**УДК 004.89:004.7**

**Шахназарян Артём Арменович,** магистрант, МИРЭА - Российский технологический университет, г. Москва

# ПОВЫШЕНИЕ ЭФФЕКТИВНОСТИ ОБНАРУЖЕНИЯ РАСПРЕДЕЛЕННЫХ СЕТЕВЫХ АТАК НА ОСНОВЕ КОЛЛЕКТИВНОГО ВЗАИМОДЕЙСТВИЯ ИНТЕЛЛЕКТУАЛЬНЫХ АГЕНТОВ

В работе рассматривается проблема обеспечения безопасности компьютерных сетей в условиях роста интенсивности распределенных кибератак. Предложен теоретический подход к построению системы обнаружения вторжений (СОВ), основанный на мультиагентной архитектуре с использованием механизма динамического консенсуса. Описана формальная модель взаимодействия агентов и алгоритм принятия решений на основе репутации. Проведена аналитическая оценка предложенного метода, демонстрирующая его преимущества в части масштабируемости и отказоустойчивости по сравнению с централизованными аналогами. Обоснована перспективность применения коллективного интеллекта для снижения вероятности ложных срабатываний в распределенных системах.

**Ключевые слова:** мультиагентные системы, информационная безопасность, обнаружение вторжений, коллективный интеллект, сетевые атаки, алгоритмы консенсуса, моделирование защиты.

The paper discusses the problem of ensuring computer network security amidst the growing intensity of distributed cyberattacks. A theoretical approach to building an intrusion detection system (IDS) based on a multi-agent architecture using a dynamic consensus mechanism is proposed. A formal model of agent interaction and a reputation-based decision-making algorithm are described. An analytical assessment of the proposed method is conducted, demonstrating its advantages in terms of scalability and fault tolerance compared to centralized

counterparts. The prospects of using collective intelligence to reduce the probability of false positives in distributed systems are substantiated.

**Keywords:** multi-agent systems, information security, intrusion detection, collective intelligence, network attacks, consensus algorithms, security modeling.

## Введение

Современные корпоративные сети характеризуются высокой степенью гетерогенности и размытостью периметра. Традиционные централизованные системы обнаружения вторжений (IDS) сталкиваются с фундаментальными ограничениями: наличием единой точки отказа и невозможностью оперативного анализа трафика в высоконагруженных распределенных сегментах [4].

Перспективным направлением развития средств защиты является применение мультиагентных систем (МАС). Согласно классическим исследованиям в области искусственного интеллекта [1, 2], децентрализованные системы, состоящие из автономных агентов, способны эффективно решать задачи в условиях неопределенности. Однако, простого внедрения агентов недостаточно; ключевой проблемой остается разработка алгоритмов их взаимодействия (кооперации) для формирования единой картины состояния безопасности сети.

Целью данной работы является разработка архитектурной модели и алгоритма коллективного взаимодействия интеллектуальных агентов, обеспечивающих повышение достоверности обнаружения атак без необходимости передачи больших объемов сырых данных на центральный сервер.

## Архитектура мультиагентной системы безопасности

Для реализации децентрализованной защиты предлагается иерархическая модель, базирующаяся на принципах построения интеллектуальных систем, описанных в работе Г.В. Рыбиной [3]. Система включает агентов, разделенных по функциональному назначению (Рисунок 1):

1. Агенты-сенсоры (Sensor Agents). Размещаются на хостах или коммутационном оборудовании. Их задача — первичная фильтрация трафика и сигнатурный анализ.
2. Агенты-аналитики (Analysis Agents). Агрегируют события от сенсоров внутри сегмента сети и выявляют аномалии методами машинного обучения.
3. Агенты-координаторы (Coordinator Agents). Образуют оверлейную сеть для обмена знаниями (метаданными) между сегментами и реализации механизма консенсуса.

