УДК 681.3

Статистические оценки решений задач оптимизации в распределенных системах.I.

Андрюхин А.И.

Донецкий национальный технический университет
alexandruckin@rambler.ru

Abstract

Andruckin A.I. The statistical estimations of the decisions of the problems to optimization in distributed systems.I. *The review of the problems to optimization in distributed systems is presented. The results of structural research and simulation are considered.*

Введение

При повышении эффективности функционирования современных распределенных вычислительных систем (РВС) решаемые задачи относятся к области целочисленной, зачастую булевой оптимизации, и как правило являются NP-трудными задачами[1-3]. На практике используют приближенные решения, которые можно получить за приемлемое время. Одним из возможных способов решения этих задач является предлагаемый подход статистической оценки эффективности решений, получаемых с помощью статистического компьютерного моделирования.

Общая постановка проблемы

Для каждой из решаемых задач в области РВС могут быть заданы возможные варианты ее решения. Каждая задача может состоять из нескольких этапов. Связи между задачами и этапами задаются в виде направленного графа, ребра которого характеризуют соотношения следования или существования между решаемыми задачами и их этапами. Направления ребер соответствуют направлениям информационных потоков. Помимо этого, задается множество узлов сети и связей между ними, виды технических средств, используемых в РВС. Требуется таким образом распределить задачи по узлам и уровням системы, чтобы наблюдался максимум целевой функции W от выбираемых решений.

Возможные временные ограничения для различных задач требуют анализа работы различных узлов РВС. Эти задачи также относятся к классу задач нелинейного булевого программирования, для которых нет эффективных точных методов решения, поэтому на практике решают частные задачи оптимизации распределения задач. Введем следующие обозначения. Пусть $i=(\overline{1,I})$ − множество задач, функций, реализуемых в системе; $j=(\overline{1,J})$ − множество обслуживающих узлов системы; W_{ij} − затраты на реализацию *i*-той задачи в *j*-м узле; $X_{ij}=1(0)$,если задача *i* решается (не решается) на *j*-м узле соответственно. Согласно [2] будем считать, что ограничениями в частных задачах оптимизации являются следующие характеристики:

1. связи между задачами, т.е. задан граф G_F ;
2. связи между узлами, т.е. задан граф G_M ;
3. общие затраты на реализацию задач в системе:

$$\sum_{i=1}^{I}\sum_{j=1}^{J} W_{ij}X_{ij} \le W_{доп};$$

4. затраты на реализацию задач в узлах:

$$\sum_{i=1}^{I} W_{ij}X_{ij} \le W_{доп}{}^{j}, i=(\overline{1,J});$$

5. загрузку каждого узла:

$$\sum_{i=1}^{I} \lambda_i t_{ij} x_{ij} \le \rho_j, j=(\overline{1,J}),$$

где λ_i − интенсивность поступления *i*-той задачи на решение;

6. общее время решения задач:

$$\sum_{i=1}^{I}\sum_{j=1}^{J} t_{ij}x_{ij} \le T;$$

время решения отдельных задач:

$$\sum_{j=1}^{J} t_{ij}x_{ij} \le \tau_i, i=(\overline{1,I}).$$

При этом в качестве критериев оптимизации используют следующие частные критерии:

1.Минимизация затрат на реализацию задач в системе

$$\mathrm{T(X)} = \min \sum_{i=1}^{I} \sum_{j=1}^{J} W_{ij} X_{ij},$$

2. Минимизация общего времени решения всех задач системы

$$\mathrm{T(X)} = \min \sum_{i=1}^{I} \sum_{j=1}^{J} t_{ij} x_{ij},$$

3. Минимизация максимального времени решения задач в системе:

