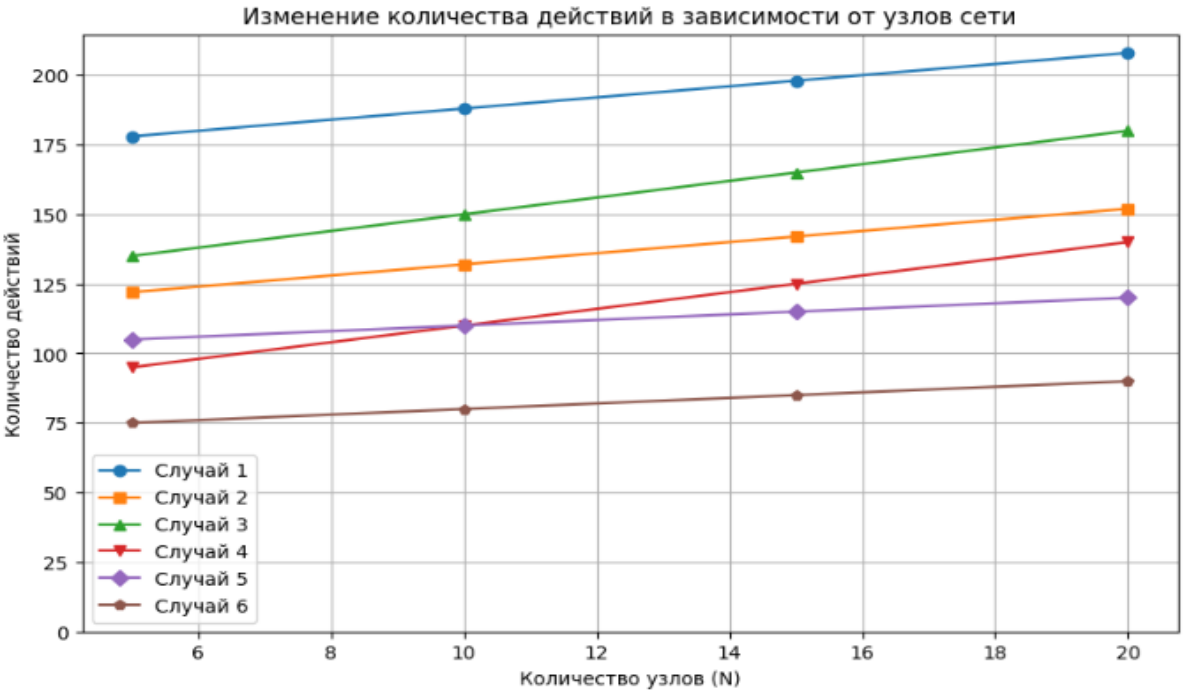

Рисунок 4 – Изменение количества действий в зависимости от количества узлов в инфраструктуре ИИ

Для моделей без возвратных действий степень полноты снижена (Случай 2 – подробная (полная) модель, Случай 4 – упрощенные модели, Случай 6 – типовая (общая) модель) по сравнению с моделями, в которых возвратные действия учтены (Случай 1 – подробная (полная) модель, Случай 3 – упрощенные модели, Случай 5 – типовая (общая) модель). Максимальный уровень охвата возможных действий и состояний демон-

стрирует подробная модель с использованием техник и тактик, учитывающая возвратные состояния (Случай 1), что актуально для повышения качества аудита в части повышения степени полноты описания атак на системы ИИ, включая ПИИ.

В некоторых аспектах предложенные модели на основе МППР превосходят модели на основе альтернативных методов. Сравнение моделей проводилось по полноте, точности, времени и адаптивности (таблица 2). Моделирование производилось с учетом следующего:

1. Выбор состояний производился из набора тактик MITRE ATLAS для тестовой системы ИИ. Используются следующие тактики: T1, T4, T5, T14.

2. Для каждой атаки при моделировании с помощью МППР строился граф состояний с учётом тактик MITRE ATLAS (например, T1 → T4 → T5 → T14).

3. Вероятности переходов между состояниями задавались на основе CVSS-оценок уязвимостей.

4. При моделировании последовательностей действий злоумышленника (при составлении моделей) использовались следующие методы: МППР (предложенные модели), деревья атак, байесовские сети, анализ с применением нейросетей.

