DOI 10.37882/2223-2966.2025.10.45

РАЗРАБОТКА МОДУЛЬНОЙ МУЛЬТИАГЕНТНОЙ АРХИТЕКТУРЫ СИСТЕМЫ ТЕСТИРОВАНИЯ ЗАЩИЩЁННОСТИ ИНФОРМАЦИОННЫХ СИСТЕМ

DEVELOPMENT OF A MODULAR MULTI-AGENT ARCHITECTURE OF THE INFORMATION SYSTEMS SECURITY TESTING SYSTEM

R. Chikaev

Summary. The article emphasizes the importance of penetration testing in the light of a growing number of cyber threats and rapid technological advancements. The use of artificial intelligence (AI) in pentesting is relevant due to the complexity of attacks, technology evolution, significant financial risks of data leaks and compliance with regulatory standards. Integration of artificial intelligence into automation testing of security of information systems significantly improves the efficiency and accuracy of the testing, provides adaptation to new challenges, and a competitive advantage in the cybersecurity market by quickly identifying and responding to emerging threats. The article substantiates the concept of distributing roles between specialized agents, each of which implements a specific stage of the attack scenario — from reconnaissance and vulnerability analysis to exploitation and consolidation in the system. A model for typifying agent actions is proposed based on the MITRE ATT&CK guide to classifying and describing cyberattacks and intrusions, which allows formalizing the behavior of system components and increasing the accuracy of results verification. The developed system demonstrates stability, scalability, and practical applicability in complex secure information systems.

Keywords: modular multi-agent system, automation of penetration testing, information security, intelligent agents, coordination of actions, MITRE ATT&CK, typing of techniques, agent architecture, efficiency evaluation, pentest, autonomy.

Чикаев Руслан Александрович
Российский экономический университет им. Г.В. Плеханова, г. Москва
Chikaev.r@edu.rea.ru

Аннотация. Статья подчеркивает важность тестирования на проникновения (пентестинга) в условиях роста киберугроз, стремительного развития технологий. Использование искусственного интеллекта в пентесте является актуальным из-за сложности атак, эволюции технологий, существенных финансовых рисков утечек данных и соблюдения нормативных стандартов. Взаимодействие и интеграция искусственного интеллекта в процессы автоматизации проведения тестирования защищенности информационных систем существенно позволяет повысить эффективность и точность тестирования, адаптируясь к новым вызовам и обеспечивая конкурентное преимущество на рынке кибербезопасности, благодаря оперативно выявлять и реагировать на возникшие угрозы. В статье обоснована концепция распределения ролей между специализированными агентами, каждый из которых реализует конкретный этап атакующего сценария — от разведки и анализа уязвимостей до эксплуатации и закрепления в системе. Предложена модель типизации действий агентов на основе руководства по классификации и описанию кибератак и вторжений — MITRE ATT&CK, позволяющая формализовать поведение компонентов системы и повысить точность верификации результатов. Разработанная система демонстрирует устойчивость, масштабируемость и практическую применимость в условиях сложных защищённых информационных системах.

Ключевые слова: модульная мультиагентная система, автоматизация тестирования на проникновение, информационная безопасность, интеллектуальные агенты, координация действий, MITRE ATT&CK, типизация техник, архитектура агентов, оценка эффективности, пентест, автономность.

Введение

Актуальность работы обусловлена необходимостью создания новой модели автоматизированного тестирования на проникновение, которая способна не просто механически выполнять запрограммированные действия, но и гибко адаптироваться к различным условиям, обучаться на предыдущем опыте и эффективно взаимодействовать с окружающей средой.

Объектом исследования является система тестирования защищенности информационных систем. Предмет исследования составляют процессы тестирования защищенности информационных систем.

Цель работы заключается в разработке общей модели мультиагентной системы для повышения эффективности процессов тестирования защищенности информационных систем. Для достижения поставленной цели в рамках исследования решается несколько вопросов: проведение анализа существующих моделей автоматизированного тестирования на проникновение информационных систем, определение архитектурных принципов построения мультиагентных систем, разработка общей концепции собственной системы, создание прототипов специализированных агентов и проведение сравнительного тестирования их эффективности. Гипотеза исследования основывается на предположении,

что построение мультиагентной системы позволит существенно повысить производительность тестирования на проникновение, осуществить расширение функционала в сравнении с существующими централизованными и монолитными решениями.

Научная новизна данного исследования заключается в формализации подхода к типизации сценариев атак с целью выявления и систематизации функциональных ролей и компетенций программных агентов, задействованных в процессе оценки защищенности информационных систем. Необходимо сформировать типовые агенты для реализации интегрированного сценария атак.

