

МЕТОДОЛОГИЯ ОПТИМИЗАЦИИ И ЭКСПЕРИМЕНТАЛЬНОЙ ОЦЕНКИ ЭФФЕКТИВНОСТИ МУЛЬТИАГЕНТНОЙ ОРКЕСТРАЦИИ

Першинов А.В. ¹

Научный руководитель – Ковальчук М. А. ¹

¹Университет ИТМО

anatoliypershinov@gmail.com

Введение

Методика использования агентных систем значительно расширяет функционал и возможности современных языковых моделей. Такие системы представляют собой совокупность нескольких LLM агентов, способных сообща решать достаточно сложные задачи. По мере увеличения количества агентов, эффективность системы всё в большей степени зависит от качества их координации: возрастают коммуникационные издержки, усиливается конкуренция за ресурсы и повышается чувствительность к частичным отказам. Современные работы связанные с мультиагентными системами подтверждают важность правильного распределения ресурсов [1].

Основная часть

В настоящее время существует множество работ и бенчмарков посвященных работе с языковыми моделями. Как правило это бенчмарки нацелены на оценивание работы отдельно взятого агента, а не системы целиком [2].

Объектом исследования являются подходы к оркестрации мультиагентных систем для разработки ПО. Ключевым отличием задач такой системы от существующих бенчмарков является необходимость построение ПО с нуля, а не внесение правок в существующие проекты.

Для исследования оркестрации создана тестовая среда, обеспечивающая запуск произвольного количества агентов с различными параметрами. Одним из требований к системе является возможность запускать существующие бенчмарки и оценивать поведение системы на примере бейслайнов [3].

Важным критерием оценивания системы, помимо времени и ресурсов, в данном случае является масштабируемость и связность предлагаемых агентами решений. Предложенные метрики позволяют количественно оценивать влияние увеличения числа агентов на эффективность координации, выявлять деградации согласованности решений и сопоставлять различные стратегии оркестрации в воспроизводимых условиях. Проведённые эксперименты направлены на сравнение базовых и оптимизированных подходов к распределению задач между агентами, а также на анализ зависимости качества итогового программного решения от параметров оркестрации и масштаба системы.

Выводы

В результате работы разработана экспериментальная методика исследования оркестрации мультиагентных систем, ориентированных на совместную разработку программного обеспечения, включающая тестовую среду, воспроизводимые сценарии и набор специализированных метрик оценки масштабируемости и связности решений. Полученные результаты подтверждают, что эффективность мультиагентных систем определяется не только возможностями отдельных агентов, но и качеством их координации, а применение оптимизационных подходов к оркестрации позволяет

повысить устойчивость и результативность коллективного решения задач. Предложенная методика может быть использована для дальнейших исследований и разработки инструментов управления мультиагентными системами на основе языковых моделей.

Литература

1. Mohammadi M., Li Y., Lo J., Yip W. Evaluation and Benchmarking of LLM Agents: A Survey // arXiv preprint. – 2025.
2. Tian A. X., Zhang R., Tang J., Cho Y. M., Li X., Yi Q. et al. Beyond the Strongest LLM: Multi-Turn Multi-Agent Orchestration vs. Single LLMs on Benchmarks // arXiv preprint. – 2025.
3. Nöther J., Singla A., Radanovic G. Benchmarking the Robustness of Agentic Systems to Adversarially-Induced Harms // arXiv preprint. – 2025.
4. Jiang T., Wang Y., Liang J., Wang T. AgentLAB: Benchmarking LLM Agents against Long-Horizon Attacks // arXiv preprint. – 2026.
5. Zhuge M., Zhao C., Ashley D., Wang W., Khizbullin D., Xiong Y. et al. Agent-as-a-Judge: Evaluate Agents with Agents // arXiv preprint. – 2024.