$$\mathrm{T(X)} = \min \left\{ \max_{j} \sum_{i=1}^{I} t_{ij} x_{ij} \right\}$$

Постановка решаемой задачи

Рассмотрим модифицированную задачу оптимизации в [3], которая является частным случаем общей проблемы распределения ресурсов в РВС [2]. Будем считать, что РВС представлена N однородными процессорными ядрами, объединенными иерархической коммуникационной средой. На уровне l размещены n_l элементов. Для каждого элемента i на уровне l определено n_{li} его прямых преемников $i \in \{1,2,...,n_l\}$. Определенная согласно графу соединений функция $L(l,i_1,i_2)$ – номер узла, являющегося ближайшим общим родителем для узлов i_1, i_2, которые находятся на уровне l. Во многих случаях расположение хранилища данных в РВС можем описать функцией *r(p,q)*. Она определяет информационные взаимодействия между частями хранилища данных, распределенными на вычислителях *p,q*.

Пусть P_l – оценка максимальной пропускной способности l-го уровня коммуникационной среды РВС. Основной характеристикой выполнения программы, которая состоит из M ветвей $i = \overline{1,M}$, является множество информационно-логических связей между ее ветвями. Пусть W_{ij} определяет для ветвей i, j объем данных, передаваемых между ними.

Задача оптимального распределения ветвей программы в подсистеме РВС заключается в нахождении бинарного вектора X_{ij} $i \in \overline{1,M}$, $j \in \overline{1,N}$ и $X_{ij} = 1$, если i-я ветвь программы назначается j-му вычислителю. Двоичный вектор X_{ij} должен минимизировать время выполнения всей программы, которое является максимальным из времен выполнения ее ветвей.

Ввиду однородности вычислителей рассматриваем только определенный фактор – время взаимодействия ветвей l в коммуникационной среде РВС. Обозначим через $t(i, j, v, s)$ время взаимодействия между ветвями i, j, распределенными на вычислители v и s соответственно. Это время можно определить, как $\frac{W_{ij}}{P_l}$, где l – есть номер уровня элемента вычислителя, который является ближайшим общим предком для узлов v, s.

Сформулируем задачу оптимального вложения ветвей программы в РВС с целью минимизации накладных расходов:

$$W(X) \rightarrow \min,\ i = 1..M\ ,$$

$$W(X) = \max\{\sum_{j=1}^{M} \sum_{p=1}^{N} \sum_{q=1}^{N} x_{ip} x_{jq} \lambda(i, j, p, q)\},$$

где λ – функция издержек при использовании ветвей программы *i,j* при назначении на узлы-вычислители p, q соответственно. Будем считать $\lambda(i, j, p, q) = r(p,q)\, t(i, j, p, q)$.

Поставленная задача является целочисленной минимаксной задачей, для которой при ее большой размерности строят приближенные решения. В работе предлагается статистический метод построения и оценки приближенного решения.

Решение проблемы

Основные этапы решения базируются на следующих сведениях.

При распределении M ветвей программы на N вычислителей имеем общее число вариантов размещения $V = \frac{N!}{(N-M)!}$.

Число V велико при достаточно больших N и M и большой разности между N и M, следовательно, можно рассматривать множество решений как генеральную совокупность.

Рассмотрим n независимых измерений некоторой величины, истинное значение которой неизвестно. Будем рассматривать результаты измерений как независимые случайные величины $X_1,...X_n$.

Поэтому, согласно [4-5] можно использовать оценку математического ожидания при неизвестном σ при помощи доверительного интервала,

$$\overline{X} - t_\gamma \frac{s}{\sqrt{n}} < M_x < \overline{X} + t_\gamma \frac{s}{\sqrt{n}} \qquad (1)$$

где γ − заданная погрешность.

Значение t_γ определяем по таблице распределения Стьюдента, согласно которого, если $n > 120$ имеем при $\gamma = 0.95$ значение $t_\gamma = 1.96$,при $\gamma = 0.99$ $t_\gamma = 2.617$, при $\gamma = 0.999$ $t_\gamma = 3.374$.