Практическая значимость работы проявляется в возможности прямого применения разработанных моделей и методик при построении комплексных систем оценки защищенности. Актуализация знаний агентов может основываться на открытых фреймворках, таких как руководство по классификации и описанию кибератак и вторжений *MITRE ATT&CK*, а также на аналитических материалах ряда коммерческих и научных организаций в области киберугроз.

Методологическую основу исследования составили: обзор актуальных научных и практических источников в сфере автоматизированного тестирования ИБ, методы функционального и архитектурного моделирования мультиагентных систем, практическая реализация и отладка прототипов AI-агентов, а также проведение эмпирических тестов в контролируемой среде. Для анализа эффективности решений применялся сравнительный метод с акцентом на измерение временны́х затрат и полноты анализа в реальных сценариях.

1. Анализ тестирования защищенности информационных систем

Стремительная эволюция ИТ-архитектуры и возникновение инновационных методов кибератак обуславливают критическую значимость систематической проверки защищенности цифровых ресурсов. Сценарное моделирование занимает лидирующую позицию среди методологий оценки безопасности. Данный подход базируется на применении специализированных программ испытаний, включающих последовательность действий по выявлению уязвимостей инфраструктуры.

Практическая реализация сценарного подхода выявляет существенные ограничения методологии. Фиксированный характер тестовых сценариев, основанных на известных векторах проникновения, затрудняет своевременное реагирование на возникновение новых угроз безопасности. Актуальность проблемы возрастает в условиях динамично развивающихся инфраструктур, характеризующихся регулярным обновлением компонентов.

Современные технологии предлагают инновационное решение существующих проблем в виде мультиагентных систем тестирования. Данные системы характеризуются способностью оперативной реконфигурации стратегии проверок при обнаружении изменений в инфраструктуре, осуществляя автоматическую модификацию механизмов тестирования на основе актуальных данных.

Функциональные возможности мультиагентных систем существенно превосходят потенциал классических методологий оценки защищенности информационных комплексов. Принципиальное отличие заключается в реализации механизма обработки информации в режиме реального времени, позволяющего оптимизировать векторы тестирования соответственно модификациям инфраструктуры, что недостижимо при использовании статических шаблонов.

Внедрение мультиагентного подхода обеспечивает качественно новый уровень гибкости процедур тестирования. Автоматизированная система осуществляет мгновенную адаптацию поведения агентов, основываясь на результатах мониторинга состояния инфраструктуры. Подобная архитектура гарантирует всестороннюю проверку защищенности с учетом динамической природы современных информационных систем, значительно повышая результативность пентестинга.

Ключевым преимуществом мультиагентных систем выступает встроенный механизм самообучения, обеспечивающий непрерывное совершенствование алгоритмов анализа сетевой инфраструктуры. Расширенная функциональность включает углубленное исследование межузлового взаимодействия, постоянный контроль сетевых потоков данных, автоматическую идентификацию нестандартного поведения системы. Применение передовых технологий минимизирует риск необнаружения критических уязвимостей, существенно повышая достоверность оценки уровня информационной безопасности в условиях эволюционирующего ландшафта киберугроз.

Мультиагентные системы (МАС) выступают инновационным направлением развития искусственного интеллекта и программной инженерии, базирующимся на принципах кооперации автономных интеллектуальных агентов. Понятие «агент» определяет программный модуль, обладающий высокой степенью самостоятельности, способностью анализировать трансформации окружающей среды и реализовывать целенаправленные действия в рамках установленных параметров функционирования [1–2].

Самостоятельность выступает фундаментальным принципом функционирования агентов. Программ-

ные модули реализуют процесс принятия решений независимо от внешнего управления, корректируя поведенческие паттерны на основе целевых установок и накопленных данных. Подобная автономность обеспечивает оперативное реагирование на модификации инфраструктуры без необходимости постоянного административного контроля.

Оперативное восприятие изменений окружающей среды характеризует реактивные свойства агентов. Способность мгновенной перестройки алгоритмов функционирования обеспечивает своевременное обнаружение возникающих угроз, многократно увеличивая продуктивность системы безопасности.

Коммуникативные возможности агентов реализуются через механизмы межагентного взаимодействия. Интеграция индивидуальных модулей в единую операционную систему достигается посредством координации действий и обмена информационными потоками. Встроенные механизмы адаптации и машинного обучения позволяют агентам совершенствовать рабочие алгоритмы в процессе эксплуатации. Непрерывная обработка поступающей информации стимулирует эволюцию поведенческих паттернов, повышая результативность системы. Способность извлекать уроки из предыдущего опыта и модифицировать стратегии соответственно трансформациям целевой среды гарантирует устойчивую эффективность комплекса при динамичном развитии инфраструктуры.

Сфера практического применения мультиагентных систем охватывает широкий спектр направлений: от координации производственных комплексов и оптимизации логистических цепочек до моделирования экономических процессов и автоматизации биржевых операций [5].

