УДК 004.75

МАТЕМАТИЧНЕ ТА ПРОГРАМНЕ ЗАБЕЗПЕЧЕННЯ ОПТИМАЛЬНОГО РОЗПОДІЛУ ІНФОРМАЦІЙНИХ РЕСУРСІВ СЕРЕД ВУЗЛІВ ОБЧИСЛЮВАЛЬНИХ МЕРЕЖ

Р.О. Тичковський
Асистент*

Г.Г. Цегелик
Доктор фізико-математичних наук, професор, завідуючий кафедрою*
*Кафедра математичного моделювання соціально-економічних процесів
Львівський національний університет імені Івана Франка
Контактний тел.: (032)-296-43-51
E-mail: kafmmsep@franko.lviv.ua

Розглянуто підхід до побудови математичних моделей оптимального розподілу копій файлів та комп'ютерних потужностей серед вузлів обчислювальної мережі. Запропоновано евристичний та генетичний алгоритм для реалізації побудованих моделей

Ключові слова: розподілені обчислення, евристичний алгоритм, генетичний алгоритм

Рассмотрен подход к построению математических моделей оптимального распределения копий файлов и компьютерных мощностей среди узлов вычислительной сети. Предложен эвристический и генетический алгоритм для реализации построеных моделей

Ключевые слова: распределенные вычисления, эвристический алгоритм, генетический алгоритм

The approach to the building of mathematical models for optimal copies of file allocation and performance of computers amongst nodes of computing networks is described. The heuristic and genetic algorithm for realization of the received mathematical models is proposed

Key words: distributed solving a tasks, heuristic algorithm, genetic algorithm

Вступ

Одним із основних напрямків розвитку інформаційних технологій є географічно розподілена обробка даних в комп'ютерній мережі, що пов'язано з необхідністю інтеграції географічно розподілених інформаційних ресурсів.

До інформаційних ресурсів, зокрема, належать: файли даних, процесорні та обчислювальні потужності, засоби передачі інформації, засоби зберігання інформації, термінали користувачів для доступу до даних, програми обробки даних. Вузлами мережі вважають: сервери баз даних, файлові сервери, маршрутизатори, комутатори тощо.

Для реалізації інформаційної інтеграції все більшого поширення набувають розподілені бази даних, які функціонують в обчислювальній мережі, та розподілені обчислення, які сумісно використовують обчислювальні потужності вузлів обчислювальної мережі.

Аналіз останніх досліджень

Проблеми оптимізації розміщення інформаційних ресурсів розглянуто у [1,2,4]. У [2,4] побудовано математичні моделі оптимального розподілу файлів серед вузлів обчислювальної мережі (ОМ). У [3,6] розглянуто методи і моделі організації доступу до інформації інтернет-серверів з використанням пошукових систем. У [5] вперше розглянуто таку організацію і перегляд сторінок серверу при заданих ймовірностях звертання до сторінок, коли досягається мінімум математичного сподівання загального часу, необхідного для пошуку потрібної сторінки користувачем.

Проблеми оптимізації продуктивності комп'ютерних мереж

Однією з проблем, що виникають при проектуванні і реалізації систем розподіленої обробки інформації, є проблема оптимального розподілу інформаційних ресурсів (файли даних, процесорні та обчислювальні потужності і.т.п.) серед вузлів обчислювальної мережі [1,2]. Математично задача формулюється так: знайти мінімальне значення функції:

$$L=\sum_{i=1}^{m}\sum_{s=1}^{n}Q_{is}x_{is}\rightarrow \min \qquad (1)$$

за умов

$$1\le \sum_{j=1}^{n}x_{ij}\le k_i \quad (i=1,2,...,m); \qquad (2)$$

$$\sum_{i=1}^{m}L_{ij}x_{ij}\le b_j \quad (j=1,2,...,n); \qquad (3)$$

$$x_{ij}\in\{0\cup 1\}\ (i=1,2,...,m;\ j=1,2,...,n); \qquad (4)$$

де n – кількість вузлів обчислювальної мережі; m – кількість файлів бази даних, які необхідно розмістити серед вузлів мережі; K_j – j-й вузлів мережі; F_i – i-й файл бази даних. Q_{is},L_{ij},b_j – невід'ємні цілі сталі, які в кожному конкретному випадку мають певний фізичний зміст;

x_{ij} –шукані величини,

$$x_{ij}=\begin{cases}1, \text{ якщо файл } F_i \text{ міститься у вузлі } K_j,\\ 0, \text{ у противному випадку}\end{cases}$$

Ця задача належить до класу NP-повних. Задачі такого типу є досить громіздкими і алгоритм повного перебору варіантів є непридатним для їх розв'язування. Тому пропонується використати евристичний та генетичний алгоритми. За допомогою цих алгоритмів, можна знайти наближений розв'язок за незначну кількість ітерацій.

