

Д. А. МУРАШОВ
Санкт-Петербургский Федеральный исследовательский центр Российской академии наук,
Санкт-Петербург

КООРДИНАЦИЯ ПРОЦЕССА БАЛАНСИРОВКИ НАГРУЗКИ НА ОСНОВЕ МУЛЬТИАГЕНТНОГО ПОДХОДА

В настоящее время существует значительное количество алгоритмов на основе мультиагентного подхода, позволяющих решать задачу балансировки нагрузки в вычислительных сетях. Однако большинство данных алгоритмов покрывает один или несколько аспектов балансировки и не позволяет выполнять координацию процесса балансировки извне.

В работе предлагается подход, позволяющий выполнять балансировку индифферентно к топологии вычислительной сети и осуществлять управление процессом балансировки, созданный на основе комбинации алгоритма балансировки нагрузки методом эстафеты и метода анализа иерархий. Приводятся результаты эксперимента, проведенного с использованием имитационной модели вычислительной сети, реализующей подход.

Введение. Алгоритмы балансировки нагрузки в вычислительных сетях на основе мультиагентного подхода часто позволяют решать задачу балансировки нагрузки более эффективно, чем модели, основанные на централизованном управлении [1]. «Ситуационная осведомленность» отдельных агентов позволяет в деталях учитывать состояние отдельных участков вычислительного кластера.

Тем не менее, значительная доля алгоритмов данного класса имеет общий недостаток. Мультиагентная архитектура подразумевает, что сущности-участники процесса являются автономными и действуют в ответ на изменение состояния внешней среды, образуемого, в том числе, действиями других агентов, исходя из опыта, накопленного в результате меж-агентных взаимодействий, или знаний, изначально заложенных в их модель принятия решений. Автономность сущностей затрудняет управление роем агентов; с другой стороны, наличие управления часто означает значительное ограничение свободы действий агента, за счет чего теряются достоинства, присущие мультиагентной архитектуре и моделям.

Существует потребность в подходах, которые позволяют сохранить преимущества мультиагентных моделей, такие как использование ситуационной осведомленности агентов при принятии решений, экономия ресурсов за счет распределения принятия решений и эмерджентное интеллектуальное поведение, но в то же время осуществлять контроль над процессом балансировки.

Предлагаемый доклад посвящен подходу, позволяющему решать данную задачу. В докладе приводится изложение разработанного алгоритма. Для апробации алгоритма была создана имитационная модель вычислительной сети, с помощью которой решалась проблема балансировки нагрузки. Был проведен вычислительный эксперимент с использованием данной модели. Приводятся результаты эксперимента.

Полимодельный подход к решению задачи балансировки нагрузки. Предлагаемый подход основан на комбинации двух моделей: алгоритма балансировки нагрузки на основе эстафеты и метода анализа иерархий.

Основополагающая идея разработанного подхода состоит в использовании особенностей структуры графа предпочтений, применяемого в методе анализа иерархий [2]. Данный граф разбивается на «уровни» (минимально возможное количество переходов по дугам дерева в пути от корневой вершины к вершине, принадлежащей уровню). Это разделение позволяет произвести декомпозицию подцелей (аспектов) процесса балансировки нагрузки. Результат декомпозиции позволяет содержательно интерпретировать полученные цели как «высокоуровневые» и «низкоуровневые».

Листья графа соответствуют альтернативам – конечному множеству, среди которого осуществляется выбор. На выбор альтернативы влияют значения весов, присвоенных дугам, соединяющих корневую вершину, промежуточные вершины («подцели») и листья.

Влияние веса дуги пропорционально близости дуги к корневой вершине. Эта особенность делает возможным управлять процессом принятия решений на «верхнем» (стратегическом) уровне, изменяя веса дуг, соединяющих корневую вершину и вершины первого уровня. В работе предлагается архитектурный подход, использующий разделение агентов на 2 типа: координатор (управляющая сущность в единственном экземпляре) и исполнители (узлы сети, реализующие процесс принятия решений). Координатор осуществляет управление сетью, исходя из текущих стратегических задач, изменяя верхние уровни дерева предпочтений. Исполнители, имея возможность изменять нижние уровни, принимают решения.

Вторая основополагающая идея заимствована из подхода к балансировке нагрузки «методом эстафеты» [3]. Согласно данному подходу, агенты (узлы сети) передают входящие задачи наименее загруженному узлу среди соседей. Цепочка передач продолжается в течение ограниченного количества шагов. В результате, в сети обеспечивается равномерное распределение задач вне зависимости от топологии.