Особую ценность мультиагентный подход демонстрирует в области кибербезопасности, обеспечивая параллельную обработку масштабных информационных потоков, своевременную идентификацию угроз и мгновенное реагирование на инциденты безопасности, что критически важно в условиях эволюционирующего ландшафта киберугроз. Внедрение МАС в процедуры пентестинга значительно расширяет возможности тестирования благодаря специализации агентов на различных аспектах анализа безопасности.

Централизованная архитектура представляет собой наиболее доступный вариант реализации МАС. Координационные функции возлагаются на специализированный элемент — управляющий узел, осуществляющий распределение задач, мониторинг состояния агентов и стратегическое планирование. Применение централизованного подхода в системах безопасности требует внедрения комплексных мер защиты координационного центра при отсутствии жестких требований к отказоустойчивости. Децентрализованные системы исключают наличие единого центра управления. Автономные агенты реализуют независимые алгоритмы принятия решений, осуществляя прямой информационный обмен. Подобная организация усложняет процессы разработки, требуя решения задач целевой синхронизации и обеспечения согласованности действий в условиях распределенной информации. Однако архитектура демонстрирует максимальную устойчивость, сохраняя работоспособность при выходе отдельных компонентов. Данное свойство критически важно при тестировании защищенности информационных систем, когда нейтрализация отдельных агентов не прерывает процесс тестирования [3].

Интеллектуальный потенциал агентов выступает критическим фактором эффективности системы. Базовые реактивные агенты функционируют по принципу «стимул-реакция», реализуя предопределенные алгоритмы без учета контекстуальной информации. Подобные модули характеризуются простотой реализации и стабильностью работы, однако демонстрируют низкую эффективность в условиях неопределенности [22–24].

Архитектурная концепция разрабатываемой мультиагентной системы базируется на принципе модульности. Функциональная организация системы предполагает специализацию агентов по конкретным направлениям пентестинга. Центральным элементом выступает Управляющий агент, реализующий функции координации компонентов, консолидации информационных потоков, администрирования жизненного цикла и оптимизации распределения задач внутри мультиагентного комплекса. функционирует как интеллектуальный медиатор, обеспечивающий интеграцию агентских модулей с внешними интерфейсами и централизованное управление информационными потоками. Функциональность агента включает системы сетевого мониторинга, менеджмент криптографических ключей и администрирование централизованного хранилища данных, аккумулирующего результаты тестирования и метаданные о проведенных атаках. Фундаментальным аспектом архитектуры выступает надежная система сетевых коммуникаций, реализованная на базе высокопроизводительной библиотеки *ZeroMQ* с интеграцией криптографического модуля *CurveZMQ*. Выбор технологического стека обусловлен требованиями к безопасности информационного обмена, минимизации латентности и обеспечению устойчивости к деструктивным воздействиям. Реализация криптографической защиты на уровне сетевых сокетов гарантирует конфиденциальность и целостность передаваемых данных.

2. Разработка основных компонентов модульной мультиагентной системы тестирования защищенности информационных систем

Архитектура разрабатываемой мультиагентной системы строится на принципе микросервисной организации: каждый агент реализован как автономный микро-сервис, функционирующий в изолированной среде (*Docker*-контейнера). Подобная структура обеспечивает эффективное масштабирование системы и возможность независимой модернизации компонентов.

Ключевым элементом архитектуры разрабатываемой мультиагентной системы выступает интеграция интеллектуального модуля *DeepSeek Code v2*, предназначенного для автоматизированной генерации, исследования и оптимизации эксплойтов. Данный компонент, базирующийся на современных языковых моделях, обеспечивает тесное взаимодействие с Управляющим Агентом, позволяя системе динамически адаптировать стратегии атак в соответствии с изменениями целевой инфраструктуры и поступлением новых данных. Механизм обработки информации в режиме реального времени обеспечивает формирование оптимальных стратегических решений, существенно повышая точность и оперативность реагирования мультиагентной системы. Данная функциональность позволяет осуществлять мгновенную адаптацию тактики тестирования в ответ на модификации сетевой инфраструктуры, максимизируя эффективность пентестинга. Автоматизированная настройка параметров функционирования оптимизирует использование вычислительных ресурсов и повышает результативность тестирования. Динамическая природа системы обеспечивает прецизионное реагирование на трансформации целевой среды, гарантируя высокую адаптивность процессов тестирования. Структурная организация разрабатываемой мультиагентной системы базируется на современных принципах *DevSecOps*, интегрируя технологии контейнеризации, оркестрации и методологии *CI/CD*. Стандартизация процессов развертывания и администрирования реализована посредством *Docker* и *Docker Compose*, автоматизация сборки осуществляется через конфигурационные файлы и *Makefile*. Межкомпонентное взаимодействие организовано через систему разделяемых директорий с многоуровневым контролем доступа и централизованным управлением ключами, обеспечивая максимальную защищенность и эффективность администрирования. Автоматизированная генерация криптографических ключей интегрирована в процесс развертывания, гарантируя безопасность информационного обмена и строгую авторизацию доступа к агентским модулям. Архитектура модульной мультиагентной системы реализует принципы масштабируемости и адаптивности к эволюционирующим требованиям информационной безопасности. Система поддерживает интеграцию дополнительных агентов, имплементацию инновационных методологий анализа и эксплуатации, внедрение передовых решений машинного обучения и искусственного интеллекта. Данные возможности обеспечивают автоматическую генерацию сценариев атак и динамическую адаптацию к изменениям условий тестирования.