Методи реалізації математичних моделей оптимального розподілу інформаційних ресурсів серед вузлів обчислювальних мереж

Для розв'язування задачі (1)–(4), розроблено евристичний алгоритм, суть якого полягає в слідуючому. Нехай $N_k=C_n^k$ – кількість можливих варіантів розподілу k копій файлу $(1\le k\le n)$, а $M_r=\sum_{k=1}^{r}N_k$ – число можливих варіантів розподілу копій файлу, який може мати від однієї до r копій. Кожному такому варіанту розподілу із множини M_r варіантів поставимо у відповідність послідовність двійкових цифр ($\varepsilon_1,\varepsilon_2,...,\varepsilon_n$), де

$$\varepsilon_j=\begin{cases}1, \text{ якщо копія файла знаходиться в } j-\text{ому вузлі};\\ 0, \text{ у противному випадку.}\end{cases}$$

Наприклад, якщо n=4, то можливим варіантам розподілу копій файлу серед чотирьох вузлів мережі будуть відповідати такі послідовності двійкових цифр: у випадку однієї копії – $(1,0,0,0)$, $(0,1,0,0)$, $(0,0,1,0)$, $(0,0,0,1)$; у випадку двох копій – $(1,1,0,0)$, $(0,1,1,0)$, $(1,0,1,0)$, $(1,0,1,0)$, $(0,1,0,1)$, $(1,0,0,1)$; у випадку трьох копій – $(1,1,1,0)$, $(1,1,0,1)$, $(1,0,1,1)$, $(0,1,1,1)$; у випадку чотирьох копій – $(1,1,1,1)$. Тому $N_1=4$, $N_2=6$, $N_3=4$, $N_4=1$, $M_1=4$, $M_2=10$, $M_3=14$, $M_4=15$.

Множину натуральних чисел від 1 до M_{k_i} позначимо через P_i. Через M_{is} позначимо множину індексів $j(j=1,2,...,n)$, яка визначається так: $j\in M_{is}$, якщо при s-му варіанті розподілу копій файлу F_i копія цього файлу міститься у вузлі K_j.

Алгоритм евристичного методу складається з двох етапів. На першому етапі знаходимо початковий розподіл файлів, який завжди буде оптимальним, якщо не враховувати умови (3). На кожному кроці другого етапу проводиться перерозподіл одного файлу із переповненого вузла таким чином, щоб досягти мінімального збільшення значення цільової функції. Виконання другого етапу алгоритму продовжується доти, доки не буде знайдено розподіл, що задовольняє умови (3).

Побудова математичних моделей оптимального використання комп'ютерних потужностей в обчислювальних мережах та методи їх реалізації

Експлуатація ОМ може передбачати не тільки доступ користувачів до інформаційних ресурсів кожного вузла, а й можливість використання комп'ютерних потужностей в будь-якому вузлі. В зв'язку з цим, виникає проблема побудови математичних моделей оптимального (з точки зору певного критерію) використання комп'ютерних потужностей в ОМ і методів їх реалізації. Нами побудовано низку математичних моделей, однією з яких є така.

Нехай n – кількість вузлів ОМ; m – кількість різних типів задач, призначених до розв'язування; m_i – кількість задач i-го типу; K_j – j-й вузол ОМ; Z_i – i-та задача; t_{ij} – час виконання задачі Z_i на ЕОМ вузла K_j; t_j – час, протягом якого можна використати ЕОМ вузла K_j. Треба так розподілити задачі між ЕОМ у мережі, щоб сумарний час розв'язування задач був мінімальний. Якщо x_{ij} – кількість задач i-го типу, що планується розв'язати на ЕОМ вузла K_j, то математична модель матиме вигляд

$$L=\sum_{i=1}^{m}\sum_{j=1}^{n}t_{ij}x_{ij}\rightarrow \min \qquad (5)$$

за умов

$$\sum_{i=1}^{m}t_{ij}x_{ij}\le t_j \quad (j=\overline{1,n}); \qquad (6)$$

$$\sum_{j=1}^{n}x_{ij}=m_i \quad (i=\overline{1,m}); \qquad (7)$$

$$x_{ij}\in\{\overline{0,m_i}\}\ (i=\overline{1,m};\ j=\overline{1,n}). \qquad (8)$$

Для розв'язування задач такого типу пропонується використати евристичний та генетичний алгоритми. За допомогою цих алгоритмів можна знайти наближений розв'язок за незначну кількість ітерацій.

