ЗАКЛЮЧЕНИЕ 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» МИНИСТЕРСТВА НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ, ФЕДЕРАЛЬНОГО ГОСУДАРСТВЕННОГО АВТОНОМНОГО ОБРАЗОВАТЕЛЬНОГО УЧРЕЖДЕНИЯ ВЫСШЕГО ОБРАЗОВАНИЯ «САНКТ-ПЕТЕРБУРГСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ АЭРОКОСМИЧЕСКОГО ПРИБОРОСТРОЕНИЯ» МИНИСТЕРСТВА НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ, ФЕДЕРАЛЬНОГО ГОСУДАРСТВЕННОГО БЮДЖЕТНОГО ОБРАЗОВАТЕЛЬНОГО УЧРЕЖДЕНИЯ ВЫСШЕГО ОБРАЗОВАНИЯ «САНКТ-ПЕТЕРБУРГСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ ТЕЛЕКОММУНИКАЦИЙ ИМ. ПРОФ. М.А. БОНЧ-БРУЕВИЧА» МИНИСТЕРСТВА ЦИФРОВОГО РАЗВИТИЯ, СВЯЗИ И МАССОВЫХ КОММУНИКАЦИЙ РОССИЙСКОЙ ФЕДЕРАЦИИ, ПО ДИССЕРТАЦИИ НА СОИСКАНИЕ УЧЕНОЙ СТЕПЕНИ КАНДИДАТА ТЕХНИЧЕСКИХ НАУК

аттестационное дело № _____

решение диссертационного совета от 10 декабря 2025 г. № 24

О присуждении Подтопельному Владиславу Владимировичу, гражданину Российской Федерации, ученой степени кандидата технических наук.

Диссертация «Модели и методика определения последовательностей атакующих воздействий на системы искусственного интеллекта при аудите информационной безопасности» по специальности 2.3.6. Методы и системы защиты информации, информационная безопасность принята к защите 08 октября 2025 года, протокол № 22 объединенным диссертационным советом 99.2.038.03, созданным на базе Федерального государственного бюджетного образовательного учреждения высшего образования «Балтийский государственный технический университет «ВОЕНМЕХ» им. Д.Ф. Устинова» Министерства науки и высшего образования Российской Федерации, Федерального государственного автономного образовательного учреждения высшего образования «Санкт-Петербургский государственный университет аэрокосмического приборостроения» Министерства науки и высшего образования Российской

Федерации, Федерального государственного бюджетного образовательного учреждения высшего образования «Санкт-Петербургский государственный университет телекоммуникаций им. проф. М.А. Бонч-Бруевича» Министерства цифрового развития, связи и массовых коммуникаций Российской Федерации, 191186, Санкт-Петербург, наб. реки Мойки, д. 61, приказ № 44/нк от 30 января 2017 года.

Соискатель Подтопельный Владислав Владимирович, 25 декабря 1978 года рождения, работает старшим преподавателем на кафедре информационной безопасности в Федеральном государственном бюджетном образовательном учреждении высшего образования «Калининградский государственный технический университет», Федеральное агентство по рыболовству.

В 2015 году соискатель окончил Федеральное государственное бюджетное образовательное учреждение высшего образования «Калининградский государственный технический университет» с присвоением квалификации инженер. В 2015 году окончил освоение программы подготовки научных и научно-педагогических кадров в аспирантуре Федерального государственного бюджетного образовательного учреждения высшего образования «Калининградский государственный технический университет».

Диссертация выполнена на кафедре информационной безопасности Федерального государственного бюджетного образовательного учреждения высшего образования «Калининградский государственный технический университет».

Научный руководитель – кандидат технических наук, Ветров Игорь Анатольевич, основное место работы: Федеральное государственное автономное образовательное учреждение высшего образования «Балтийский федеральный университет имени Иммануила Канта», Образовательно-научный кластер «Институт высоких технологий», методический руководитель по укрупнённой группе специальностей и направлений подготовки 10.00.00 «Информационная безопасность».

Оппоненты: 1. Рытов Михаил Юрьевич, доктор технических наук, доцент, основное место работы: Федеральное государственное образовательное учреждение высшего образования «Брянский государственный технический университет», кафедра систем информационной безопасности, заведующий кафедрой ; 2. Браницкий Александр Александрович, кандидат технических наук, доцент, основное место работы: акционерное общество «Клаудран», программист, дали положительные отзывы о диссертации.

