ОТЗЫВ НА АВТОРЕФЕРАТ ДИССЕРТАЦИИ
Подтопельного Владислава Владимировича
«Модели и методика определения последовательностей атакующих воздействий на системы искусственного интеллекта при аудите информационной безопасности», представленную на соискание учёной степени кандидата технических наук по научной специальности 2.3.6. – Методы и системы защиты информации, информационная безопасность

В эпоху цифровой трансформации внедрение искусственного интеллекта (ИИ) в корпоративные системы становится ключевым элементом модернизации бизнес-процессов. Этот шаг порождает новые вызовы в сфере информационной безопасности, требуя пересмотра традиционных методов аудита и мониторинга информационных систем. Быстрое развитие технологий ИИ подчеркивает необходимость разработки новых подходов для оценки потенциальных рисков и уязвимостей, связанных с использованием интеллектуальных систем.

В условиях обеспечения кибербезопасности становится критическим не только выявление существующих угроз, но и прогнозирование возможных направлений атак, а также адаптация методов аудита к новым реалиям, вызванным интеграцией ИИ в современные информационные системы. Представленная диссертация В.В. Подтопельного посвящена решению актуальных задач в области аудита информационных систем с учётом специфики интеграции в них подсистем ИИ (ПИИ). Цель исследования – определить последовательности действий при предполагаемых атаках на системы ИИ, чтобы повысить качество аудита их безопасности. Для этого были разработаны модели и методика углублённого изучения сопряжения действий, направленных на эксплуатацию уязвимостей.

Диссертация включает следующие части:

1. Введение, в котором определены объект и предмет исследования, цели, актуальность, задачи и методы, положения на защиту, краткое содержание по главам.

2. Глава 1 включает анализ моделирования атакующих воздействий с использованием марковских моделей принятия решений (МППР). Рассмотрены особенности аудита систем ИИ с учетом MITRE ATLAS и методик ФСТЭК России. Приведена сравнительная таблица методов.

3. Глава 2 содержит описание специфики применения МППР для атак на ПИИ. Определены функция ценности, вознаграждения, математические основы графа атаки. Выделены два режима моделирования: on-line (с возвратом в состояния) и off-line.

4. Глава 3 включает модели формирования последовательностей атак на основе МППР. Использован метод Value Iteration.

Разработаны три типа моделей:

A. Типовая модель для общего анализа атак.

B. Упрощенные модели с тактиками MITRE ATLAS.

C. Подробные модели без ограничений переходов.

Для каждого типа описаны ограничения и рекомендации. Формально описаны сценарии атак, включая вычислительные средства.

5. Глава 4 содержит экспериментальную оценку моделей.

6. Заключение содержит основные результаты и выводы.

Автореферат демонстрирует высокий уровень структурированности и логическую последовательность работы, включая все необходимые разделы, а также чётко сформулированные цели, задачи и результаты исследования. Ключевые научные достижения диссертации включают:

1. Разработку и обоснование комплекса моделей для выявления критичных последовательностей действий злоумышленников при атаках на системы ИИ. Следует отметить, что комплекс моделей подразумевает распределение моделей по степени детализации, при этом используются современные метрики уязвимостей, базы знаний об атакующих воздействиях на системы ИИ злоумышленников.

2. Создание методики использования разработанных моделей при аудите информационной безопасности инфраструктуры предприятий и организаций, что позволяет оптимизировать процесс оценки рисков и повышения уровня защищённости.

3. Разработку алгоритма применения разработанных моделей.

Однако, 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 с MITRE ATLAS.

Указанное замечание не снижает общей оценки диссертационного исследования. В заключение, автореферат позволяет сделать вывод, что

диссертация **Владислава Владимировича Подтопельного** представляет собой качественную научно-квалификационную работу, сочетающую актуальность темы, новизну подхода и практическую значимость. Работа удовлетворяет требованиям ВАК и соответствуют паспорту научной специальности 2.3.6 – Методы и системы защиты информации, информационная безопасность. Диссертационная работа соответствует всем требованиям п. 9-14 «Положения о присуждении ученых степеней», утвержденным постановлением Правительства Российской Федерации от 24.09.2013 г. № 842. Автор Владислав Владимирович Подтопельный заслуживает присуждения ученой степени кандидата технических наук, что подтверждает его высокий уровень компетентности и фундаментальное знание в области информационной безопасности.

Профессор кафедры вычислительной математики и кибернетики
Северо-Кавказского федерального университета
доктор физико-математических наук, доцент
Тебуева Фариза Биляловна

«06» 11 2025 г.

Адрес организации: 355017, г. Ставрополь, ул. Пушкина, д. 1, ФГАОУ ВО «Северо-Кавказский федеральный университет». Тел.: (865) 295-68-08
Адрес электронной почты: ftebueva@ncfu.ru

Подпись Ф.Б. Тебуевой заверяю:

ПОДПИСЬ УДОСТОВЕРЯЮ:
начальник отдела по работе с сотрудниками УКА
Д С Горбачева