Таблица 2 - Сравнение результатов моделирования

Сравниваемые модели	Полнота (%)	Точность (%)	Время выполнения (с.)	Адаптивность (1-5)
Разработанные модели (МППР)	90	95	120	5
Модели на основе деревьев атак	70	80	60	2
Модели на основе нейронной сети	75	85	300	3
Модели на основе байесовской сети	80	88	200	4

Разработанные модели позволяют формировать и анализировать все возможные пути атак, включая многошаговые атаки, редкие и неочевидные комбинации тактик. При добавлении новых уязвимостей или тактик, модель МППР требует лишь обновления матриц переходов и вознаграждений, а не полной перестройки. Для изучения аспектов моделирования атак на ПИИ использовалась простая нейронная сеть (семь слоев, из которых пять скрытых полносвязанных слоев) и сборка генеративных состязательных сетей (GAN).

Архитектура программного решения, которое представлено системой построения последовательности атакующих воздействий на ПИИ на основе разработанных моделей МППР, включает в свой состав следующие элементы: модуль сбора данных о параметрах атак (клиентская часть); модуль анализа данных и сохранения результатов анализа (серверная часть); анализатор событий или других параметров; консультаци-

онная подсистема, использующая МППР. Специфика последовательности работы программного решения в части реализации разработанных моделей отражена на рисунке 5.

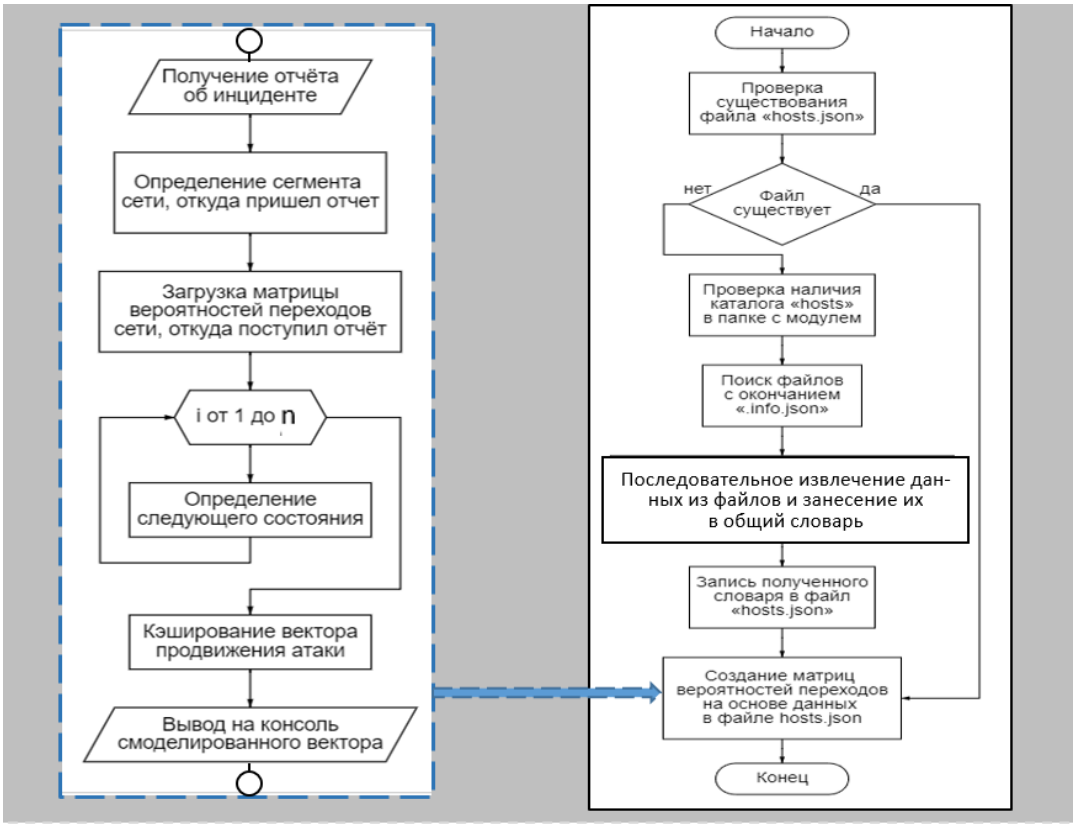

Рисунок 5 - Специфика последовательности работы предложенной системы (в части реализации разработанных моделей)

В заключении приведены основные научно-практические результаты, полученные в ходе диссертационного исследования.