Ведущая организация Федеральное государственное бюджетное учреждение науки «Санкт-Петербургский Федеральный исследовательский центр Российской академии наук», г. Санкт-Петербург, в своем положительном заключении, подписанном Саенко Игорем Борисовичем, доктором технических наук, профессором, главным научным сотрудником лаборатории Проблем компьютерной безопасности; Новиковой Евгенией Сергеевной, кандидатом технических наук, доцентом, старшим научным сотрудником лаборатории Проблем компьютерной безопасности; утвержденным Ронжиным Андреем Леонидовичем, доктором технических наук, профессором, профессором РАН, директором Федерального государственного бюджетного учреждения науки «Санкт-Петербургский Федеральный исследовательский центр Российской академии наук», указала, что диссертация является законченной научно-квалификационной работы, в которой на основании выполненных автором исследований разработаны модели, методика и алгоритм определения последовательностей атакующих воздействий на подсистемы искусственного интеллекта, а также архитектура программного решения для автоматизации моделирования атак. Представленные автором исследования результаты обеспечивают повышение полноты и обоснованности сценариев атак, применяемых при аудите информационной безопасности систем искусственного интеллекта (ИИ), и могут быть использованы в практических задачах анализа защищенности. Диссертация соответствует требованиям п. 9-14 Положения о порядке присуждения ученых степеней, утвержденного Постановлением Правительства Российской Федерации № 842 от 24.09.2013 г. (с последующими

изменениями), а ее автор – Подтопельный Владислав Владимирович заслуживает присуждения ученой степени кандидата технических наук по специальности 2.3.6. Методы и системы защиты информации, информационная безопасность.

Соискатель имеет 58 опубликованных работ, в том числе по теме диссертации 28, из них в рецензируемых научных изданиях, рекомендованных ВАК, – 5, в том числе 5 в изданиях, соответствующих искомой специальности и отрасли науки, а также: 2 работы в изданиях, индексируемых в международных базах цитирования; 7 результатов интеллектуальной деятельности; 21 статью в других научных журналах, сборниках научных статей, трудов и материалах конференций. Из 28 работ 14 работ опубликованы соискателем без соавторства. Общий объём авторского вклада в работы (без результатов интеллектуальной собственности) составляет 8,3 печ.л. из общего количества 11,8 печ.л. Диссертация не содержит недостоверных сведений об опубликованных соискателем ученой степени работах.

Наиболее значительные научные работы 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, индексируемых в МБЦ:

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.

Результаты 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.

На диссертацию и автореферат поступили отзывы: официального оппонента Рытова М.Ю., д.т.н., доц., заведующего кафедрой систем информационной безопасности Федерального государственного бюджетного образовательного учреждения высшего образования «Брянский государственный технический университет»; официального оппонента Браницкого А.А., к.т.н., доц., программиста акционерного общества «Клаудран»; ведущей организации СПб ФИЦ РАН; Шевцовой Г.А., к.и.н., доц., директора Института информационных наук и технологий безопасности, заведующей кафедрой Информационной безопасности Федерального государственного автономного образовательного учреждения высшего образования «Российский государственный гуманитарный университет»; Тебуевой Ф.Б., д.ф.-м.н., доц., профессора кафедры вычислительной математики и кибернетики, доцента Федерального государственного автономного образовательного учреждения высшего образования «Северо-Кавказский федеральный университет»; Серпенинова О.В.,

к.т.н., доц., профессора кафедры информационной безопасности Федерального государственного бюджетного образовательного учреждения высшего образования «Ростовский государственный экономический университет (РИНХ)»; Конева А.А., к.т.н., доц., старшего научного сотрудника Института системной интеграции и безопасности Федерального государственного автономного образовательного учреждения высшего образования «Томский государственный университет систем управления и радиоэлектроники»; Щербова И.Л., к.т.н., проректора Федерального государственного бюджетного образовательного учреждения высшего образования «Донецкий национальный технический университет»; Петрова В.И., к.т.н., доц., заведующего кафедрой основ радиотехники и защиты информации Федерального государственного бюджетного образовательного учреждения высшего образования «Московский государственный технический университет гражданской авиации»; Новикова С.Н., д.т.н., доц., и.о. заведующего кафедрой безопасности информации и технологий ГБОУ ВПО «Сибирский государственный университет телекоммуникаций и информатики»; Супруна А.Ф., к.т.н., доц., доцента Института компьютерных наук и кибербезопасности Федерального государственного автономного образовательного учреждения высшего образования «Санкт-Петербургский политехнический университет Петра Великого»; Шабуровой А.В., д.э.н., доц., директора Института оптики и технологий информационной безопасности Федерального государственного образовательного учреждения высшего образования «Сибирского государственного университета геосистем и технологий»; Кущева С.С., к.т.н., начальника 55 кафедры Автоматизированных систем управления (и информационной безопасности) и Будникова С.А., д.т.н., проф., профессора 55 кафедры автоматизированных систем управления (и информационной безопасности) Федерального государственного казенного военного образовательного учреждения высшего образования «Военный учебно-научный центр Военно-воздушных сил «Военно-воздушная академия имени профессора Н.Е. Жуковского и Ю.А. Гагарина»; Милославской Н.Г. д.т.н., доц., профессора

