

**РАЗРАБОТКА МАТЕМАТИЧЕСКОГО И АЛГОРИТМИЧЕСКОГО ОБЕСПЕЧЕНИЯ
ИНТЕЛЛЕКТУАЛЬНОГО ПРОЕКТИРОВАНИЯ АППАРАТНО-ПРОГРАММНЫХ
КОМПЛЕКСОВ ОБРАБОТКИ ИНФОРМАЦИИ В РАСПРЕДЕЛЕННЫХ
ВЫСОКОПРОИЗВОДИТЕЛЬНЫХ СИСТЕМАХ***

Рассматриваются вопросы выбора эффективных структур и состава высокопроизводительных аппаратно-программных комплексов обработки информации. Проводится сравнительный анализ централизованного и распределенного подходов реализации таких комплексов. Приводятся результаты патентного поиска, отражающего современные тенденции и мировой уровень в данной предметной области.

Ключевые слова: многопроцессорные системы, распределенные вычисления, многоагентные алгоритмы.

В последние годы четко сформировалась определенная тенденция развития средств вычислительной техники: рост производительности компьютеров происходит в основном не за счет увеличения производительности единственного вычислительного узла (центрального процессора), а за счет распараллеливания вычислительного процесса по нескольким узлам компьютера или даже по компьютерной сети. Такой подход не содержит принципиальных ограничений в повышении скорости и точности работы систем обработки информации. Однако одновременно с ростом производительности вычислительных систем растет и сложность решаемых с их помощью задач. Особенно актуальными сейчас, в свете повсеместного распространения компьютерных сетей и Интернета, являются задачи обработки и анализа информации в распределенных системах. Разнообразие задач обработки и анализа данных не позволяет говорить о какой-либо универсальной аппаратно-программной реализации вычислительного комплекса, которому были бы по силам все задачи. Таким образом, массовое использование технологий анализа данных сдерживается трудностями разработки вычислительных комплексов, предназначенных для решения конкретных задач, что во многом объясняется отсутствием математического аппарата моделирования функционирования специализированных распределенных вычислительных систем.

Можно утверждать, что сегодня в области больших вычислений конкурируют две идеи. Первая идея предполагает использование централизованных вычислительных комплексов (суперкомпьютеров), вторая – использование распределенных вычислительных систем. Особенностью распределенных многопроцессорных вычислительных систем, в отличие от локальных суперкомпьютеров, является возможность неограниченного наращивания производительности за счет масштабирования. Так, проект BOINC (Berkeley Open Infrastructure for Network Computing), являю-

щийся примером реализации распределенных вычислений, обладает средней производительностью 5 000 TeraFLOPS, что сравнимо с производительностью суперкомпьютера, возглавляющего рейтинг TOP500 (8 162 TeraFLOPS).

По доступности для использования суперкомпьютеры также можно сравнивать с распределенными вычислительными системами. Если для суперкомпьютеров естественным ограничением является их стоимость, то применение распределенных вычислений связано лишь с организационными и алгоритмическими сложностями.

Проведен патентный поиск и выполнен аналитический обзор данной предметной области. Исследования позволили сделать следующие выводы:

– при проектировании распределенных систем и алгоритмов используются в основном подходы к повышению производительности вычислений частного запроса (выбор вычислительного элемента (процессора), оптимизация нагрузки на систему, структура контролера и т. д.). При этом не оптимизируется архитектура комплекса в целом, а, как известно, транзакционные издержки и режим работы главного сервера могут замедлять систему в целом даже при сверхпроизводительных локальных серверах;

– разработчики не используют распараллеливание задач на уровне алгоритмов (островные модели, кооперационные и конкурирующие схемы, комитеты решателей), которые существенно повышают эффективность решения задач даже на последовательных компьютерах;

– комплексные решения предлагаются крупными разработчиками (IBM, Deutsche Telecom, Intel и др.) и в общем случае являются системами, ориентированными исключительно на специфику, технологии и технические решения разработчика, что существенно снижает область применения;

– проблема обработки слабо структурированной информации практически не рассматривается;

* Работа выполнена при финансовой поддержке ФЦП «Исследования и разработки по приоритетным направлениям развития научно-технологического комплекса России» на 2007–2013 гг. НИР 2011-1.9-519-005-042 и ФЦП «Научные и научно-педагогические кадры России» НИР 2011-1.2.1-113-025, 2011-1.2.2-215-021.