Подчеркнем, что согласно [4] «при больших n указанные соотношения имеют место даже тогда, когда случайные величины $X_1,...X_n$ имеют распределение, отличное от нормального». На основе *n* генерируемых случайных решений задачи производится оценка распределения значений целевой функции и тем самым приближенная оценка значений приемлемых решений.

Согласно неравенства Чебышева, если случайная величина X имеет конечные математическое ожидание μ и дисперсию σ^2 , то $P(|X- \mu| \geq k\sigma) \leq k^{-2}$.

Отсюда можем записать, что

$$P(X \leq \mu - k\sigma) \leq k^{-2} \qquad (2)$$

Основываясь на этом соотношении, можно определять при задаваемых k вероятностные оценки значения X. Понятно, что исходя из конкретики нашей задачи k < μ/ σ.

Имея оценки математического ожидания и дисперсии с определенной точностью, можем использовать распределения порядковых статистик для оценки точного решения [6].

Оценка с помощью порядковых статистик

Дальнейшее использование статистических оценок связано с использованием порядковых статистик. Напомним, что если мы расположим значения целевой функции W_i в порядке возрастания их значения и обозначим элементы этой возрастающей последовательности через $W^{(1)}$, $W^{(2)}$, $W^{(3)}$,…, $W^{(n)}$, то каждый из $W^{(i)}$ называется порядковой статистикой. Ясно, что $W^{(1)}$-наименьшее значение, $W^{(n)}$-наибольшее значение из всех $W^{(i)}$.Возможно использование соотношения [6]:

$$MW^{(1)} \geq \mu-(n-1)\ \sigma/(2n-1)^{1/2} \qquad (3)$$

Согласно соотношения (3) и данным таблицы 4.2 из [6], в которой выполнена оценка размаха для произвольного, нормального и равномерного распределений, можем построить определенные границы значения точного решения в зависимости от n .

На основании вышеизложенных теоретических сведений имеем следующую схему статистической оценки и прогноза значения целевой функции W(X_{ij}) для определенной структуры вычислительной среды и для заданных связей между ветвями рассматриваемой задачи.

1.Выполняем n раз генерацию случайного решения задачи, т.е. определение значения X_{ij} и рассчитываем значение целевой функции W(X_{ij}).

2.Определяем согласно (1) оценки математического ожидания M[W(X_{ij})].

3.На основании (2) при задаваемых k< μ/ σ, где μ= M[W(X_{ij})], σ =D[W(X_{ij})] выполняем оценку вероятности значений W.

Экспериментальные результаты

Рассмотрим распределенную вычислительную среду из 12 вычислителей на рис.1. из [1]. Таким образом задается соответствующий граф G_M . Считаем, что число параллельных ветвей программы равно 8.

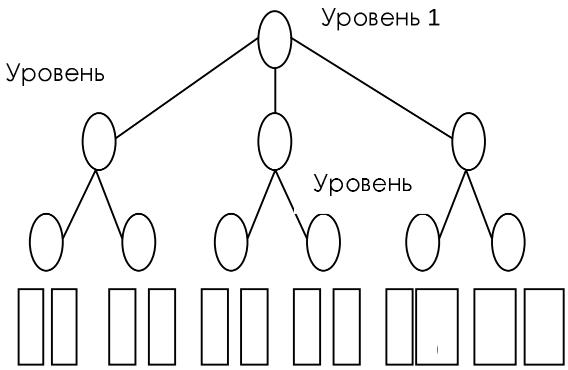

Рисунок 1 – Структура РВС.

Рассмотрим три варианта информационно-логических связей между ветвями программы, которые представлены первыми 8 ветвями (из 16) известных тестов HPL, NPB Multigrid, NPB Conjugate Gradient. Соответствующие им графы G_F представлены конкретными матрицами инцидентности в программной реализации.