кафедры информационной безопасности банковских систем Института интеллектуальных кибернетических систем Федерального государственного автономного образовательного учреждения высшего образования «Национальный исследовательский ядерный университет «МИФИ»; Сабанова А.Г., д.т.н., доц., главного эксперта научно-технического комплекса технологий информационного общества АО «Научно-исследовательский и проектно-конструкторский институт информатизации, автоматизации и связи на железнодорожном транспорте»; Осипова М.Н., к.ф.-м.н., доц., заведующего кафедрой безопасности информационных систем Федерального государственного автономного образовательного учреждения высшего образования «Самарский национальный исследовательский университет имени академика С.П. Королева»; Мазина А.В., д.т.н., проф., заведующего кафедрой защиты информации Калужского филиала Федерального государственного бюджетного образовательного учреждения высшего образования «Московский государственный технический университет имени Н.Э. Баумана»; Шакурского М.В., д.т.н., доц., заведующего кафедрой информационной безопасности Федерального государственного бюджетного образовательного учреждения высшего образования «Поволжский государственный университет телекоммуникаций и информатики».

Все отзывы положительные, но имеются замечания. Доказательства ряда утверждений излагаются излишне подробно, что увеличивает объем диссертационной работы и, соответственно, затрудняет чтение. В пункте 2.4 модель нарушителя чрезвычайно сжата, что требует для понимания специфики формирования и влияния модели нарушителя на процесс моделирования атак. Специфика моделирования внутреннего нарушителя в главе 3 учитывается, однако не выделяется в отдельное описание. Стиль изложения материала диссертационного исследования сложен для понимания сути работы. Во второй главе описываются виды доступа, которые используются при определении параметров моделей второго типа в третьей главе. При этом для моделей третьего типа упоминается только возможность их применения. Для моделей второго типа указывается применимость к описанию атак на инфраструктуру объекта,

подробно не объяснено, почему приоритет отдается анализу и построению атак на логические компоненты систем ИИ. Подробно не описывается, в каких случаях допустимо определять искомые последовательности без средств автоматизации. В ряде фрагментов главы, посвященной подробному построению моделей и алгоритмов (разделы 3.2, 3.3, 3.4), материал изложен чрезмерно детально. Было бы целесообразно предварять их более компактным обобщенным описанием модели, что позволило бы уменьшить громоздкость последующего изложения. Хотя в диссертации активно используется база знаний MITRE ATLAS, специфика атак на вычислительные модели ИИ представлена менее полно, чем инфраструктурные и сетевые аспекты. Проблематика атак на данные, модели и механизмы обучения освещена преимущественно через тактики ATLAS с использованием стандарта для оценки степени опасности уязвимостей CVSS, однако глубокая связь между внутренними уязвимостями моделей ИИ и построением марковских стратегий остается раскрытой не в полной мере. Включение в качестве приложений исходных кодов разработанных программных компонентов повысило бы уровень воспроизводимости результатов и упростило бы практическое использование предложенного программного решения. Работа обоснованно опирается на базу MITRE ATLAS и частично на методические документы ФСТЭК, однако современные международные источники по угрозам для ИИ освещены слабо. Это несколько сужает контекст исследования, особенно с учетом активного развития международных стандартов и таксономий угроз в области искусственного интеллекта. Оформление отдельных элементов диссертационной работы не в полной мере соответствует ГОСТ Р 7.0.11–2011, а оформление библиографических записей не полностью соответствует ГОСТ 7.1 и ГОСТ 7.80. Из автореферата не вполне ясно, почему связь классов атак на системы ИИ с предполагаемыми моделями основывается на степени известности моделей ИИ. В автореферате автор указывает, но подробно не разъясняет, почему тактики и техники «Методики оценки угроз безопасности информации ФСТЭК России» (приложение 11) считаются дополнительными, а не основными, как в случае MITRE ATLAS. По тексту автореферата недостаточно ясно, каким образом