2.1. Типизация сценария действия агентов на основе руководства по классификации и описанию кибератак и вторжений (*MITRE ATT&CK*)

Создание сценариев тестирования на проникновение в условиях, приближенных к реальным, создает необходимость системного подхода к описанию действий каждого из агентов. В основе анализа лежит концепция типизации агентов [11], при которой каждый агент формализуется как исполнитель конкретных техник из базы знаний *MITRE ATT&CK*. Предложенная модель опирается на трехуровневое представление мультиагентной системы в виде ориентированного графа G (A, T, I), где:

A — множество агентов (Агент_сканер_портов, Агент, сканирующий уязвимости и др.),

T — множество техник и тактик (например, T1046, T1210 и др.),

I — множество инструментов, обеспечивающих реализацию (например, *nmap*, *Metasploit, Kerberoast*).

Использование данной модели дает возможность формализовать логику поведения агентов для обеспечения сопоставимости с индустриальными стандартами и задать методологическую основу, которая затем будет использоваться для осуществления масштабирования, проектирования и верификации действий агентов в рамках атакующих сценариев. Типизация представляет собой процесс взаимосвязи действий агентов к строго определенным техникам и тактикам, которые описаны в рамках матрицы *MITRE ATT&CK*. Данный подход позволяет обеспечить не только формальное поведение агентов, но и связать каждый шаг атаки с конкретной стадией жизненного цикла злоумышленника, начиная от разведки и заканчивая техниками повышения привилегий при закреплении в системе, а также этапами постэксплуатации. Данный шаг позволяет повысить эффективность действий агентов, упростить контроль и адаптацию к различным условиям сетевой инфраструктуры.

В архитектуре мультиагентной системы каждый агент обладает закрепленным за ним набором техник и тактик, которые в свою очередь соотносятся с их функциональным профилем.

Для примера: агент, отвечающий за сканирование портов, действует согласно техникам T1046 (Network

Таблица 1.

Матрица инцидентности W_{AT} — соответствие агентов реализуемым техникам руководства по классификации и описанию кибератак и вторжений (*MITRE ATT&CK*)

Агент	T1046	T1018	T1203	T1068	T1059	T1210	T1003	T1558
Агент сканирования портов	1	1	0	0	0	0	0	0
Агент анализа уязвимостей	0	0	1	1	0	0	0	0
Агент эксплуатации	0	0	0	0	1	1	0	0
Агент атак на Active Directory	0	0	0	0	0	0	1	1

Service Discovery) и T1018 (Remote System Discovery), суть которых заключается в действиях по сбору информации об открытых портах, удаленных хостах, определении типа операционных систем.

Агент, осуществляющий анализ уязвимостей, использует технику T1203 (*Exploitation for Client Execution*) в совокупности с T1068 *(Exploitation for Privilege Escalation),* которые ориентируются на данные, полученные в ходе разведки, что позволяет эффективно выявлять уязвимые сервисы, уязвимости которых подвержены эксплуатации.

Агент, отвечающий за эксплуатацию уязвимостей, действует на основе техник *T1059 (Command and Scripting Interpreter)* и *T1210 (Exploitation of Remote Services).* В его задачи входит реализация активных фаз атаки, а именно: выполнение кода на целевой системе, эксплуатация эксплоитов на основе обнаруженных уязвимостей, а также инициирование взаимодействия с удаленными сервисами. Типизация в данном случае позволяет задать ограничения по допустимым действиям, исключив нежелательное или случайное отклонение от обозначенных сценариев.

Агент, который реализует атаки на Active Directory, осуществляет свои действия в рамках техник *T1003 (Credential Dumping)* и *T1558 (Steal or Forge Kerberos Tickets),* что дает возможность моделировать реальные угрозы, которые связаны с компрометацией AD-инфраструктуры. Типизация действий данного агента необходима как для автоматизации, так и для соблюдения ограничений в рамках регулируемой среды. Это создает возможность для гибкого расширения мультиагентной системы, интеграцию с другими инструментами и платформами.

Наличие типизированных моделей существенно упрощает аудит и тестирование агентов, позволяя осуществить формализованное и ожидаемое поведение, задействованное в процессе атаки.