– говоря об интеллектуальных информационных системах, разработчики чаще всего имеют в виду использование разных эвристик в предметной области. При этом методы обработки информации могут быть традиционными, не связанными с интеллектуальными информационными технологиями. Семантика решаемых задач не используется;

– частные работы по применению методов и алгоритмов интеллектуальных информационных технологий не ориентированы на широкое распространение, так как требуют привлечения специалистов и предметных экспертов как на этапе проектирования интеллектуальных технологий, так и при их настройке и использовании.

Таким образом, очевидно, что задача проектирования аппаратно-программных комплексов обработки информации в распределенных системах является актуальной.

В [1] проводится обзор существующих математических моделей многопроцессорных систем, рассматриваются различные подходы к созданию таких моделей. Подробно исследуется модель, где процесс функционирования многопроцессорных систем представляется замкнутой системой массового обслуживания. Предполагается, что состав и характеристики узлов вычислительного комплекса напрямую зависят от специфики решаемой задачи. Так, для реализации наиболее часто применяемых алгоритмов и операций используются специализированные процессоры, с индивидуальными характеристиками. Приводятся алгоритмы расчета производительности, надежности и стоимости проектируемых комплексов. Таким образом, в [1] задача проектирования эффективного вычислительного комплекса формулируется в виде задачи оптимизации:

$$\begin{aligned} P_b(n, m_1, \dots, m_i, \dots, m_N, T_{01}, \dots, T_{0i}, \dots, T_{0N}) &\rightarrow \max, \\ H_b(n, m_1, \dots, m_i, \dots, m_N, T_{01}, \dots, T_{0i}, \dots, T_{0N}) &\rightarrow \max, \\ C_b(n, m_1, \dots, m_i, \dots, m_N, T_{01}, \dots, T_{0i}, \dots, T_{0N}) &\rightarrow \min. \end{aligned}$$

В данной модели приняты следующие обозначения: P_b – критерий оценки производительности; H_b – критерий оценки надежности; C_b – критерий оценки стоимости; n – количество шин; m_i – количество процессоров i -го типа, $1 \leq m_i \leq 16$, $i = \overline{1, N}$; T_{0i} – среднее время выполнения одной команды в процессоре i -го типа, $T_{0i} > 0$, $i = \overline{1, N}$; определяется исходя из реализуемой в специализированном процессоре функции.

В реальных задачах количество типов процессоров может достигать нескольких десятков, а количество специализированных процессоров каждого типа – многих десятков. Допустим, что имеется десять типов спецпроцессоров и возможно включение в систему до 20 спецпроцессоров каждого типа. Тогда при использовании рассматриваемой в [1] аналитической модели имеется около 20^{10} возможных комбинаций. Если осуществлять поиск оптимальной конфигурации не только по структуре вычислительного комплекса, но и по быстродействию спецпроцессоров, то общее число вариантов системы возрастет примерно до 10^{28} . Кроме того, при выборе метода решения данной зада-

чи необходимо учитывать следующие факторы: переменные задачи выражены в различных шкалах измерения (целочисленные и вещественные переменные), оптимизация производится по нескольким экстремальным критериям одновременно, априорные сведения о свойствах целевых функционалов отсутствуют (оптимизация производится только по измерениям функционалов в фиксированных точках). Метод оптимизации должен учитывать условия, накладываемые на критерии. Таким образом, при выборе эффективной конфигурации вычислительного комплекса возникает необходимость решения задачи многокритериальной условной оптимизации с алгоритмически заданными функциями разношкальных переменных высокой размерности.

В [2] процесс функционирования распределенной вычислительной системы (ГРИД-системы) представлен в виде системы массового обслуживания. Приводятся модели расчета производительности и надежности ГРИД-системы. Процесс проектирования эффективной конфигурации ГРИД-системы также сводится к решению многокритериальной задачи условной оптимизации.

Всеми необходимыми качествами для решения таких задач обладают многоагентные методы стохастической оптимизации, имитирующие естественный отбор и поведение социальных групп в природе. К таким методам относятся генетический алгоритм, алгоритм генетического программирования, алгоритм оптимизации роем частиц (PSO). Однако задача эффективного применения данных методов сравнима по сложности с исходной задачей проектирования вычислительных комплексов, требуется разработка и исследование методов автоматического проектирования и настройки интеллектуальных информационных технологий. Данная проблема может быть решена с использованием самонастраивающихся многоагентных стохастических алгоритмов, включая коэволюционные схемы и вероятностные алгоритмы [3; 4].