проверяется на значимость каждое действие с учётом его стоимости (стр. 17), а также порядок определения адаптивности сравниваемых моделей (стр. 21, табл. 2). Внимание акцептируется на атаках, производимых злоумышленником на модели ИИ, в то время как атаки на вычислительную инфраструктуру ИИ рассматриваются опосредованно. В обозначении некоторых величин отсутствует единообразие, в частности, используется одинаковое обозначение «R» для различных сущностей – действий злоумышленника по возврату в пройденные состояния (курсивом) и функции вознаграждения (без курсива), что вносит определённую путаницу, т.к. в табл. 1 автореферата в первом столбце функция вознаграждения (R), вероятность переходов (P) и множество действий (A) вводятся без курсива, а дальше по тексту A и P используются с курсивом, а R без курсива, кроме второго столбца табл. 1. Из материала автореферата недостаточно ясно, на каком этапе аудита следует применять предложенную методику. Из автореферата не ясно, каким образом в моделях учитывается специфика распространенных в России подходов и методик формирования сценариев атак. Положения, выносимые на защиту диссертации, должны быть привязаны к пунктам специальности 2.3.6, что чётко показывает результаты автора в представленном диссертационном исследовании. В работе можно отметить некоторую ограниченность исследования, связанную с использованием MITRE ATLAS, как основы построения моделей. В пунктах научной новизны отсутствуют количественные показатели предлагаемых методик, алгоритмов. Отмечаются только качественные параметры. В автореферате недостаточно подробно рассмотрены особенности MITRE ATLAS, важные для применения в моделях на основе марковских процессов принятия решений. Из материала автореферата недостаточно ясно, как влияет предложенная функция вознаграждения на модели с учетом специфики применяемого режима моделирования. Из текста автореферата не очевидно достижение цели работы, так как не приведен критерий выбора моделей (максимум полноты) и количественные значения повышения степени полноты описания атак на системы ИИ при аудите их информационной безопасности. Утверждения о «превосходстве»

предложенных моделей по критерию полноты требуют более строгой аргументации. Кроме того, судя по содержанию таблицы 2, должна была проводиться многокритериальная оценка с привлечением дополнительных показателей «Точность», «Время выполнения» и «Адаптивность». Недостаточная конкретизация научной новизны. Пункты научной новизны (стр. 6-7) сформулированы как декларации о намерениях или возможностях предлагаемого аппарата («появится возможность...», «позволяет выявлять...»), а не как конкретные, уже достигнутые научные результаты. Требуется более четкое формулирование, в чем именно заключается новизна разработанных моделей, методики и алгоритма по сравнению с известными аналогами. В автореферате в явном виде не представлена формализованная постановка задачи исследования и структурно-логическая схема исследования; что затрудняет понимание роли и места разработанных моделей и методик при аудите информационной безопасности. Из материала автореферата недостаточно ясно, какова специфика прототипов, с которыми производится сравнение разработанных автором моделей. В разделе «новизна» для разработанных моделей и методики нет чёткого изложения аспектов новизны результатов исследования. В частности, в каждом пункте перечисления явно не хватает общепринятого оборота «отличающийся от известных...». В разделе «соответствие требованиям ВАК» отсутствует указание областей исследования специальности 2.3.6, которым соответствуют выдвинутые на защиту положения. Автор диссертации значительное внимание уделяет базе знаний MITRE ATLAS, в то время как другие базы и методики, применимые к области ИБ ИИ, рассматриваются как вспомогательные. В исследовании не описывается подробно специфика применения моделей в сопряжении с другими подходами, которые используются сегодня в современных системах аудита и методиках аудитов; недостаточно проанализирована таблица 2 и последовательность работы программного решения – рисунок 5, что связано, скорее всего, с ограниченным объёмом автореферата. Описание применения моделей с учетом методики аудита рассматривается либо в контексте MITRE ATLAS, либо в её альтернативах. В ATLAS представлено 56 техник, а в ATT&CK

196, это может не охватывать все возможные угрозы для сложной ИИ-системы. Угрозы в области ИИ быстро эволюционируют, ATLAS требует регулярного обновления. Задержки в добавлении новых техник могут снижать его актуальность. В материале автореферата недостаточно подробно свойства MITRE ATLAS, которые важны для формирования предлагаемых моделей.

