

УДК 004.75

МЕТОДЫ МУЛЬТИАГЕНТНОГО ВЗАИМОДЕЙСТВИЯ В РАСПРЕДЕЛЁННЫХ СЕТЯХ ДОСТАВКИ КОНТЕНТА

Арсентьев Георгий Михайлович,

магистрант, Московский государственный технический университет им. Н.Э. Баумана
(МГТУ им. Н.Э. Баумана)

Аннотация

В статье рассматриваются методы мультиагентного взаимодействия между узлами в распределённых сетях доставки контента. Показано, что отсутствие координации между узлами приводит к снижению эффективности сети за счёт дублирования контента и невозможности использования кешей соседей. Анализируются три сложившихся семейства методов координации: эвристические методы, федеративное обучение и мультиагентное обучение с подкреплением. Для каждого семейства рассмотрены принципы работы, основные алгоритмы, сильные и слабые стороны. Отдельно обсуждаются гибридные архитектуры, объединяющие несколько уровней координации. Сделан вывод о доминирующем направлении развития и обозначены перспективные исследовательские задачи.

Ключевые слова: сеть доставки контента, кэширование, мультиагентные системы, федеративное обучение, обучение с подкреплением, координация агентов, распределённые системы.

METHODS OF MULTI-AGENT INTERACTION IN DISTRIBUTED CONTENT DELIVERY NETWORKS

Arsentyev Georgy Mikhailovich,

Master's Student, Bauman Moscow State Technical University (BMSTU)
e-mail: arsentevgm@student.bmstu.ru

ABSTRACT

The paper reviews methods of multi-agent interaction between nodes in distributed content delivery networks. It is shown that the lack of coordination between nodes reduces network efficiency due to content duplication and the inability to reuse neighboring caches. Three established families of coordination methods are analyzed: heuristic protocols for exchanging cache content metadata, federated learning, and multi-agent reinforcement learning. For each family, working principles, key algorithms, strengths, and limitations are considered. Hybrid architectures combining multiple coordination levels are discussed separately. The dominant direction of development is identified, and promising research problems are outlined.

Keywords: content delivery network, caching, multi-agent systems, federated learning, reinforcement learning, agent coordination, distributed systems.

Введение

Современные системы доставки контента представляют собой распределённые сети кеширующих узлов, расположенных близко к конечным пользователям. Узлы принимают на себя основную часть запросов и обращаются к исходному серверу только при отсутствии запрошенного объекта в локальном кеше. Объём кеша каждого узла существенно меньше суммарного объёма файлов источника, поэтому узлы вынуждены постоянно принимать решения, какие объекты сохранять, а какие вытеснять. От качества этих решений зависят и нагрузка на источник, и задержки при обслуживании пользователей.

Если узлы принимают решения исключительно на основании собственных наблюдений, эффективность сети заметно снижается. Соседние узлы со схожими паттернами запросов хранят одни и те же популярные объекты. Совокупный охват исходного каталога сетью оказывается значительно меньше суммы локальных ёмкостей, что особенно ощутимо в области малопопулярного контента. При промахе узел вынужден обращаться к источнику даже в том случае, когда запрошенный объект уже находится в кеше соседа, что создаёт лишнюю нагрузку на восходящие каналы.

Также проблемой является замедленная реакция на изменения популярности объектов. Каждый узел самостоятельно накапливает статистику с нуля, в то же время обмен этой информацией с соседями позволил бы быстрее перестраиваться. Кроме того, в подобной системе невозможна специализация узлов. При наличии координации соседи могли бы неявно разделять каталог между собой, расширяя совокупную ёмкость сети, однако в изолированном режиме у узла нет ни оснований, ни данных для того, чтобы уступить хранение объекта соседу.

Перечисленные ограничения определяют интерес к мультиагентному взаимодействию между узлами как к одному из центральных направлений в исследовании систем доставки контента.

Эвристические методы координации

Исторически первое и до сих пор широко применяемое в индустрии семейство методов основано на фиксированных правилах координации, не требующих обучения. Внутри него выделяются два принципиально различных подхода.

Первый подход состоит в периодической рассылке соседям компактного представления содержимого кеша. Стандартным средством для этого служат фильтры Блума — вероятностные структуры данных, позволяющие проверять принадлежность объекта множеству при очень малом размере представления. Для десяти тысяч объектов с вероятностью ложного срабатывания в один процент достаточно порядка двенадцати килобайт [1]. При промахе узел сначала проверяет фильтры соседей и, если объект потенциально найден, обращается к соседу вместо источника. Этот механизм был реализован в академических прототипах и в ранних версиях прокси-сервера Squid под названием Cache Digests. Главное достоинство подхода состоит в естественной децентрализации, при которой каждый узел самостоятельно решает откуда запрашивать объект. Основные ограничения связаны с тем, что данные о содержимом соседей быстро теряют актуальность между обновлениями, а накладные расходы на широковебательную рассылку плохо масштабируются с ростом числа узлов.