ОСНОВНЫЕ ВЫВОДЫ И РЕЗУЛЬТАТЫ РАБОТЫ

В рамках данной работы были представлены модели на основе МППР, методика и алгоритм, позволяющие формировать последовательности атакующих воздействий на системы ИИ (ПИИ). Эти последовательности позволяют построить более полные и точные сценарии атак, которые применяются при аудите ИБ систем ИИ. Основные результаты работы следующие:

1. Произведен анализ существующих подходов поиска и оценки событий безопасности в информационных системах с элементами ИИ. Определена специфика применения МППР для формирования последовательности атакующих воздействий на ПИИ (поиска наиболее опасных последовательностей), при этом учтены особенности функционирования ПИИ, их инфраструктурные особенности (интерфейсы доступа, их

программно-аппаратное обеспечение и связи между компонентами системы). Осуществлен выбор способов описания уязвимостей и методик описания атак.

2. Разработаны модели определения последовательностей и анализа атакующих воздействий на ПИИ. Также приведены типы моделей описания атак, которые позволяют повысить полноту описания возможных атакующих воздействий, выявить наиболее опасные из них.

3. Разработана методика, позволяющая применять модели МППР с учетом их детализации в процессе составления сценариев атак. Разработан и предложен алгоритм формирования последовательностей атакующих воздействий.

4. Разработана архитектура программного решения, в том числе программные компоненты системы построения последовательности атакующих воздействий на ПИИ на основе разработанных моделей МППР. Программное решение дает возможность автоматизировать процессы моделирования атак с учетом разработанной методики.

Эксперименты и теоретические оценки подтверждают работоспособность моделей при учете различных режимов моделирования. При этом прослеживается превосходство предложенных моделей МППР по критерию полноты описания последовательности атакующих воздействий над моделями альтернативных методов (с учетом уровня описания и режима моделирования). Следует отметить высокую степень адаптивности моделей (возможность манипулировать наборами состояний и действий). Также модели позволяют учитывать непредсказуемость поведения злоумышленника при недостаточной информации о мотивации нарушителя. Результаты работы позволяют повысить качество аудита защищенности систем ИИ посредством уточнения и полного описания сценариев атак на ПИИ, что в итоге может значительно повысить уровень защищенности ПИИ.

Полученные результаты работы соответствуют специальности 2.3.6. Методы и системы защиты информации, информационная безопасность.

СПИСОК РАБОТ, ОПУБЛИКОВАННЫХ ПО ТЕМЕ ДИССЕРТАЦИИ

Публикации в рецензируемых научных изданиях, рекомендованных ВАК

1. Подтопельный, В.В. Особенности формирования вектора современных сетевых атак / И.А. Ветров, В.В. Подтопельный // Вестник СибГУТИ. – 2022. – № 3 (59). – С. 3 - 13.

2. Подтопельный, В. В. Особенности построения нейросетей с учетом специфики их обучения для решения задач поиска сетевых атак / И.А. Ветров, В.В. Подтопельный // Доклады ТУСУР. – 2023. – Т. 26, № 2. – С. 42 - 50.

3. Подтопельный, В.В. Формирование вектора сетевых атак с учетом специфики связей техник и тактик / И.А. Ветров, В.В. Подтопельный // Вестник СибГУТИ. – 2023. – Т. 17, № 4. – С. 49 - 61.

4. Подтопельный, В. В. Особенности моделирования атак на модели машинного обучения с использованием марковских процессов принятия решений / В.В. Подтопельный // Доклады ТУСУР. – 2024. – Т. 27, № 2. – С. 21 - 30.

5. Подтопельный, В. В. Исследование специфики моделирования компьютерных атак с использованием марковских процессов принятия решений и q-обучения/ В.В. Подтопельный // Информация и безопасность. – 2024. – Т. 27, вып.3 – С. 421 - 440.

Публикации в изданиях, входящих в международные базы цитирования

6. Podtopelny, V. Reliability Assessment of Segments of the Digital Ecosystems / V. Podtopelny, A. Babaeva // Ecosystems Without Borders. EcoSystConfKlgtu 2021. Lecture Notes in Networks and Systems. – Springer, Cham, 2022.– Vol. 474.– pp. 261 - 269.

7. Podtopelny, V. The Specifics of Determining the Value of Segments of Digital Ecosystems / V. Podtopelny, A. Babaeva // Ecosystems Without Borders 2023. EcoSystConfKlgtu 2023. Lecture Notes in Networks and Systems. – Springer, Cham, 2023.– Vol. 705. – pp. 189 - 197.

Свидетельства о результатах интеллектуальной деятельности

8. Подтопельный, В.В. Модуль анализа параметров сетевых атак для СОВ / Д.В. Куделка, В.В. Подтопельный // Свидетельство о регистрации программы для ЭВМ № 2020664040, 06.11.2020. Заявка № 2020662457 от 20.10.2020.

