УДК 004.056

Валеев С.С., Гузаиров М.Б.
Уфимский университет науки и технологий, Уфа

ГЕНЕРАТИВНЫЕ МНОГОАГЕНТНЫЕ СИСТЕМЫ И ЗАДАЧИ ЗАЩИТЫ ИНФОРМАЦИИ НА БАЗЕ КОНЦЕПЦИИ НУЛЕВОГО ДОВЕРИЯ

Аннотация. Распределенные мультиагентные системы представляют собой перспективное направление развития технологий искусственного интеллекта. Они позволяют объединить усилия множества интеллектуальных агентов, размещенных на разных серверных площадках для совместного решения масштабных и комплексных задач. Эта парадигма решает проблему ограниченности локальных вычислительных

мощностей и дает возможность задействовать специфические компетенции отдельных моделей машинного обучения. Рассматривается аспект задач, связанных с вопросами обеспечения информационной безопасности в этих системах. Обсуждается решение задач защиты информации в мультиагентной распределенной системе на базе концепции нулевого доверия в рамках мультиагентных технологий.

Ключевые слова: программный агент, мультиагентные системы, распределенные вычисления, угрозы безопасности, концепция нулевого доверия.

Распределенные мультиагентные системы (РМС) позволяют объединить гетерогенные вычислительные системы для повышения эффективности достижения цели и относятся к классу распределенных вычислительных систем [1]. Рассмотрим далее основные особенности архитектуры РМС.

Агенты в рамках РМС могут иметь разные характеристики и специализации. Например, один агент может отвечать за обработку естественного языка, другой – за компьютерное зрение, третий – за принятие решений и управление ресурсами. Каждый агент представляет специализированный сервис, тем самым, предоставляя свои сервисы и знания другим участникам сетевого взаимодействия.

Для эффективной координации действий всех участников РМС необходима единая система протоколов обмена информацией между агентами. Это достигается благодаря использованию открытых API-интерфейсов, обеспечивающих совместимость и интеграцию компонентов независимо от среды исполнения и особенностей конкретных вычислительных платформ.

Несмотря на децентрализованный характер вычисления, система имеет иерархическую архитектуру, т. е. централизованная управляющая подсистема, размещенная в РМС, координирует процессы всей вычислительной системы. Она определяет порядок взаимодействий между агентами, распределяет задачи и ресурсы, обеспечивает мониторинг производительности и надежности каждой подсистемы.

Данный подход к организации вычислений имеет как преимущества, так и недостатки. К преимуществам можно отнести:

– возможность подключения новых вычислительных узлов, что позволяет повысить эффективность системы без значительного увеличения нагрузки на существующие подсистемы;

– повышение отказоустойчивости на основе разнообразия функционала агентов и разделению функций между ними;

– разделение задач между несколькими узлами, что позволяет в ряде случаев снижать требования к мощности отдельного сервера и оптимизировать процессы использования вычислительных ресурсов.

К недостаткам, с точки зрения информационной безопасности, следует отнести:

– нарушение конфиденциальности, связанное с распределенным характером архитектуры РМС (отсутствие единого периметра безопасности);

– так как агенты расположены на разных серверах, злоумышленники могут реализовать проникновение в РМС через наименее защищенный узел, или возможна реализация распределенной атаки (эффект слабого звена);

– обмен конфиденциальной информацией между удаленными подсистемами требует гармонизации политик безопасности, что не всегда возможно по разным причинам (необходима разработка стандартов безопасности);

– низкое качество сетевых соединений или перебои с доступом к сети могут привести к потере синхронизации между агентами, что нарушает работоспособность системы;

– сбой одного вычислительного узла способен вызвать каскадный эффект, негативно влияющий на производительность и безопасность всей сети;

– сложность мониторинга и диагностики системы, поэтому выявление причины сбоев требует больших затрат, особенно если проблема возникает одновременно на нескольких уровнях иерархии РМС;

– центр должен постоянно отслеживать состояние всех агентов, обеспечивая балансировку нагрузок и своевременное реагирование на изменения условий эксплуатации;

– «бесплатный проезд» – поставщики услуг могут эксплуатировать общие ресурсы без равноценного вклада, заменяя запрошенную мощную БЯМ (большая языковая модель) на менее простую модель для снижения своих затрат на вычисления.

Рассмотрим далее предлагаемые решения в рамках концепции нулевого доверия [2].

Концепция нулевого доверия предполагает отсутствие автоматического доверия ко всем субъектам и объектам внутри корпоративной сети, независимо от их расположения – внутренний или внешний доступ требует постоянного подтверждения подлинности и проверки полномочий каждого субъекта перед предоставлением доступа к ресурсам [3,4]. Роль многоагентной системы защиты информации (МСЗИ) становится особенно значимой, поскольку она способна обеспечить гибкое и адаптивное управление политиками безопасности в РМС путем динамического контроля состояния информационной среды организации.