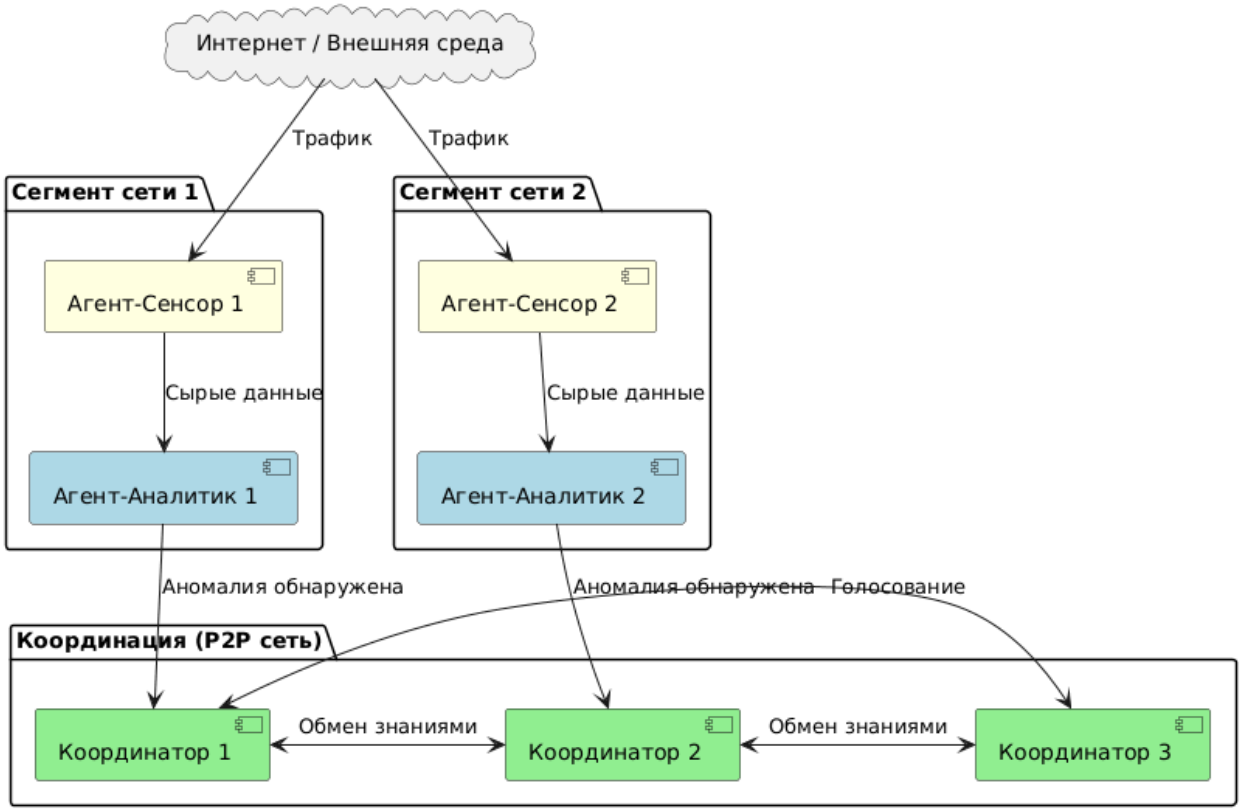


Рисунок 1. Схема взаимодействия агентов в распределенной системе обнаружения атак

Ключевым отличием предложенной архитектуры является отсутствие главного управляющего узла. Решение о блокировке принимается коллегиально, что повышает живучесть системы: выход из строя одного или нескольких координаторов не парализует работу всей системы защиты.

**Алгоритм коллективного принятия решений**

В основу взаимодействия положен алгоритм взвешенного голосования. Каждый агент-координатор обладает параметром «репутация» ($R_i$), который отражает историческую достоверность его вердиктов.

Процесс принятия решения описывается следующим алгоритмом (Рисунок 2):

1. При обнаружении подозрительной активности локальный агент инициирует запрос «Check_Request» к соседним узлам.
2. Соседние узлы проверяют свои логи на наличие коррелирующих событий (например, запросов с того же IP-адреса).
3. Формируется вектор голосов, где голос каждого агента взвешивается на его репутацию.
4. Если взвешенная сумма превышает пороговое значение T, атака считается подтвержденной.