Обобщая – формируется следующая последовательность. Координатор изменяет верхние уровни графа. Исполнитель, наблюдая за непосредственным окружением в кластере, изменяет нижние уровни. В момент поступления задачи исполнителю последний до-формирует граф, добавляя в него листья (альтернативы), каждый из которых соответствует соседней вершине, кроме одного, который соответствует самому исполнителю, принимающему решение. Выполняется рекурсивная вычислительная процедура:

$$W_{(N)}^{(a,N-2)} = \text{АНР} \left(\left\{ W_{(N)}^{(i,N-1)} \right\}_i, W_{(N-1)}^{(v,N-2)} \right), \quad (1)$$

где N – уровень, на котором расположены альтернативы (считая от корневой вершины, соответствующей уровню 0), $W_H^{(a,G)}$ – вектор, содержащий оценки предпочтительности всех вершин уровня H в контексте вершины a , расположенной на уровне G , АНР(...) – стандартная вычислительная процедура получения оценок альтернатив, используемая в методу анализа иерархий [2]. Целью рекурсивной процедуры является формирование вектора $W_{(N)}^{(0)}$, т.е. оценки альтернатив в контексте корневой вершины (индекс вершины опущен, потому что существует только одна корневая вершина).

По своей сути, рекурсивная процедура представляет собой последовательное применение стереотипной вычислительной процедуры из метода анализа иерархий, описанной для трехуровневого графа предпочтений, но в случае графа с большим количеством уровней.

Интеграция вышеописанной интерпретации метода анализа иерархий с разделением графа предпочтений и метода эстафеты позволяет принимать решение более гибко: учитывать задачи, стоящие перед процессом балансировки нагрузки, в момент передачи поступающей задачи между соседними узлами сети. Разделение задач на стратегические и тактические позволяет также учитывать локальные состояния сети перед принятием решения. Достоинство метода анализа иерархий состоит в том, что он не накладывает ограничений на реализацию процесса принятия решений агентом.

Вычислительный эксперимент. Для проведения вычислительного эксперимента была создана модель гетерогенной вычислительной сети с изменяющейся топологией. Узлам сети и связям между ними назначались различные величины производительности и пропускной способности соответственно. Изменчивость топологии моделировалась случайным отключением и подключением узлов сети с различными интервалами.

В сети моделировалась передача вычислительной задачи между узлами сети с применением вышеописанного полимодельного подхода.

По мере работы имитационной модели осуществлялся сбор статистики, состоящей из усредненных технических показателей вычислительного кластера.

Было проведено несколько инстанций эксперимента. По результатам эксперимента были построены графики, демонстрирующие зависимости между установленными стратегическими целями процесса балансировки и техническим состоянием вычислительного кластера.

Заключение. Среди возможных способов дальнейшего развития подхода, предложенного в настоящем докладе, следует выделить поиск технических средств, позволяющих реализовать предложенный подход на практике.

Исследования, выполненные по данной тематике, проводились при частичной финансовой поддержке грантов РФФИ (№№20-08-01046), в рамках бюджетной темы FFZF-2022-0004.

ЛИТЕРАТУРА

1. **Biswas T., Kula P., Ray A.K.** A novel scheduling with multi-criteria for high-performance computing systems: an improved genetic algorithm-based approach. *Engineering with Computers*, 35(4), 2019. С.1475–1490.
2. **Saaty Thomas L.** "What is the analytic hierarchy process?" *Mathematical models for decision support*. Springer, Berlin, Heidelberg, 1988. 109–121.
3. **Gorodetskii V.** Self-organization and multiagent systems: I. models of multiagent self-organization. *Journal of Computer and Systems Sciences International* 51(2), 256–281 (2012)

D.A.Murashov, (St. Petersburg Federal Research Center of the Russian Academy of Sciences, St. Petersburg)

Agent-based coordination of load balancing processes

There exist a significant amount of agent-based algorithms capable of solving the problem of load balancing in computational networks. However, a significant part of them only covers one or few narrow aspects of the process, and do not provide facilities for controlling the process from outside. In this paper, an approach to load balancing based on combination of analytic hierarchy process with relay-race load balancing algorithm was proposed. It enables control over the process of load balancing, while being indifferent to the network's topology. The paper offers a description of the approach and experimental results acquired from a computational experiment using simulation of a heterogeneous network employing the algorithm.