Одним из примеров программы-агента для задач однокритериальной условной и безусловной оптимизации с алгоритмически заданными целевыми функциями вещественных, целочисленных, бинарных или смешанных переменных, программы, обладающей значительно меньшим, чем стандартный генетический алгоритм, числом настраиваемых параметров, является так называемый коэволюционный генетический алгоритм [5].

Основная проблема использования эволюционных процедур – высокая сложность настройки параметров алгоритма для решения конкретной задачи оптимизации. Поэтому использование эволюционных процедур, в том числе генетических алгоритмов, требует очень высокой квалификации лица, принимающего решение (ЛПР), как в области данных алгоритмов, так и в проблемной области. Нужны также значительные временные затраты. Одним из перспективных подходов для решения указанной проблемы является коэволюционный алгоритм. Ранее были разработаны коэволюционные алгоритмы, решающие задачи безус-

ловной оптимизации [4]. Однако большинство практических задач относится к классу задач условной оптимизации, т. е. имеются ограничения в виде неравенств или равенств.

Козволюционный генетический алгоритм – это несколько параллельно действующих стандартных генетических алгоритмов. Отдельный генетический алгоритм называется подпопуляцией. Каждая из подпопуляций оптимизирует заданную функцию и обладает своей стратегией оптимизации. При этом популяции «борются» за ресурс, который в течение работы алгоритма перераспределяется в пользу более эффективной из них. Стандартный козволюционный алгоритм состоит из следующих этапов [4]:

- выбора индивидуальных алгоритмов;
- задания параметров козволюционного алгоритма (размера общего ресурса, величины интервала адаптации, размера штрафа «проигравшего» алгоритма, размера «социальной карты») для всех индивидуальных алгоритмов;
- работы выбранных алгоритмов в течение интервала адаптации;
- оценки алгоритмов;
- перераспределения ресурсов;
- миграции лучших индивидов во все подпопуляции.

Перераспределение ресурсов позволяет «доминировать» алгоритму, обладающему наилучшими настройками для оптимизации имеющейся целевой функции. Миграция позволяет одновременно использовать преимущества различных алгоритмов – в первую очередь, локальную и глобальную сходимости. Изменение размеров ресурсов происходит путем сокращения каждой проигравшей популяции на некоторый процент (определенный заранее размер штрафа) и увеличением победившей популяции на число, равное сумме потерь проигравших. Иными словами, после каждого интервала адаптации определяется один победивший алгоритм (подпопуляция). При этом размер каждой подпопуляции не может быть меньше определенного уровня, называемого «социальной картой».

В [5] предлагается новый метод перераспределения ресурсов между алгоритмами, получивший название «турнирный». Суть метода заключается в том, что после каждого интервала адаптации не определяется единственный победивший алгоритм, а происходит перераспределение ресурсов в каждой паре подпопуляций («турнир каждого с каждым»). Данный метод перераспределения ресурсов позволяет провести ранжирование подпопуляций по эффективности оптимизации для текущей целевой функции. Результат такого ранжирования – перераспределение ресурсов между всеми подпопуляциями в соответствии с их эффективностью. Турнирный метод перераспределения ресурсов тестировался в сравнении со стандартным на десяти тестовых функциях (задачах безусловной оптимизации). В качестве показателя эффективности работы алгоритма использовалась надежность – отношение числа запусков, в которых с заданной точностью был найден оптимум, к общему числу запусков. По ранговому критерию Вилкоксона турнирный

метод оказался значительно лучше стандартного на четырех тестовых функциях и не хуже на оставшихся шести, что свидетельствует о целесообразности предложенного подхода.

Для адаптации генетического алгоритма к решению задач условной оптимизации необходимо ввести в процедуру специальные методы учета ограничений, используемые в ГА. В данной работе рассматривались следующие: метод «смертельных» штрафов, «лечение» локальным поиском, метод статических штрафов, метод динамических штрафов, метод адаптивных штрафов.

Для апробации предложенного алгоритма была решена задача проектирования распределенной вычислительной сети для решения задачи прогнозирования объемов продаж товаров в аптеках ГПКК «Губернские аптеки» [6]. Данные по задаче следующие:

- средняя вычислительная сложность одной итерации вычислений на одной ветви алгоритма решения задачи – 10 000 оп.;
- средняя вычислительная сложность управления одной ветвью алгоритма решения задачи – 5 000 оп.;
- средний объем данных клиент-серверного обмена – 200 КБ.

Для построения GRID-системы имеется в наличии 153 клиентских узла, параметры которых варьируются в следующих пределах:

- производительность от 4 200 до 17 345 MFLOPS;
- скорость каналов передачи данных от 10 до 27 МБит/с;
- стоимость аренды за 1 ч использования от 1,03 руб. до 3,82 руб.