Выбор оппонентов и ведущей организации обосновывается известностью среди научных специалистов в соответствующей области науки. Определяющие критерии включают наличие научных публикаций в данной и смежных областях, а также способность к квалифицированной оценке актуальности, теоретической значимости и практической ценности представляемой диссертации. Рытов М.Ю., д.т.н., доц., возглавляет подразделение, учебная и научная деятельность которого сосредоточена в области защиты информации, исследовании специфики атакующих воздействий на компьютерные системы. Рытов М.Ю. является одним из ведущих специалистов в исследовании проблем использования ИИ-систем при обеспечении информационной безопасности автоматизированных систем; Браницкий А.А., к.т.н., доц., ведет активную научную деятельность в области использования ИИ-систем при решении задач обеспечения информационной безопасности, а также в области распознавания сетевых атак, в том числе направленных на компрометацию ИИ-систем, реализуя накопленный научный опыт в рабочих процессах своей профессиональной деятельности. Ведущая организация Санкт-Петербургский Федеральный исследовательский центр Российской академии наук является одним из основных научных центров исследования проблем защиты информации. Отзыв был подготовлен сотрудниками, научная деятельность которых заключается в изучении аспектов реализации компьютерных атак, поиске способов их предотвращения. Также сотрудниками лаборатории выполнены множество грантов и государственных контрактов, направленных на повышение уровня защищенности современной информационной среды от действий злоумышленников (известные ученые: Котенко И.В., Саенко И.Б., Новикова Е.С., Чечулин А.А. и др.).

Диссертационный совет отмечает, что на основании выполненных соискателем исследований: разработаны модели, алгоритм, методика определения последовательностей атакующих воздействий на системы искусственного интеллекта, которые, в отличие от известных, учитывают специфику устройства и организации работы ИИ-систем, превосходят таковые по критерию полноты, а также позволяют выявить наиболее опасные последовательности, где в качестве показателя опасности состояния последовательности выступает значение функции полезности, что в итоге повышает степень полноты описания сценариев атак, а также положительно влияет на качество аудита информационной безопасности ИИ-систем, в том числе, когда они выступают в качестве подсистем других информационных систем;

предложены модели определения последовательности атакующих воздействий, которые, в отличие от известных, учитывают инфраструктурную организацию, режимы работы ИИ-систем, специфику их уязвимостей, уровни доступности и контролируемости атакующих воздействий, стохастический характер действий злоумышленника, что отражено в вводимых режимах моделирования, этапах атак, при этом модели предусматривают интеграцию различных метрических систем и баз знаний, используемых при описании атак, угроз и уязвимостей; алгоритм, который, в отличие уже ранее предложенных, предусматривает введение с учетом специфики моделирования отдельной дополнительной аналитической части, позволяющей изучать динамику состояний в зависимости от действий злоумышленника; методика, в отличие от уже используемых, включает в свой состав уровни детализации последовательностей атакующих воздействий с последующим формированием сценария атак для использования в процессах аудита информационной безопасности ИИ-систем с использованием актуальных баз знаний и методологий в области информационной безопасности искусственного интеллекта (MITRE ATLAS, CVSS и др.), что с учетом условий моделирования дает возможность довести до 90% степень полноты описания атак на ИИ-системы;

доказана перспективность использования разработанных моделей, алгоритма, методики определения последовательностей атакующих воздействий на ИИ-системы для повышения качества аудита по показателю полноты;

введены изменения в порядок проведения аудита информационной безопасности ИИ-систем, новая трактовка состояний и действий компрометации ИИ, используемых при моделировании атакующих воздействий.

Теоретическая значимость исследования обоснована тем, что: доказано повышение степени полноты описания атак на системы ИИ, определения наиболее опасной последовательности атакующих воздействий при использовании предлагаемых моделей. При этом они способны сохранять актуальность интерпретации состояний и действий с учетом использования современных баз знаний MITRE ATLAS;

применительно к проблематике диссертации результативно (эффективно, то есть с получением обладающих новизной результатов) использованы марковские процессы принятия решения, модели машинного обучения с применением аналитико-статистических методов;

изложены положения диссертационного исследования, раскрывающие специфику разработанных моделей, доказательства их работоспособности и эффективности, специфика алгоритма и методики их применения при решении задач аудита информационной безопасности ИИ, условия применения и порядок интеграции состояний моделей и методик описания атак на ИИ-системы;

раскрыты особенности формирования атакующих последовательностей на ИИ-системы с учетом различных режимов моделирования, специфики компонентного состава ИИ- систем;

изучены специфика определения действий злоумышленника при атаках на ИИ с учетом уровня доступности ресурсов атакуемых систем, метрических оценок уязвимостей CVSS, особенностей действий злоумышленника;

проведена модернизация существующих математических моделей на основе марковских процессов принятия решений в части их применимости к

задачам аудита информационной безопасности ИИ-систем, обеспечивающих получение новых результатов по теме диссертации.