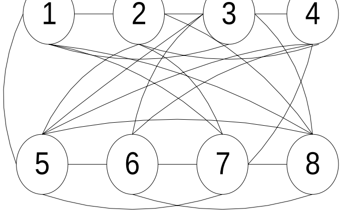

Рисунок 2 – Структура теста подсхемы HPL

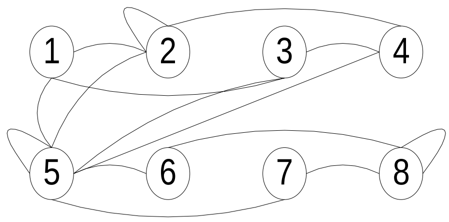

Рисунок 3 – Структура теста подсхемы NPB Conjugate Gradient.

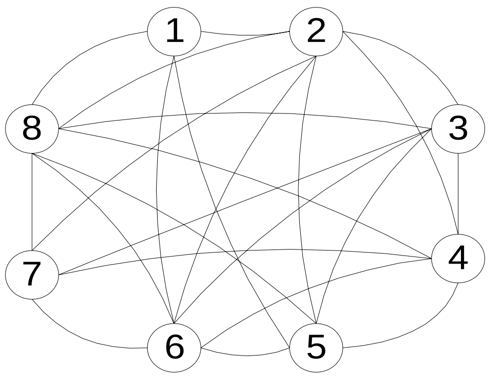

Рисунок 4 – Структура теста подсхемы NPB Multigrid

В табл.1 представлены типичные результаты расчетов для этих схем при определенных параметрах. Конкретные значения P_l , W_{ij} определялись по схеме P_l=(1,20,100) при l=1,2,3 соответственно, а $W_{ij}=1$для всех *i,j*.

В табл.2 представлены результаты расчетов для случая, когда W_{ij} случайная величина, имеющая равномерное распределение от 1 до 200.

Заключение и выводы

Анализ представленных и аналогичных им расчетов, при которых варьировались значения P_l и рассматривались различные параметры равномерного распределения W_{ij} , позволяет сделать следующие выводы.

Точное решение W_T имеет значение, которое согласно неравенства Чебышева оценивается вероятностью 0.02-0.01. Для рассмотренных вариантов и аналогичных примеров отношение размаха $(m-W_T)/\sigma$ находится в интервале 4.5-7.

Дальнейшие шаги в развитие этого подхода связаны с:

1) варьированием вида и параметров распределения W_{ij} *и* P_l*;*

2) использованием других топологических структур распределенных вычислительных систем;

3) расширением области применения статистического подхода для иных задач с соответствующими критериями, и в частности известной задачи оптимального распределения хранилища данных в вычислительных узлах [7];

4)использованием генетических алгоритмов для поиска оптимальных решений.

Литература

1. Гэри М., Д. Джонсон. Вычислительные машины и труднорешаемые задачи. М.: Издательство "Мир", 1982 г. - 416 с.
2. Пономаренко В.С., Листровой С.В., Минухин С.В., Знахур С.В. Методы и модели планирования ресурсов в GRID-системах. СИД "ИНЖЭК", 2008. с 408.
3. Хорошевский В.Г.,Курносов М.Г. Моделирование алгоритмов вложения параллельных программ в структуры распределенных вычислительных систем//Сборник трудов конференции Моделирование-2008,т.2, Киев,14-16 мая 2008 г.с.435-440.
4. Ван дер Варден Б.Л. Математическая статистика. М.:Издательство иностранной литературы, 1960, с.139.
5. Гмурман В.Е. Теория вероятностей и математическая статистика. М.:Высшая школа, 1972,с.101.
6. Дейвид Г.Порядковые статистики. М.:Наука,1979 г.-336 с.
7. Хорошевский В.Г.Архитектурные концепции, анализ и организация функционирования вычислительных систем//Сборник трудов конференции Моделирование-2008,т.2, Киев,14-16 мая 2008 г.с.15-24.

Таблица 1.Статистические результаты с одинаковыми весами.