3. Оценка эффективности использования модульной мультиагентной системы по критерию покрытия

Для осуществления оценки эффективности взят один из ключевых показателей эффективности мультиагентной системы — полнота покрытия техник, которая соответствует реальной атакующей активности, осуществляемой командой тестировщиков на проникновение. В отличие от классических, статических средств анализа использование мультиагентной системы предполагает распределение ролей между автономными агентами, каждый из которых реализует определенный сценарий, аналогичный действиям специалистов наступательной безопасности. В подобных условиях имеется возможность достаточно полно оценить масштаб охвата техник и тактик, задействованных системой, и соотнести его с допустимым уровнем, который соответствует ручной работе опытной команды пентестеров.

Для оценки полноты покрытия функционала пентест-команды вводим коэффициент покрытия σ, который определяет соотношение количества реализуемых агентами техник с числом техник, которые могут быть реализованы в реальной инфраструктуре командой из восьми специалистов наступательной безопасности:

$$\sigma = \frac{|T_{realized}|}{|T_{expected}|},$$

где: T(realized) — количество техник, покрытых агентами системы, T(expected) — ориентировочный минимум техник, реализуемый командой из восьми специалистов в реальной инфраструктуре.

Следующая формула позволяет оценить относительную эффективность автоматизации:

$$E = \frac{K_a \cdot T_r}{T_x}$$

где: K_a — число одновременно работающих агентов, T_r — среднее время выполнения задачи вручную (в минутах), T_x — среднее время выполнения задачи агентом.

Пример:

если K_a = 6; T_r = 45 мин; T_x = 5 мин, то E = 54.

Формула учитывает количество одновременно работающих агентов, средний показатель времени продолжительности выполнения задачи при ручном тестировании, а также среднее время выполнения задачи в автоматическом режиме. Таким образом, введенная в рамках данного исследования метрика эффективности — E демонстрирует совокупность эффектов от масштабируемости и скорости реакции агентов мультиагентной системы, которые действуют в рамках координируемого сценария.

Пример расчета демонстрирует, что при наличии шести действующих агентов общая эффективность достигает более чем 50-кратного значения. Иными словами, это означает, что мультиагентная система обеспечивает более чем пятикратное ускорение на каждого агента, что в совокупности подтверждает значительное преимущество разработанного подхода.

3.1 Оценка эффективности использования модульной мультиагентной системы по временному покрытию

В рамках данного исследования при оценке эффективности мультиагентной системы важным значением является временно́й аспект функционирования агентов. В условиях реального тестирования на проникновение атаки выполняются в ограниченном промежутке времени, в течение которого должны пройти основные этапы сценария, включая разведку, где осуществляется сбор информации о целевой системе, эксплуатация уязвимостей, получение привилегированного доступа и закрепление в инфраструктуре. По этой причине важна не только полнота покрытия техник, но и способность системы работать синхронно и без задержек, что является критически важной характеристикой.

Для осуществления формализованной оценки поведения системы в зависимости от охвата техник и времени выполнения сценария целесообразно введение функции плотности вероятности успешной атаки, которая отражала бы устойчивость системы к усложнению инфраструктуры и ограниченности времени.

Данный график построен для разных значений параметра *t*, который отражает сложность инфраструктуры (и соответственно — продолжительность тестирования). Чем выше значение *t*, тем плавнее рост ее функции и ниже ее максимум, что демонстрирует влияние усложнения инфраструктуры на снижение эффективности при фиксированном покрытии. Функция также демонстрирует типичное затухание, а именно при малых значениях σ вероятность мала, затем наблюдается пик эффективности, а после него — экспоненциальное снижение.

Иными словами, применима формула, основанная на плотности охвата графа техник:

$$P(\sigma,t) = \frac{\sigma^2}{t^2} \cdot e^{-\sigma/\left(2t^2\right)},$$

где σ отражает коэффициент покрытия техник, которые реализованы агентами, а *t* — временно́й показатель, от-