По надежности все клиентские узлы разбиваются на две группы с интенсивностями отказов 0,000 01 и 0,000 001.

Также задано расписание работы всех клиентских узлов.

Выбор серверного узла проектируемой GRID-системы не производится по причине того, что предприятие-заказчик предоставляет свой серверный узел со следующими параметрами:

- количество процессоров – 2;
- производительность каждого процессора сервера – 11 500 MFLOPS;
- стоимость аренды – 0;
- интенсивность отказов процессоров сервера – 0,000 001.

На проектируемую GRID-систему наложены следующие ограничения:

- количество клиентских узлов GRID-системы должно быть не менее 15 и не более 20;
- коэффициент готовности GRID-системы не менее 99,9 %;
- минимальная производительность GRID-системы не менее 50 GFLOPS.

При решении задачи была получена аппроксимация парето-множества, состоящая из шести точек, а также фронта Парето.

По значениям стоимости и производительности полученных конфигураций GRID-систем был выбран вариант состава проектируемой системы (см. таблицу).

Состав выбранной конфигурации GRID-системы

Производительность, MFLOPS	Скорость канала данных, МБит/с	Стоимость 1 часа аренды, руб.	Время работы		Интенсивность отказов
			Начало работы	Конец работы	
5 295	12	1,24	8:00	18:00	0,000 01
7 600	16	1,64	8:00	19:00	0,000 01
6 420	21	1,49	круглосуточно		0,000 01
7 020	15	1,68	22:00	9:00	0,000 01
8 410	19	1,95	12:00	19:00	0,000 01
10 250	25	2,35	9:00	21:00	0,000 001
9 330	23	1,93	9:00	21:00	0,000 01
8 465	12	1,87	8:00	18:00	0,000 001
6 930	15	1,65	22:00	9:00	0,000 01
11 270	18	2,38	8:00	19:00	0,000 001
9 430	15	2,04	9:00	19:00	0,000 001
7 515	20	1,60	круглосуточно		0,000 01
10 650	24	2,21	12:00	19:00	0,000 001
8 800	19	1,90	8:00	19:00	0,000 01
6 900	11	1,53	9:00	21:00	0,000 01
6 750	21	1,42	круглосуточно		0,000 01
11 000	27	2,36	19:00	8:00	0,000 01
11 000	27	2,36	19:00	8:00	0,000 01
11 000	27	2,36	19:00	8:00	0,000 01
17 345	24	3,64	0:00	6:00	0,000 001

Построенная конфигурация GRID-системы обладает следующими характеристиками:

- средняя производительность – 81,522 GFLOPS;
- стоимость – 466,71 руб. в сутки;
- коэффициент готовности – 99,92 %;
- минимальная производительность – 51 GFLOPS.

Таким образом, полученный в ходе решения задачи вариант распределенной вычислительной системы полностью удовлетворяет исходным ограничениям.

Несмотря на растущую популярность и множество успешных проектов по реализации распределенных вычислений, отсутствует формальное обоснование эффективности применяемых решений. Часто для успешного решения вычислительных задач распределенными или централизованными многопроцессорными системами используются аппаратные средства, заведомо превышающие необходимый уровень. Это многократно повышает стоимость таких проектов и в целом снижает их эффективность. Описываемое в данной статье исследование направлено на повышение обоснованности применения распределенных вычислительных систем. Разработка математического и алгоритмического обеспечения интеллектуального проектирования аппаратно-программных комплексов обработки информации в распределенных высокопроизводительных системах позволит значительно повысить эффективность их работы и расширить сферу их применения.

Библиографические ссылки

1. Панфилов И. А. Модели и алгоритмы выбора эффективной конфигурации многопроцессорных систем обработки информации и управления : автореф. дис. ... канд. техн. наук. Красноярск, 2006.
2. Автоматизация выбора эффективной конфигурации грид-систем / С. Н. Ефимов, Е. С. Семенкин, В. В. Тынченко, В. С. Тынченко // Программные продукты и системы. 2009. № 1. С. 46.
3. Об одной модификации вероятностного генетического алгоритма для решения сложных задач условной оптимизации / А. Ю. Ворожейкин, Т. Н. Гончар, И. А. Панфилов, Е. А. Сопов, С. А. Сопов // Вестник СибГАУ. 2009. № 4. С. 79–84.
4. Жукова М. Н. Коэволюционный алгоритм решения сложных задач оптимизации : автореф. дис. ... канд. техн. наук. Красноярск, 2004.
5. Семенкин Е. С., Семенкин Е. С. Коэволюционный генетический алгоритм решения сложных задач условной оптимизации // Вестник СибГАУ. 2010. № 2 (23). С. 17–21.
6. Ефимов С. Н., Тынченко В. В. Формирование GRID-системы эффективной архитектуры для распределенного решения сложных задач // Системы управления и информ. технологии. 2009. № 2(36). С. 4–8.