Число испытаний	n=200	n=1000
Подсхема HPL точное решение min=252		
Найденное решение	min=501.0	min=411.000000
Среднее, ср.кв.отк.	579.90 53.119	576.79 52.35
Оценка среднего по Стьюденту	P(572.54< T <587.26)=0.95 P(570.07< T <589.73)=0.99 P(567.23< T <592.57)=0.999	P(573.55< T <580.04)=0.95 P(572.46< T <581.13)=0.99 P(571.21< T <582.38)=0.999
Оценки согласно неравенства Чебышева	P(T < 526.78) <=1.0 P(T < 473.66) <=0.25 P(T < 420.54) <=0.111 P(T < 367.42) <=0.0625 P(T < 314.30) <=0.040 P(T < 261.18) <=0.027778 P(T < 208.06) <=0.020408 P(T < 154.94) <=0.015625 P(T < 101.82) <=0.012346 P(T < 48.70) <=0.010	P(T < 524.43) <=1.0 P(T < 472.07) <=0.25 P(T < 419.71) <=0.111 P(T < 367.36) <=0.0625 P(T < 315.00) <=0.040 P(T < 262.64) <=0.027778 P(T < 210.28) <=0.020408 P(T < 157.92) <=0.015625 P(T < 105.56) <=0.012346 P(T < 53.20) <=0.010000 P(T < 0.84) <=0.008264
Подсхема NPB Multigrid точное решение min=318		
Найденное решение	min=602.0	min=412.0
Среднее, ср.кв.отк.	625.17 39.12	621.103 39.706
Оценка среднего по Стьюденту	P(619.74< T <630.59)=0.95 P(617.93< T <632.40)=0.99 P(615.83< T <634.50)=0.999	P(618.64< T <623.56)=0.95 P(617.81< T <624.38)=0.99 P(616.86< T <625.33)=0.999
Оценки согласно неравенства Чебышева	P(T < 586.047) <=1.0 P(T < 546.924) <=0.25 P(T < 507.80) <=0.111 P(T < 468.67) <=0.0625 P(T < 429.55) <=0.04 P(T < 390.43) <=0.027778 P(T < 351.31) <=0.020408 P(T < 312.18) <=0.015625 P(T < 273.06) <=0.012346 P(T < 233.94) <=0.010000 P(T < 194.81) <=0.008264 P(T < 155.696765) <=0.006944 P(T < 116.57) <=0.005917 P(T < 77.45) <=0.005102 P(T < 38.32) <=0.004444	P(T < 581.39) <=1.0 P(T < 541.68) <=0.250000 P(T < 501.98) <=0.111111 P(T < 462.27) <=0.062500 P(T < 422.56) <=0.040000 P(T < 382.86) <=0.027778 P(T < 343.15) <=0.020408 P(T < 303.44) <=0.015625 P(T < 263.74) <=0.012346 P(T < 224.03) <=0.010000 P(T < 184.32) <=0.008264 P(T < 144.62) <=0.006944 P(T < 104.91) <=0.005917 P(T < 65.20) <=0.005102 P(T < 25.50) <=0.004444
Подсхема NPB Conjugate Gradient точное решение min=164		
Найденное решение	min=412.000000	min=412.000000
Среднее, ср.кв.отк.	509.39 66.78	514.32 69.87
Оценка среднего по Стьюденту	P(500.13< T <518.64)=0.95 P(497.03< T <521.74)=0.99 P(493.45< T <525.32)=0.999	P(509.99< T <518.65)=0.95 P(508.53< T <520.10)=0.99 P(506.86< T <521.77)=0.999
Оценки согласно неравенства Чебышева	P(T < 442.600421) <=1.0 P(T < 375.810841) <=0.25 P(T < 309.021262) <=0.111 P(T < 242.231683) <=0.0625 P(T < 175.442104) <=0.04	P(T < 444.44) <=1.0 P(T < 374.57) <=0.25 P(T < 304.70) <=0.111 P(T < 234.82) <=0.0625 P(T < 164.95) <=0.04 P(T < 95.079) <=0.0277 P(T < 25.205) <=0.020

Таблица 2.Статистические результаты со случайными весами.