Формально итоговый вердикт $V$ определяется как:

$$V = \begin{cases} 1, \text{если} \sum_{i=1}^{N} (R_i \cdot d_i) > T \\ 0, \text{иначе} \end{cases}$$

где $N$ – число агентов в кластере, $d_i$ – решение $i$-го агента (1 – атака, 0 – норма).

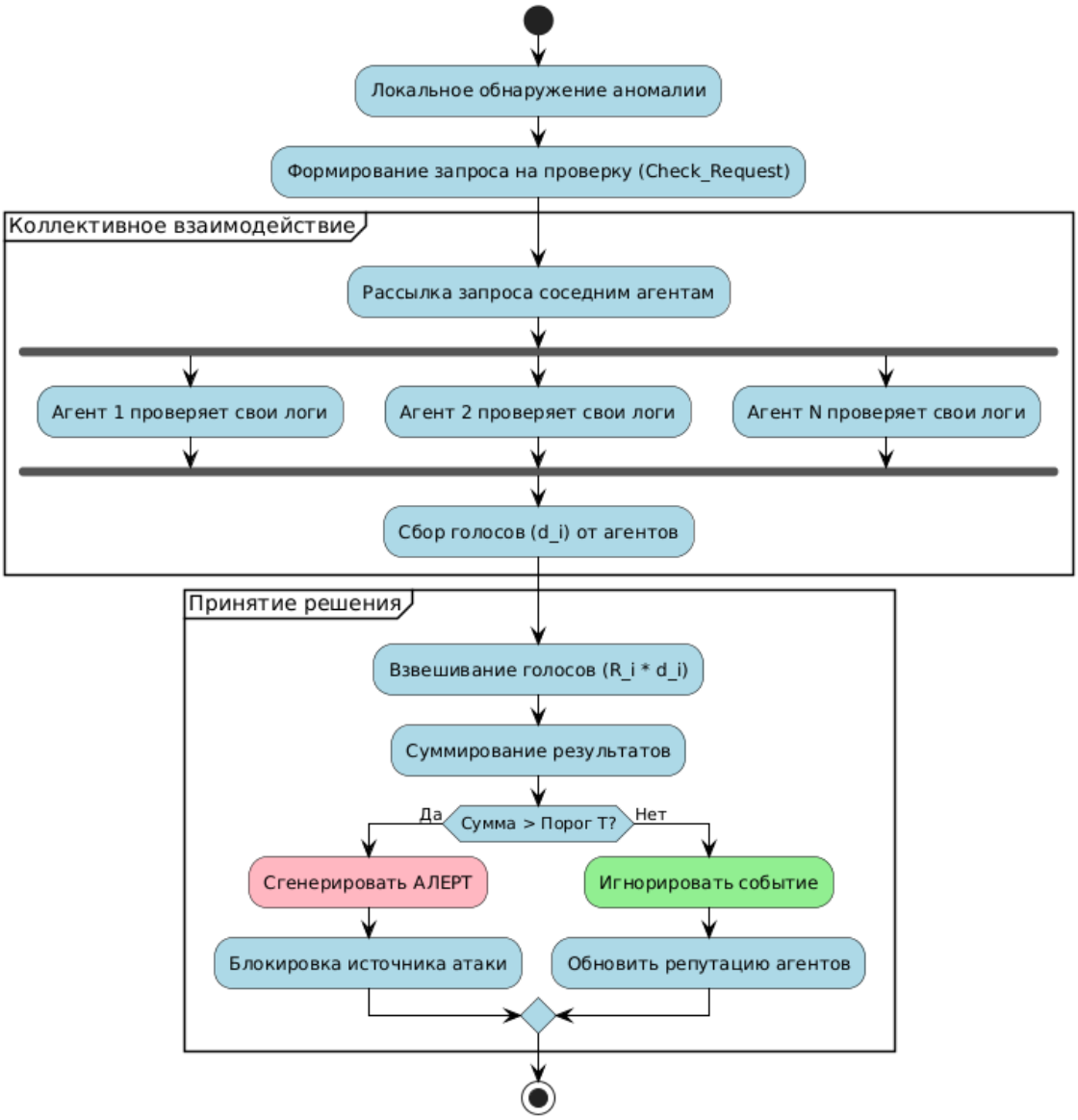


Рисунок 2. Блок-схема алгоритма достижения консенсуса в группе агентов

**Аналитическая оценка эффективности подхода**

Вместо прямого экспериментального сравнения, которое требует развертывания масштабного полигона, проведем теоретический анализ характеристик предложенной модели в сравнении с классическими централизованными системами.