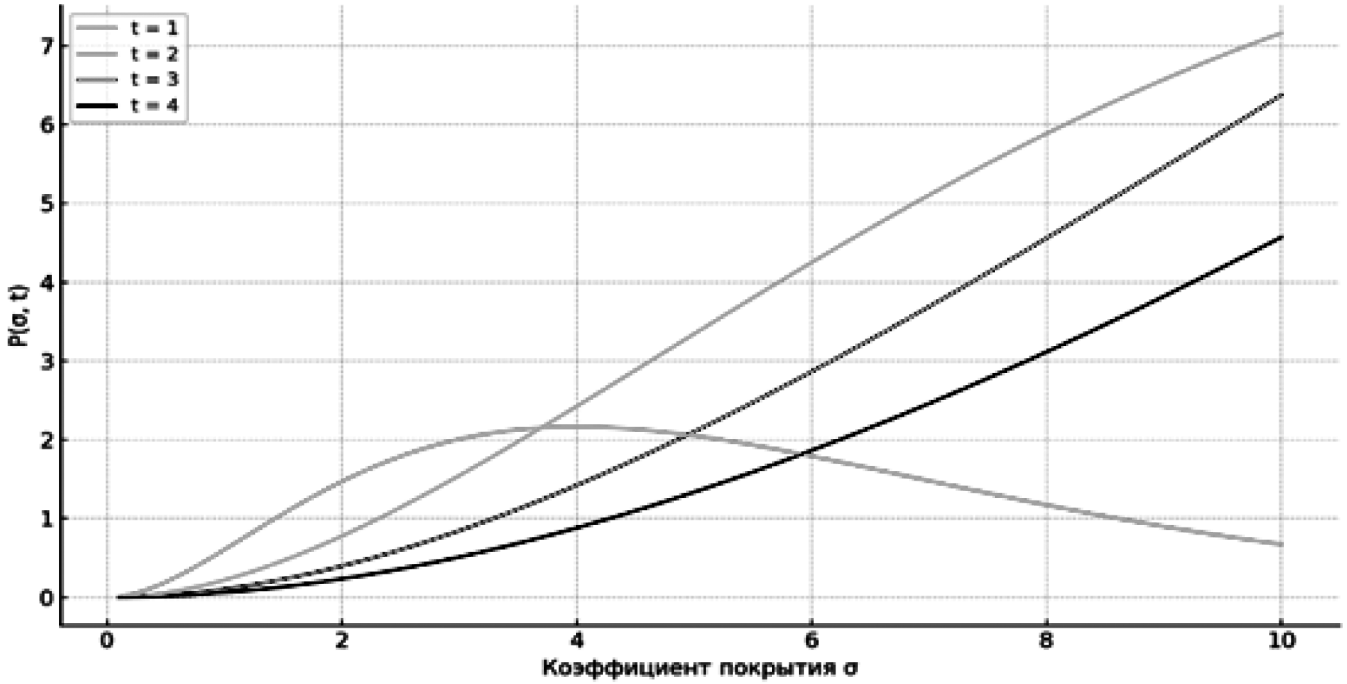

Рис. 1. График функции вероятности успешной атаки

ражающий затраченное системой время на осуществление проводимого тестирования.

Показатель *t* отражает условную оценку времени, которое необходимо для реализации сценария тестирования, пропорционален сложности атакуемой инфраструктуры и включает в себя совокупность факторов, а именно: объем используемых сервисов, количество целевых узлов, глубину сети, наличие защитных механизмов и иных ограничений, которые влияют на продолжительность выполнения атаки.

$P(\sigma,t)$ — функция плотности вероятности успешной атаки, моделирующая эффективность системы в условиях заданного покрытия техник и сложности атакуемой инфраструктуры. Чем выше значение этой функции, тем больше вероятность, что агенты смогут успешно атаковать целевую систему. **e** — основание натурального логарифма, используемое для экспоненциального затухания функции при увеличении сложности инфраструктуры.

$\frac{\sigma^2}{2t^2}$ — экспоненциальный член, который моделирует резкое снижение вероятности успешной атаки при недостаточном покрытии техник (малое значение σ) или при высокой сложности инфраструктуры (большое значение *t*). На практике значение *t* может быть установлено экспертно, на основании числа целевых хостов, глубины сети, наличия механизмов защиты и среднего времени, которое затрачивается на компрометацию аналогичных сред. Эта функция моделирует устойчивость системы по аналогии с вероятностной плотностью распределения Гаусса: при недостижении критического σ система не обеспечивает достаточной устойчивости.

Таким образом, типизация агентов и соответствие их функций техникам *MITRE ATT&CK* позволяют провести комплексную оценку эффективности разрабатываемой мультиагентной системы. За счет параллельности, модульности и интеграции с генеративным интеллектом *DeepSeek Code v2* платформа демонстрирует показатели, сравнимые и в ряде сценариев превосходящие работу группы специалистов. Система формализует знания, адаптируется к инфраструктуре и по метрике σ покрывает до 88–94 % ключевых техник атак на типовые корпоративные сети.

Таким образом, разработанная мультиагентная система может не только эффективно воспроизводить действия команды из более чем восьми пентестеров, но и формировать базу знаний для дальнейшего обучения и масштабируемой адаптации, что делает ее полноценной интеллектуальной *offensive*-платформой.