9. Подтопельный, В.В. Программа для ЭВМ «NNSCA» / А.Ю. Майстренко, В.В. Подтопельный // Свидетельство о регистрации программы для ЭВМ 2021665446, 27.09.2021. Заявка № 2021664445 от 17.09.2021.

10. Подтопельный, В.В. Программа анализа сценариев атак на системы корпоративного типа при аудите / И.А. Ветров, В.В. Подтопельный // Свидетельство о регистрации программы для ЭВМ 2024689273, 05.12.2024. Заявка № 2024688698 от 22.11.2024.

11. Подтопельный, В.В. Программа экспертного анализа угроз информационной безопасности / И.А. Ветров, В.В. Подтопельный // Свидетельство о регистрации программы для ЭВМ 2024662827, 30.05.2024. Заявка № 2024619240 от 25.04.2024.

12. Подтопельный, В.В. Программа для ЭВМ «BotnetSniffer» / Н.А. Семенов, В.В. Подтопельный //Свидетельство о регистрации программы для ЭВМ RU 2024616878, 26.03.2024. Заявка № 2024614885 от 12.03.2024.

13. Подтопельный, В.В. Гибридная система поиска, анализа и прогнозирования событий безопасности в распределённой информационной системе / В.В. Подтопельный, Н.А. Семенов, А.А. Кожевникова, А.А. Подтереба // Свидетельство о регистрации программы для ЭВМ RU 2024665096, 27.06.2024. Заявка № 2024663319 от 13.06.2024.

14. Подтопельный, В.В. Программа анализа атак отравления на наборы данных / И.А. Ветров, А.И. Сацута, В.В. Подтопельный // Свидетельство о регистрации программы для ЭВМ 2025619455, 16.04.2025. Заявка № 2025617891 от 04.04.2025.

Публикации в других изданиях

15. Подтопельный В.В. Особенности адресной защиты в АСУ ТП /В. В. Подтопельный // В сборнике: Балтийский морской форум. Материалы VI Международного Балтийского морского форума, в 6 томах. – 2018. – С. 456 - 462.

16. Подтопельный В.В. Особенности сбора данных в многомодульной системе обнаружения вторжений / В. В. Подтопельный // В сборнике: Балтийский морской форум. Материалы VII Международного Балтийского морского форума: в 6 т.– 2019. – С. 332 - 336.

17. Подтопельный В.В. Особенности активного аудита информационной безопасности АСУ предприятия /В. В. Подтопельный // Научный аспект. – 2019. – Т.1, № 4.– С. 86 - 90.

18. Хватов Д.А., Ковтун А.И., Подтопельный В.В. Проблемы аудита информационной безопасности АСУ ТП / Д.А. Хватов, А.И. Ковтун, В. В. Подтопельный // Вестник Балтийского федерального университета им. И. Канта. Серия: Физико-математические и технические науки. – 2019. – № 4. – С. 67 - 75.

19. Подтопельный В.В. Особенности классификации сигнатур систем обнаружения вторжений в АСУТП /В. В. Подтопельный // Modern Science. – 2019. – № 12-1. – С. 605 - 608.

20. Подтопельный, В.В. Особенности формирования SIEM-правил в АСУТП / В.В. Подтопельный // Научный аспект. – 2020. – Т. 4. – № 4. – С. 480 - 484.

21. Подтопельный, В.В. Сравнительный анализ технологий аудита информационной безопасности сетевой инфраструктуры диспетчерского уровня АСУТП / В.В. Подтопельный // В сборнике: VIII Балтийский морской форум. Материалы Международного балтийского морского форума: в 6-ти томах. Калининград, 2020. – С. 306 - 311.

22. Подтопельный, В.В. Особенности формирования сигнатурных последовательностей для обнаружения сетевых атак в АСУТП / В.В. Подтопельный // Modern Science. – 2020. – № 12-3. – С. 303 - 307.

23. Подтопельный, В.В. Особенности подготовки активного аудита информационной безопасности АСУТП / И.А. Ветров, В.В. Подтопельный // Вестник Балтийского федерального университета им. И. Канта. Серия: Физико-математические и технические науки. – 2021. – № 1. – С. 5 - 11.

24. Подтопельный, В.В. Определение пригодности правил обнаружения сетевых вторжений и их математическая оценка / И.А. Ветров, В.В. Подтопельный // Вестник Балтийского федерального университета им. И. Канта. Серия: Физико-математические и технические науки. – 2021. – № 1. – С. 11 - 18.