Оценка надежности и отказоустойчивости. В централизованной системе вероятность отказа $P_{sys}$ равна вероятности отказа сервера $P_{srv}$. В предложенной мультиагентной модели, благодаря распределенному консенсусу [5], отказ системы наступает только при одновременной компрометации большинства узлов (k>N/2). Согласно теории надежности, вероятность отказа кластера стремится к нулю

при увеличении числа узлов, что делает систему устойчивой к DDoS-атакам, направленным на инфраструктуру защиты.

Оценка нагрузки на сеть. Традиционные SIEM-системы требуют передачи всех логов в центр. Объем трафика $Traffic_{SIEM} \approx \sum data_i$.

В предложенной модели передаются только вердикты и метаданные (сигналы тревоги). Объем служебного трафика при голосовании:

$$Traffic_{MAS} \approx N \cdot Size(Message)$$

Поскольку размер сообщения с вердиктом ничтожно мал по сравнению с сырым трафиком ($Size(Message) \ll data_i$), предложенный подход существенно снижает нагрузку на каналы связи, что критически важно для сетей IoT и промышленных систем.

Снижение вероятности ошибок (False Positives). Локальные системы часто дают ложные срабатывания на легитимные всплески активности. Коллективная верификация позволяет отфильтровать такие события: если агент A видит аномалию, а агенты B и C (находящиеся в других сегментах, но видящие тот же поток) ее не подтверждают, итоговый вердикт будет отрицательным. Это теоретически снижает частоту ошибок первого рода.

Заключение

В работе предложена архитектура мультиагентной системы обнаружения вторжений, использующая механизм взвешенного голосования. Теоретический анализ показал, что переход от централизованной обработки данных к коллективному взаимодействию интеллектуальных агентов позволяет:

1. Обеспечить функционирование системы защиты даже при выходе из строя части узлов.
2. Снизить нагрузку на магистральные каналы связи за счет локальной обработки данных.
3. Повысить достоверность обнаружения распределенных атак за счет перекрестной проверки гипотез агентами.

Дальнейшие исследования могут быть направлены на математическое моделирование оптимальных пороговых значений консенсуса для различных типов сетевых топологий.

## Список литературы


1. Городецкий В.И. Современное состояние и перспективы индустриальных применений многоагентных систем // Управление большими системами: сборник трудов. 2017. № 1. С. 44–59.
2. Тарасов В.Б. От многоагентных систем к интеллектуальным организациям: философия, психология, информатика. // УРСС, 2002. 352 с.
3. Рыбина Г.В. Основы построения интеллектуальных систем: учебное пособие. // Москва: Финансы и статистика, 2021. 432 с.
4. Макаренко С.И. Информационная безопасность: учебное пособие для студентов вузов. // Ставрополь: СФ МГГУ им. М.А. Шолохова, 2009. 372 с.
5. Таненбаум Э., Ван Стеен М. Распределенные системы. Принципы и парадигмы. // СПб.: Питер, 2003. 887 с.


## References


1. Gorodetsky V.I. Current state and prospects of industrial applications of multi-agent systems. // Large Scale Systems Control, 2017, no. 1, P. 44–59.
2. Tarasov V.B. From multi-agent systems to intellectual organizations: philosophy, psychology, informatics. // Moscow, URSS Publ., 2002. 352 p..
3. Rybina G.V. Fundamentals of building intelligent systems: textbook. // Moscow, Finansy i statistika Publ., 2021. 432 p.
4. Makarenko S.I. Information Security: textbook for university students. // Stavropol, SF MGGU im. M.A. Sholokhova Publ., 2009. 372 p.
5. Tanenbaum A., Van Steen M. Distributed Systems. Principles and Paradigms. // St. Petersburg, Piter Publ., 2003. 887 p.