Число испытаний	n=10000	n=1000
Подсхема HPL Точное решение min= 15821		
Найденное решение	min=42547.00	min=45302.000000
Среднее, ср.кв.отк.	69449.93 8257.81	69231.61 8184.55
Оценка среднего по Стьюденту	P(69240.98< T <69658.88)= 0.95 P(69170.94< T <69728.92)= 0.99 P(69090.24< T <69809.63)= 0.999	P(68724.32< T <69738.89)= 0.95 P(68554.28< T <69908.94)= 0.99 P(68358.36< T <70104.86)= 0.999
Оценки, согласно неравенства Чебышева.	P(T < 61192.12) <= 1.00 P(T < 52934.30) <= 0.25 P(T < 44676.49) <= 0.11 P(T < 36418.68) <= 0.06 P(T < 28160.86) <= 0.04 P(T < 19903.05) <= 0.03 P(T < 11645.23) <= 0.02 P(T < 3387.42) <= 0.02	P(T < 61047.06) <= 1.00 P(T < 52862.50) <= 0.25 P(T < 44677.95) <= 0.11 P(T < 36493.40) <= 0.06 P(T < 28308.85) <= 0.04 P(T < 20124.30) <= 0.03 P(T < 11939.75) <= 0.02 P(T < 3755.19) <= 0.02
Подсхема NPB Multigrid Точное решение min= 23466		
Найденное решение	min=54783.0	min=58991.0
Среднее, ср.кв.отк.	80631.70 8738.75	80663.00 8474.30
Оценка среднего по Стьюденту	P(80410.58< T <80852.82)= 0.95 P(80336.45< T <80926.94)= 0.99 P(80251.05< T <81012.34)= 0.999	P(80137.76< T <81188.24)= 0.95 P(79961.69< T <81364.31)= 0.99 P(79758.83< T <81567.17)= 0.999
Оценки, согласно неравенства Чебышева.	P(T < 71892.95) <= 1.00 P(T < 63154.20) <= 0.25 P(T < 54415.46) <= 0.11 P(T < 45676.71) <= 0.06 P(T < 36937.96) <= 0.04 P(T < 28199.22) <= 0.03 P(T < 19460.47) <= 0.02 P(T < 10721.73) <= 0.02 P(T < 1982.98) <= 0.01	P(T < 72188.70) <= 1.00 P(T < 63714.41) <= 0.25 P(T < 55240.11) <= 0.11 P(T < 46765.81) <= 0.06 P(T < 38291.51) <= 0.04 P(T < 29817.22) <= 0.03 P(T < 21342.92) <= 0.02 P(T < 12868.62) <= 0.02 P(T < 4394.32) <= 0.01
Подсхема NPB Conjugate Gradient Точное решение min= 6056		
Найденное решение	min=35547.0	min=35547.0
Среднее, ср.кв.отк.	60824.00 10614.71	60861.50 10356.85
Оценка среднего по Стьюденту	P(60615.95< T <61032.05)= 0.95 P(60546.21< T <61101.79)= 0.99 P(60465.86< T <61182.14)= 1.00	P(60219.57< T <61503.42)= 0.95 P(60004.40< T <61718.60)= 0.99 P(59756.47< T <61966.52)= 0.999
Оценки, согласно неравенства Чебышева.	P(T < 50209.29) <= 1.00 P(T < 39594.58) <= 0.25 P(T < 28979.88) <= 0.11 P(T < 18365.17) <= 0.06 P(T < 7750.46) <= 0.04	P(T < 50504.64) <= 1.00 P(T < 40147.79) <= 0.25 P(T < 29790.94) <= 0.11 P(T < 19434.09) <= 0.06 P(T < 9077.24) <= 0.04

Поступила в редколлегию 02.03.2009