Второй подход устраняет необходимость в обмене информацией о содержимом, заменяя её детерминированным правилом размещения. Каждый объект отображается на конкретный узел внутри группы через хеш-функцию от его идентификатора. Узел,

обрабатывающий запрос, по тому же хешу определяет, какой из соседей хранит объект, и обращается к нему напрямую. Дублирование устраняется в принципе, обмен метаданными становится излишним. Каноническая реализация этого принципа – протокол Cache Array Routing Protocol (CARP) [2]. Его современные разновидности на основе согласованного хеширования обеспечивают устойчивость к изменениям состава группы: добавление или удаление узла требует переноса лишь небольшой доли объектов. На этом подходе построена координация периферийных кластеров в крупных промышленных сетях доставки контента – Akamai, Cloudflare, Fastly. Слабая сторона этого подхода состоит в отсутствии гибкости, так как правило размещения не учитывает реальную популярность объектов.

Главное достоинство всего семейства состоит в предсказуемости поведения и низких накладных расходах. Главный недостаток – отсутствие адаптации к опыту из-за фиксированных правил, которые не способны использовать информацию об эффективности уже принятых решений. В современных гибридных архитектурах эвристический транспортный слой сохраняется как нижний уровень, поверх которого надстраиваются обучающиеся компоненты.

Федеративное обучение

Второе семейство методов организует координацию через обмен параметрами обучаемых моделей. Каждый агент обучает собственную локальную модель, предсказывающую популярность объектов или ценность их хранения, на собственных данных о запросах. Веса моделей периодически агрегируются – централизованно или через одноранговый обмен с соседями. Базовый алгоритм такой агрегации (федеративное усреднение) работает следующим образом. Координатор рассылает участникам общую модель, каждый делает несколько шагов обучения на своих данных и отправляет обратно только изменённые веса. Координатор усредняет полученные обновления и формирует новую общую модель для следующего раунда [3].

Этот подход имеет несколько преимуществ. Во-первых, сохраняется приватность данных: записи о запросах не покидают узел, что важно при регуляторных ограничениях на передачу пользовательских трасс. Во-вторых, сеть приобретает устойчивость к изменениям нагрузки благодаря тому, что локальные модели быстро реагируют на новые паттерны, а усреднённая модель сглаживает случайные колебания. В-третьих, узлы с малой нагрузкой получают возможность пользоваться опытом более загруженных соседей, поскольку у них самих недостаточно данных для обучения качественной модели. Применение этой схемы к упреждающему кешированию на периферийных узлах продемонстрировало значительное снижение нагрузки на источник по сравнению с фиксированными эвристиками [4].

Основная слабость базового алгоритма связана с неоднородностью нагрузок. Агенты, обслуживающие районы с разным профилем спроса, имеют существенно различающиеся распределения запросов. Прямое усреднение их моделей размывает локальную специализацию, и общая модель оказывается компромиссом, уступающим по качеству локальным моделям, обученным на собственных данных каждого узла. Эту проблему решает персонализированное федеративное обучение, при котором каждый узел сохраняет собственную надстройку над общей моделью. Конкретные механизмы могут различаться: предварительное мета-обучение, регуляризация локальной модели к общей или разделение модели на общую часть и индивидуальную «голову» под каждого агента [5]. Другое расширение – иерархическое федеративное обучение, при котором агенты группируются в кластеры с частой синхронизацией внутри кластера и редкой между кластерами [6].

Это направление представляется одним из наиболее перспективных для практических мультиагентных систем кэширования. Оно сочетает естественную децентрализацию, защищённость пользовательских данных и хорошую теоретическую базу, накопленную в смежных областях машинного обучения на распределённых данных.

Мультиагентное обучение с подкреплением

Третье и наиболее активно исследуемое в академической литературе семейство рассматривает каждый узел сети как агента, обучающегося политике принятия решений на основе вознаграждения. Задача формализуется как марковская игра: состояние сети меняется в зависимости от совместных действий всех агентов, и каждый агент стремится максимизировать суммарное вознаграждение, наблюдая лишь часть глобального состояния. Подходы внутри этого семейства различаются степенью координации агентов во время обучения и характером информации, которой они обмениваются.

Простейшая концепция — независимое обучение, при котором каждый агент применяет стандартный одноагентный алгоритм, считая всех остальных частью внешней среды. Этот подход прост в реализации и не требует никакой координационной инфраструктуры. Главная теоретическая проблема состоит в нестационарности среды. Пока один агент обучается, остальные тоже изменяют свои политики, и распределение наблюдаемых переходов между состояниями смещается, что нарушает гарантии сходимости одноагентных алгоритмов. Тем не менее, на практике метод часто работает приемлемо, особенно при слабом взаимодействии между агентами. Конкретные реализации обычно строятся на основе классических одноагентных алгоритмов (например, DQN).