Преимущества использования МСЗИ в рамках концепции нулевого доверия:

– благодаря наличию множества программных агентов, которые работают параллельно и выполняют мониторинг разных участков РМС,

МСЗИ способна значительно сократить время обнаружения угроз и ускорить реакцию на инциденты безопасности. Каждый агент МСЗИ собирает локальную информацию о своем сегменте сети РМС, обрабатывает ее и передает информацию другим агентам МСЗИ. Это ускоряет процесс выявления аномалий в РМС и повышает общую эффективность обнаружения атак, обеспечивают постоянное наблюдение за всеми элементами РМС, охватывая даже удаленные вычислительные узлы и ресурсы, позволяя эффективно контролировать безопасность больших территориально распределенных сетей;

– каждое событие, происходящее в любом узле сети РМС, фиксируется соответствующим агентом МСЗИ и передается на обработку другим компонентам системы МСЗИ. Такой подход помогает выявлять скрытые каналы передачи данных, злоупотребления привилегиями и другие типы сложных атак, которые могут остаться незамеченными традиционными средствами мониторинга системы;

– разделение функций между агентами МСЗИ позволяет детально исследовать каждое потенциально опасное действие в РМС отдельно, снижая нагрузку на центральный кластер обработки данных и повышая точность аналитики. Использование распределенных алгоритмов машинного обучения в МСЗИ также дает возможность быстро выявить новые виды угроз и своевременно обновлять политики безопасности в РМС.

Рассмотрим далее набор основных методов адаптации политик безопасности в РМС. Для эффективного функционирования РМС в среде нулевого доверия используются следующие методы адаптации правил безопасности:

– применение многофакторной аутентификации в случае любого отклонения поведения агента РМС или процесса от ожидаемых норм, обеспечивая дополнительный уровень уверенности в подлинности запросов;

– мониторинг активности в сети: постоянный контроль изменений трафика, поведенческих моделей агентов РМС и процессов помогает вовремя идентифицировать возможные атаки и предотвратить реализацию угроз;

– проверка целей внедрения нового кода в РМС, оценка его функциональности и потенциального воздействия на РМС.

Перечисленные меры дополняют друг друга, создавая комплексную защитную среду, способную оперативно адаптироваться к изменениям условий эксплуатации и угрозам.

Таким образом, использование МСЗИ для реализации задач защиты информации в РМС представляет собой необходимый инструмент в рамках концепции нулевого доверия, обеспечивающий требуемую степень надежности РМС и эффективности процедур управления рисками.

Список использованных источников:

1. Тарасов В.Б. Агенты, многоагентные системы, виртуальные сообщества: стратегическое направление в информатике и искусственном интеллекте // Новости искусственного интеллекта. 1998. № 2. С. 5–63.
2. Многоагентные системы как технологическая база реализации концепции нулевого доверия / С.С. Валеев, Н.В. Кондратьева, М.Б. Гузаиров, А.С. Исмагилова // Вестник Российского нового университета. Серия: Сложные системы: модели, анализ и управление. 2024. № 3. С. 116–123. DOI 10.18137/RNU.V9187.24.03.P.116.
3. Валеев С.С., Кондратьева Н.В. Особенности проектирования систем безопасности на базе архитектуры нулевого доверия // ivdon.ru – 2023. URL: http://www.ivdon.ru/uploads/article/pdf/IVD_68__8_valeev_kondratyeva_v2.pdf_72458b243f.pdf (дата обращения: 30.04.2025).
4. Валеев С.С., Кондратьева Н.В., Мельников А.В. Архитектура предприятия и архитектура нулевого доверия // info-secur.ru – 2023. URL: https://www.info-secur.ru/index.php/ ojs/article/download/413/371/ (дата обращения: 30.04.2025).

Valeev S.S., Guzairov M.B.
Ufa University of Science and Technology, Ufa

GENERATIVE MULTI-AGENT SYSTEMS AND INFORMATION SECURITY TASKS BASED ON THE CONCEPT OF ZERO TRUST

Abstract. Distributed multi-agent systems represent a promising area for the development of artificial intelligence technologies. They allow us to combine the efforts of many intelligent agents located on different server sites to jointly solve large-scale and complex tasks. This solves the problem of limited local computing power and makes it possible to use the specific competencies of individual machine learning models. The aspect of tasks related to the issues of ensuring information security in these systems is considered. The solution of information security problems in a multi-agent distributed system based on the concept of zero trust within the framework of multi-agent technologies is discussed.

Keywords: software agent, multi-agent systems, distributed computing, security threats, zero trust concept.