



Высшая школа экономики

Прикладной анализ данных и
искусственный интеллект

Санкт-Петербург, 2025

Адаптивные топологии мульти- агентных LLM систем с *human-in-the-loop*

Дипломная работа Соснина Тимофея. ПАДИИ-4



Адаптивные топологии мульти-агентных LLM систем с human-in-the-loop

- Объект: мульти-агентные LLM системы, где агенты обмениваются сообщениями
- Фокус: как топология связей + роль человека влияют на эффективность решения сложных задач
- Гипотеза: оптимальная топология зависит от типа задачи и должна динамически адаптироваться; интеграция человека требует системного подхода
- Что хочу получить в итоге: рекомендации “какая конфигурация для каких задач” + экспериментальная база



Предметная область: топологии и роли человека

Топология = кто с кем общается и как течет информация

- Статичные топологии: Star, Chain, Mesh, Debate, Hierarchical
- Adaptive: переключение топологий по ходу решения

Human-in-the-loop: роли человека в системе

- Coordinator / Reviewer / Judge / Peer / Monitor



Постановка задачи

Проблема: одна фиксированная схема коммуникации и одна роль человека редко оптимальны для всех фаз сложной задачи

- Сложные задачи имеют фазы: планирование -> выполнение -> проверка
- Идея работы: динамически переключать топологию и адаптировать роль человека под текущую фазу

Исследовательские вопросы:

- Как топологии влияют на качество, стоимость (токены) и время в разных типах задач?
- Дает ли адаптация преимущество над статичными конфигурациями?
- Какая роль человека наиболее эффективна и где ее “ставить” в топологии?



Актуальность: кому и зачем

Почему это важно:

- В реальных LLM-пайплайнах нужно качество при ограничениях по времени и бюджету токенов
- Интеграция человека добавляет стоимость: время, когнитивная нагрузка, количество вмешательств
- Цель: найти конфигурации, которые дают лучший баланс: качество vs токены/стоимость vs время vs нагрузка на человека

Где применимо: AI-assisted инструменты для программирования, анализа, исследований; collaborative AI в индустрии



Новизна и отличие от аналогов

Новизна подхода: объединить в одном исследовании

- сравнение топологий в контролируемых условиях с едиными метриками
- изучение runtime adaptation (переключение топологии во время решения)
- анализ роли/позиции человека в разных топологиях
- совместная оптимизация: топология + адаптация + human-in-the-loop

Отличие от связанных работ (по формулировке из контекста):

- G-Designer: выбор топологии один раз в начале, без адаптации
- Scaling LLM-MAS: фокус на scaling и статичных конфигурациях
- DylAN: меняет состав команды, но не топологию коммуникации



План работы и текущий прогресс

План экспериментов (4 этапа):

1. Baseline: статичные топологии без человека
2. Human-in-the-loop: добавление ролей человека и сравнение с baseline
3. Adaptive: переключение топологий и ролей по фазам
4. Анализ: статистика, паттерны, практические рекомендации

Ожидаемые артефакты:

- benchmark dataset результатов по типам задач
- design guidelines: когда какая топология и как встраивать человека