I. A. Panfilov, N. A. Smirnov, V. A. Terskov

DEVELOPMENT OF MODELS AND ALGORITHMS OF INTELLIGENT DESIGN OF SOFTWARE AND HARDWARE COMPLEXES OF INFORMATION SYSTEMS PROCESSING IN DISTRIBUTED HIGH-PERFORMANCE SYSTEMS

The authors consider approaches to selection of effective structures and composition of high-performance hard- and software systems for information processing. A comparative analysis of centralized and distributed approach to implementation of such systems is performed. The results of patent search, which reflects the contemporary trends and global level in this object domain, are presented.

Keywords: multiprocessor systems, distributed computing, multi-agent algorithms.

© Панфилов И. А., Смирнов Н. А., Терсков В. А., 2011

УДК 519.866

П. Н. Победаш, О. Э. Семенкина, С. И. Сенашов

АГРЕГИРОВАННАЯ МОДЕЛЬ РЕАЛЬНЫХ ИНВЕСТИЦИЙ С ЧАСТИЧНО НЕОПРЕДЕЛЕННЫМИ СПРОСОМ И ФОНДОУДАЧЕЙ ПРОИЗВОДСТВЕННЫХ АКТИВОВ*

Предложено приложение z-преобразования к оценке эффективности инвестиционных проектов в условиях частичной неопределенности спроса на производимую продукцию и фондоудачи основных производственных фондов.

Ключевые слова: многошаговая задача линейного программирования, z-преобразование.

Рассмотрим задачу, являющуюся, с одной стороны, обобщением задачи из работы [1] на случай двух критериев, а с другой – ее частным вариантом, когда спрос на производимую продукцию или максимальную фондоудачу по некоторым видам производственных активов неизвестен, т. е. когда статистические данные об этих важнейших рыночных характеристиках инвестиционного проекта (ИП) отсутствуют, являются недостоверными или неполными (в силу инновационности проекта, коммерческой тайны, намеренной дезинформации и т. п.).

Сформулируем задачу следующим образом. Предприятие имеет капитал K_0 . При этом государственный орган (ГО) для реализации ИП выделяет инвестиции не более величины I_0 на приобретение основных производственных фондов (ОПФ) n видов. Спрос на производимую продукцию и максимальная фондоудача заданы лишь на первые n_1 и n_2 видов ОПФ соответственно. Необходимо найти стоимость (количество) всех приобретаемых в моменты $t = 1, \dots, T$ ОПФ каждого вида, при которых дисконтированные суммы собственных средств предприятия и его налоговых поступлений в ГО являются наибольшими за все время T действия ИП. При этом предполагаются выполненными следующие основные предпосылки:

1) учитываются налоги, составляющие большую часть затрат предприятия: налог на добавленную стоимость (НДС), налог на прибыль (НП), налог на имущество (НИ), единый социальный налог (ЕСН) и отчисления в фонд оплаты труда (ФОТ);

2) предприятие имеет достаточные запасы сырья;

3) срок T действия ИП меньше сроков T_k службы единицы ОПФ каждого типа: $T < T_k$ ($k = 1, \dots, n$);

4) на ОПФ каждого вида производится лишь один тип продукции.

С учетом перечисленных предпосылок сформулированная выше задача описывается двухкритериальной многошаговой задачей линейного программирования (МЗЛП), которую обозначим как модель $A(n_1, n_2)$:

$$x_k(t+1) = x_k(t) + u_k(t) \quad (k = 1, \dots, n; t = 0, \dots, T-1),$$

$$x_{n+1}(t+1) = -\sum_{k=1}^n x_k(t) / T_k + x_{n+1}(t) + \sum_{k=1}^n u_k(t) \quad (t = 0, \dots, T-1),$$

$$x_{n+2}(t+1) = -\alpha_2 x_{n+1}(t) + x_{n+2}(t) - \sum_{k=1}^n u_k(t) + u_{2n+1}(t) + u_{2n+2}(t) \quad (t = 0), \quad (1)$$

* Работа выполнена при финансовой поддержке АВЦП «Развитие научного потенциала высшей школы» (НИР 2.1.1/2710).