Заключение

Исследование фокусируется на разработке инновационной мультиагентной платформы, обеспечивающей динамическую адаптацию к изменениям операционной среды и автономное выполнение задач посредством интеграции с DeepSeek Code v2 для интеллектуальной генерации эксплойтов. Также возможно интеграция и с иными интеллектуальными модулями, либо создания собственного, основанного на принципах глубокого машинного обучения. Концептуальная архитектура разработанной в рамках данного исследования мультиагентной системы существенно превосходит традиционные подходы к автоматизации тестирования на проникновение, реализуя принципы мультиагентной организации с распределением функциональных ролей и централизованным интеллектуальным управлением. Функциональная специализация агентов обеспечивает четкую сегментацию задач и зон ответственности компонентов системы. Применение микросервисной архитектуры и контейнерной изоляции гарантирует отказоустойчивость, адаптивность к инфраструктурным изменениям и надежность межагентных коммуникаций. Реализованный подход минимизирует критические точки отказа, повышая устойчивость системы к непрогнозируемым модификациям сетевой среды.

Экспериментальные исследования демонстрируют исключительную производительность системы, обеспечивая 50-кратное увеличение эффективности по сравнению с традиционным подходом к пентесту. Количественные показатели подтверждают: при задействовании шести агентских модулей длительность выполнения стандартных операций сократилась с 45 до 5 минут, формируя уникальные характеристики производительности и адаптивности разработанной мультиагентной системы. Комплексное тестирование в условиях, моделирующих реальные корпоративные инфраструктуры, подтверждает существенные преимущества данной модульной мультиагентной системы. Реализуемая мультиагентная система демонстрирует превосходные показатели автоматизации процессов, интеллектуальной эволюции поведенческих паттернов агентов и адаптивности к трансформациям целевой среды. По ключевым метрикам эффективности — оперативность реагирования, глубина аналитики, динамическая адаптация — разработанная мультиагентная система демонстрирует значительное превосходство над существующими решениями при сохранении открытой архитектуры, обеспечивающей возможности модификации и масштабирования функционала.