Более продвинутой концепцией — централизованное обучение с децентрализованным исполнением. Идея состоит в том, что во время обучения используется глобальная информация о наблюдениях и действиях всех агентов, а во время работы в реальной системе каждый агент действует только на основании собственных локальных наблюдений. Этот компромисс снимает проблему нестационарности среды на этапе обучения и при этом сохраняет возможность децентрализованного исполнения. Первой реализацией этой парадигмы является алгоритм MADDPG [7], в котором каждый агент имеет собственного актора, действующего локально, и общего критика, видящего полную картину сети. Современный сильный базовый метод для кооперативных задач — MAPPO, адаптация алгоритма PPO к мультиагентному режиму [8]. Главное ограничение парадигмы состоит в том, что сложность централизованного критика растёт с числом агентов, и для крупных сетей такой подход плохо масштабируется.

Для решения проблемы масштабирования используется концепция декомпозиции функции ценности. Вместо одной общей функции на всю сеть её представляют как комбинацию индивидуальных функций ценности отдельных агентов: либо в виде простой суммы, либо как более общую монотонную композицию. Каждый агент обучает свою компоненту, но коэффициенты композиции оптимизируются относительно общего вознаграждения сети, что обеспечивает кооперативное поведение без необходимости поддерживать модель глобального критика. Такая декомпозиция особенно естественна для задачи кэширования, поскольку глобальная стоимость сети складывается из локальных затрат на обслуживание запросов. Конкретные алгоритмы этого класса — VDN с аддитивной декомпозицией и QMIX с монотонной композицией [9]. Их применение к кооперативному кэшированию на периферийных узлах демонстрирует устойчивое преимущество над эвристическими методами и независимым обучением [10].

Методы мультиагентного обучения с подкреплением обладают наибольшей теоретической мощностью среди всех рассмотренных семейств. Они способны находить сложные кооперативные стратегии, которые не выводятся из явных правил и не сводятся к

локальным предсказаниям. Однако их применение сопряжено с рядом серьезных сложностей, таких как подбор гиперпараметров системы и достижение устойчивой сходимости.

Гибридные архитектуры

Анализ современной литературы показывает, что наиболее успешные системы строятся как гибриды, объединяющие несколько описанных уровней. Типичная архитектура содержит нижний слой обмена метаданными о владении объектами, средний слой локальной обучаемой модели ценности, верхний слой федеративного обмена параметрами моделей между узлами, а также кооперативную функцию вознаграждения, явно учитывающую глобальные интересы сети [11].

Такая структура отражает понимание, что ни одно семейство методов в одиночку не покрывает всех требований. Эвристические протоколы не адаптируются к меняющимся паттернам нагрузки. Мультиагентное обучение с подкреплением медленно сходится и нестабильно в больших системах. Федеративное обучение не использует специфическую структуру задачи кэширования и не координирует решения на коротких временных масштабах. Гибридная архитектура позволяет каждому слою решать подзадачи, для которых он наиболее применим. Обмен метаданными обеспечивает актуальную картину сети, локальная модель отвечает за быстрые решения о вытеснении, федеративный уровень обеспечивает обмен опытом и долгосрочную адаптацию, кооперативная функция выполняет согласование локальных и глобальных интересов.

Заключение

Отсутствие координации между узлами сетей доставки контента порождает четыре фундаментальных эффекта: дублирование контента, невозможность переиспользования соседских кешей, замедленную адаптацию к изменениям нагрузки и отсутствие специализации узлов. Совокупность этих эффектов делает мультиагентную координацию необходимым компонентом эффективной системы.

Современные методы мультиагентного взаимодействия можно разделить на три группы. Эвристические протоколы просты и предсказуемы, но не обучаются на опыте. Федеративное обучение сочетает децентрализацию, защищенность пользовательских данных и адаптивность, однако чувствительно к неоднородности распределений запросов на разных узлах. Мультиагентное обучение с подкреплением обладает наибольшей теоретической мощностью при наибольшей сложности реализации.

Доминирующим направлением развития являются гибридные архитектуры, объединяющие эвристический транспортный слой, локальную обучаемую модель ценности, федеративный обмен опытом и кооперативную функцию вознаграждения.