Значение полученных соискателем результатов исследования для практики подтверждается тем, что: разработаны процессы проведения аудита информационной безопасности, методика, модели и алгоритм их использования **внедрены** в рабочие процессы предприятий, ведущих деятельность в области обеспечения защиты информации: ООО «Радиоэлектронные системы» (г. Екатеринбург), ООО «Центр защиты информации» (г. Калининград);

определены пределы и перспективы практического использования разработанных моделей, алгоритмов, методик в процессах аудита информационной безопасности систем ИИ;

созданы алгоритм, модели и методика определения последовательностей атакующих воздействий на системы ИИ;

представлены рекомендации по дальнейшему совершенствованию разработанных моделей и методики.

Оценка достоверности результатов исследования выявила: для экспериментальных работ результаты получены в разработанной программной среде, которая включает в себя комплекс обученных моделей машинного обучения, в том числе нейронные сети;

теория построена на основе марковских процессов принятия решений, аналитико-статистических методах для теоретических работ;

идея базируется на анализе практики и обобщении передового опыта в области построения атакующих последовательностей на информационные системы и результатах исследований в области компрометации моделей и систем искусственного интеллекта;

использованы сравнение авторских данных и данных, полученных ранее по рассматриваемой тематике;

установлено, что результаты применения разработанных моделей в соответствии с предложенными алгоритмом и методикой показывают высокую степень полноты описания атакующих воздействий на системы ИИ;

использованы современные методики сбора и обработки исходной информации, современные базы знаний в области компрометации ИИ-систем, модели машинного обучения .

Личный вклад соискателя состоит в разработке моделей определения последовательности атакующих воздействий на системы искусственного интеллекта, в том числе, когда они выступают в качестве подсистем других информационных систем, алгоритма и методики формирования последовательности, в разработке программного обеспечения, реализующего модели поиска атакующих воздействий с использованием предложенной методики и алгоритма, апробации результатов исследования, разработке лабораторной среды испытания моделей.

В ходе защиты диссертации были высказаны следующие критические замечания о том, что необходима более точная формулировка научной новизны выносимых на рассмотрение положений и решаемых задач, а также требуется более точное и корректное отображение параметров точности и адаптивности сравниваемых моделей определения атакующих последовательностей.

Соискатель Подтопельный В.В. в ходе заседания **ответил** на задаваемые ему вопросы, **согласился с замечаниями** и привел собственную аргументацию.

Диссертационный совет установил, что диссертация «Модели и методика определения последовательностей атакующих воздействий на системы искусственного интеллекта при аудите информационной безопасности» является законченной научно-квалификационной работой и соответствует требованиям п. 9 Положения о присуждении ученых степеней, предъявляемым к кандидатским диссертациям, а также пунктам 3, 16 паспорта научной специальности 2.3.6. Методы и системы защиты информации, информационная безопасность.

На заседании 10 декабря 2025 года объединенный диссертационный совет принял решение присудить Подтопельному В.В. ученую степень кандидата технических наук за решение научной задачи по разработке на основе методов марковских процессов принятия решений моделей, алгоритма и методики определения, построения и анализа последовательностей атакующих воздействий

на системы ИИ с учетом современной специфики формирования сценариев атак и рекомендаций в области ИБ при аудите информационных систем.

При проведении тайного голосования объединенный диссертационный совет в количестве 17 человек, из них 5 докторов наук по научной специальности рассматриваемой диссертации, участвовавших в заседании, из 23 человек, входящих в состав совета, проголосовали: за – 17, против – нет, недействительных бюллетеней – нет.

Председатель диссертационного совета,
доктор технических наук, профессор Киричек Руслан Валентинович

Ученый секретарь диссертационного совета,
кандидат технических наук, доцент Владыко Андрей Геннадьевич

12 декабря 2025 года