ЛИТЕРАТУРА

1. Michael Wooldridge. An Introduction to MultiAgent Systems. ISBN 978-0471496915.
2. Katia Sycara. Multiagent Systems. DOI: 10.1145/220605551.
3. Shlomo Zilberstein. Decentralized Partially Observable Markov Decision Processes (Dec-POMDPs).
4. Magnus Egerstedt. Graph Theoretic Methods in Multiagent Networks. ISBN 978-1107070945.
5. Alessio Lomuscio. MCMAS: A Model Checker for the Verification of Multi-Agent Systems.
6. Frank Dignum. Socially Aware Agents. DOI: 10.1007/978-3-319-91848-8_10.
7. Leon van der Torre. Deontic Logic and Normative Multiagent Systems. ISBN 978-3030402921.
8. Jakob Foerster, Shimon Whiteson. Counterfactual Multi-Agent Policy Gradients. arXiv:1705.08926.
9. Mikayel Samvelyan et al. The StarCraft Multi-Agent Challenge. arXiv:1902.04043.
10. Stefano V. Albrecht et al. Deep Reinforcement Learning for Multi-Agent Interaction. arXiv:2208.01769.
11. Сизов В.А., Киров А.Д. Разработка модели актуализации профессиональных компетенций специалиста по кибербезопасности в условиях информационного противоборства. УДК 372.8, 004.056. DOI 10.54835/18102883_2023_34_6. URL: https://old.aeer.ru/files/io/m34/art_6.pdf
12. ГОСТ Р 56875–2016. Информационные технологии (ИТ). Системы безопасности комплексные и интегрированные. Типовые требования к архитектуре и технологиям интеллектуальных систем мониторинга для обеспечения безопасности предприятий и территорий. URL: https://docs.cntd.ru/document/1200132478.
13. Приказ ФСБ России от 11 мая 2023 г. № 213 «Об утверждении порядка осуществления мониторинга защищенности информационных ресурсов». URL: https://www.garant.ru/products/ipo/prime/doc/406876800/.
14. Приказ ФСТЭК России от 14 апреля 2023 г. № 64 «Требования по безопасности информации». URL: https://fstec.ru/dokumenty/vse-dokumenty/spetsialnye-normativnye-dokumenty/trebovaniya-po-bezopasnosti-informatsii-utverzhdeny-prikazom-fstek-rossii-ot-14-aprelya-2023-g-n-64.
15. Приказ ФСБ России от 18 марта 2025 г. № 117 «Об утверждении Требований о защите информации, содержащейся в государственных информационных системах». URL: https://www.garant.ru/products/ipo/prime/doc/411649021/.
16. Приказ ФСТЭК РФ от 11 февраля 2013 г. № 17 «Об утверждении требований о защите информации, не составляющей государственную тайну, содержащейся в государственных информационных системах». URL: https://normativ.kontur.ru/document?documentId=481221&moduleId=1.
17. ГОСТ Р 53114–2008. Защита информации. Обеспечение информационной безопасности в организации. Основные термины и определения. URL: https://meganorm.ru/Data/484/48411.pdf.
18. ГОСТ Р 59516–2021. Информационные технологии (ИТ). Страхование рисков информационной безопасности. Рекомендации по выбору и применению. URL: https://docs.cntd.ru/document/1200179668.
19. Приказ ФСТЭК России от 27 октября 2022 г. № 187 «Требования по безопасности информации». URL: https://fstec.ru/dokumenty/vse-dokumenty/spetsialnye-normativnye-dokumenty/trebovaniya-po-bezopasnosti-informatsii-utverzhdeny-prikazom-fstek-rossii-ot-27-oktyabrya-2022-g-n-187.
20. Приказ ФСТЭК России от 30 июля 2018 г. № 131 «Требования по безопасности информации, устанавливающие уровни доверия к средствам технической защиты информации и средствам обеспечения безопасности информационных технологий». URL: https://docs.cntd.ru/document/565278856.
21. Приказ ФСТЭК России от 25 декабря 2017 г. № 239 «Об утверждении Требований по обеспечению безопасности значимых объектов критической информационной инфраструктуры РФ». URL: https://fstec.ru/dokumenty/vse-dokumenty/prikazy/prikaz-fstek-rossii-ot-25-dekabrya-2017-g-n-239.
22. Лихтенштейн В.Е., Конявский В.А., Росс Г.В., Лось В.П. Мультиагентные системы: самоорганизация и развитие. — М.: Финансы и статистика, 2018. — 264 с.
23. Бессмертный И.А. Искусственный интеллект. Введение в многоагентные системы. — М.: Юрайт, 2024. — 148 с.
24. Радченко И.А. Интеллектуальные мультиагентные системы: учебное пособие. — СПб.: БГТУ, 2016.
25. Ивашкин Ю.А. Мультиагентное моделирование в имитационной системе Simplex3. — М.: 2024.
26. Тарасов В.Б. От многоагентных систем к интеллектуальным организациям. — М.: Эдиториал, 2002.
27. Белов Е.Б., Лось В.П., Мещеряков Р.В., Шелупанов А.А. Основы информационной безопасности: учебное пособие. — М.: Горячая линия — Телеком, 2020.
28. Бузов Г.А., Калинин С.В., Кондратьев А.В. Защита от утечки информации по техническим каналам: учебное пособие. — М.: Горячая линия — Телеком, 2005. — 416 с.
29. Запечников С.В., Милославская Н.Г., Толстой А.И. Информационная безопасность открытых систем. Часть 1: учебник для вузов. — М.: 2019.
30. Граймс Р. Как противостоять хакерским атакам. Уроки экспертов по информационной безопасности. — М.: 2021.
31. Скабцов Н. Kali Linux в действии. Аудит безопасности информационных систем. — М.: 2020.
32. Кукшкина Н.С., Воронкин Е.Ю. Исследование возможности применения мультиагентных систем для организации работы технической поддержки внутри организации. — 2023. URL: https://geosib.sgugit.ru/upload/geosibir/sborniki/2023/tom-6/148-152.pdf.
33. Гриценко В.И., Гладун А.Я., Журавлев Ю.Д., Несен М.В. Модель мультиагентной системы для е-бизнеса и технология ее программной реализации. — 2016. URL: https://core.ac.uk/download/pdf/38331895.pdf.
34. Матвеева А.А., Ким Ю.В., Викснин И.И. Методы обеспечения информационной безопасности коммуникационных каналов в мультиагентных робототехнических системах. — 2022. URL: https://cyberleninka.ru/article/n/metody-obespecheniya-informatsionnoy-bezopasnosti-kommunikatsionnyh-kanalov-v-multiagentnyh-robototehnicheskih-sistemah.
35. Печеркин С.А. Взаимодействие агентов в мультиагентных системах. — 2021. URL: https://cyberleninka.ru/article/n/vzaimodeystvie-agentov-v-multiagentnyh-sistemah.
36. Лахтер М.Д. Мультиагентные технологии как инструмент перехода к экономике знаний. — 2020. URL: https://cyberleninka.ru/article/n/multiagentnye-tehnologii-kak-instument-perehoda-k-ekonomike-znaniy.
37. Колесникова Г.И. Искусственный интеллект: проблемы и перспективы. — 2023. URL: https://cyberleninka.ru/article/n/iskusstvennyy-intellekt-problemy-i-perspektivy.
38. Капустин Ф.А. Информационная безопасность и защита информации в современном обществе. — 2021. URL: https://cyberleninka.ru/article/n/informatsionnaya-bezopasnost-i-zaschita-informatsii-v-sovremennom-obschestve-1.
39. Журнал «Информационная безопасность». Различные статьи по информационной безопасности. — 2025. URL: https://www.itsec.ru/articles.

Журнал «Современная наука: актуальные проблемы теории и практики»