Список литературы:

1. Fan L., Cao P., Almeida J., Broder A. Z. Summary cache: a scalable wide-area Web cache sharing protocol // IEEE/ACM Transactions on Networking. 2000. Vol. 8, № 3. P. 281–293.
2. Valloppillil V., Ross K. W. Cache Array Routing Protocol v1.0 // Internet-Draft, IETF. February 1998. URL: <https://datatracker.ietf.org/doc/html/draft-vinod-carp-v1-03> (дата обращения: 29.03.2026).
3. McMahan H. B., Moore E., Ramage D., Hampson S., Arcas B. A. Communication-efficient learning of deep networks from decentralized data // Proceedings of AISTATS 2017. P. 1273–1282.
4. Yu Z., Hu J., Min G., Lu H., Zhao Z., Wang H., Georgalas N. Federated learning based proactive content caching in edge computing // Proceedings of IEEE GLOBECOM 2018. P. 1–6.

5. Tan A. Z., Yu H., Cui L., Yang Q. Towards personalized federated learning // IEEE Transactions on Neural Networks and Learning Systems. 2023. Vol. 34, № 12. P. 9587–9603.
6. Liu L., Zhang J., Song S. H., Letaief K. B. Client-edge-cloud hierarchical federated learning // Proceedings of IEEE ICC 2020. P. 1–6.
7. Lowe R., Wu Y., Tamar A., Harb J., Abbeel P., Mordatch I. Multi-agent actor-critic for mixed cooperative-competitive environments // Advances in Neural Information Processing Systems 30 (NeurIPS 2017). P. 6379–6390.
8. Yu C., Velu A., Vinitsky E., Gao J., Wang Y., Bayen A., Wu Y. The surprising effectiveness of PPO in cooperative multi-agent games // Advances in Neural Information Processing Systems 35 (NeurIPS 2022). P. 24611–24624.
9. Rashid T., Samvelyan M., de Witt C. S., Farquhar G., Foerster J., Whiteson S. QMIX: monotonic value function factorisation for deep multi-agent reinforcement learning // Proceedings of ICML 2018. P. 4292–4301.
10. Zhong C., Gursoy M. C., Velipasalar S. Deep multi-agent reinforcement learning based cooperative edge caching in wireless networks // Proceedings of IEEE ICC 2019. P. 1–6.
11. Wang X., Wang C., Li X., Leung V. C. M., Taleb T. Federated deep reinforcement learning for Internet of Things with decentralized cooperative edge caching // IEEE Internet of Things Journal. 2020. Vol. 7, № 10. P. 9441–9455.

References:

1. Fan L., Cao P., Almeida J., Broder A. Z. Summary cache: a scalable wide-area Web cache sharing protocol. IEEE/ACM Transactions on Networking, 2000, vol. 8, no. 3, pp. 281–293.
2. Valloppillil V, Ross KW. Cache Array Routing Protocol v1.0. Internet-Draft, IETF, February 1998. Available at: <https://datatracker.ietf.org/doc/html/draft-vinod-carp-v1-03>. Accessed March 29, 2026.
3. McMahan HB, Moore E, Ramage D, Hampson S, Arcas BA. Communication-efficient learning of deep networks from decentralized data. In: Proceedings of AISTATS 2017. 2017. pp. 1273–1282.
4. Yu Z, Hu J, Min G, Lu H, Zhao Z, Wang H, Georgalas N. Federated learning based proactive content caching in edge computing. In: Proceedings of IEEE GLOBECOM 2018. 2018. pp. 1–6.
5. Tan A. Z., Yu H., Cui L., Yang Q. Towards personalized federated learning. IEEE Transactions on Neural Networks and Learning Systems, 2023, vol. 34, no. 12, pp. 9587–9603.
6. Liu L, Zhang J, Song SH, Letaief KB. Client-edge-cloud hierarchical federated learning. In: Proceedings of IEEE ICC 2020. 2020. pp. 1–6.
7. Lowe R, Wu Y, Tamar A, Harb J, Abbeel P, Mordatch I. Multi-agent actor-critic for mixed cooperative-competitive environments. In: Advances in Neural Information Processing Systems 30 (NeurIPS 2017). 2017. pp. 6379–6390.
8. Yu C, Velu A, Vinitsky E, Gao J, Wang Y, Bayen A, Wu Y. The surprising effectiveness of PPO in cooperative multi-agent games. In: Advances in Neural Information Processing Systems 35 (NeurIPS 2022). 2022. pp. 24611–24624.

9. Rashid T, Samvelyan M, de Witt CS, Farquhar G, Foerster J, Whiteson S. QMIX: monotonic value function factorisation for deep multi-agent reinforcement learning. In: Proceedings of ICML 2018. 2018. pp. 4292–4301.
10. Zhong C, Gursoy MC, Velipasalar S. Deep multi-agent reinforcement learning based cooperative edge caching in wireless networks. In: Proceedings of IEEE ICC 2019. 2019. pp. 1–6.
11. Wang X., Wang C., Li X., Leung V. C. M., Taleb T. Federated deep reinforcement learning for Internet of Things with decentralized cooperative edge caching. IEEE Internet of Things Journal, 2020, vol. 7, no. 10, pp. 9441–9455.