Генетичні алгоритми розв'язування задач дискретної оптимізації

Для розв'язування задачі (1)-(4) за допомогою генетичного алгоритму, необхідно закодувати оптимізуючі параметри в двійкові рядки (хромосоми). Ці параметри повинні задовольняти умови (2)-(4). Вибравши певного вигляду їх кодування, можна добитися того, що умови (2),(4) задачі будуть задовольнятись автоматично. Специфіка кодування полягає в тому, що допустимий розв'язок задачі представляться у вигляді вектора $R=\left\{R^{(1)},...,R^{(m)}\right\}$, де $R^{(i)}\left(1\le R^{(i)}\le |P_i|\right)$ означає, що копії i-го файла розподілені по $R^{(i)}$-му варіанту, і кодується вектор R, а не масив $X=\left[x_{ij}\right]_{i,j=1}^{m,n}$. Закодований вектор R має двійкове представлення $S=\left\{S^{(1)},S^{(2)},...,S^{(m)}\right\}$ довжини $q=\sum_{j=1}^{m}q_i$, де q_i обчислюються з такої нерівності:

$$2^{q_i-1}<\left(N_i-1\right)\le 2^{q_i}-1 .$$

Для врахуваня умови (3) треба скористатися методикою штрафних функцій, тобто функцію корисності в генетичному алгоритмі задати в такому вигляді:

$$f=L+\sum_{j=1}^{n}A_j\max\left\{0,\sum_{i=1}^{m}L_{ij}x_{ij}-b_j\right\}.$$

де A_j –коефіцієнти масштабування. Ознакою зупинки генетичного алгоритму може бути умова:

$$\left|\frac{f_{max}(P(t))}{f_{avg}(P(t))}-\frac{f_{max}(P(t-1))}{f_{avg}(P(t-1))}\right|\le\varepsilon .$$

Математичне моделювання оптимального доступу до інформації серверів з боку користувачів

Припустимо, що інформація, яка зберігається на віддаленому сервері, розміщена на N сторінках. Вважатимемо, що всі сторінки розбиті на n блоків по m сторінок в кожному (N=n·m) і пошук користувачем потрібної сторінки відбувається шляхом послідовного читання блоків сторінок і їх послідовного переляду. Нехай

p_i – ймовірність звертання до i-ої сторінки;

$a_0=b_0+d_0m$ – час читання блоку сторінок, де b_0,d_0 –деякі сталі;

t_0 – середній час перегляду однієї сторінки користувачем;

E – математичне сподівання загального часу, необхідного для пошуку сторінки.

Тоді E виразиться формулою

$$E=\sum_{i=1}^{n}\sum_{j=1}^{m}\left(ia_0+\left((i-1)m+j\right)t_0\right)p_{(i-1)m+j} .$$

Нами знайдено явний вираз для E як у випадку рівномірного розподілу ймовірностей звертання до сторінок, так і для таких законів нерівномірного розподілу ймовірностей як [2,4]: Зіпфа, узагальнений та"бінарний" розподіл.

Виведено співвідношення для знаходження значень параметрів n і m, при яких математичне сподівання E досягає мінімуму.

1. У разі рівномірного розподілу ймовірностей звертання до сторінок для E, одержуємо вираз

$$E=\frac{1}{2}\left(\left(\frac{N}{m}+1\right)\left(b_0+d_0m\right)+\left(N+1\right)t_0\right).$$

Функція E досягає мінімуму при $m=\left(\frac{b_0}{d_0}N\right)^{1/2}$. При цьому $n=\left(\frac{d_0}{b_0}N\right)^{1/2}$.

2. Нехай ймовірності звертання до сторінок задовольняють "бінарний" закон розподілу. Тоді аналогічно як в [4] для E маємо

$$E=\frac{2^m}{2^m-1}a_0+2t_0 .$$

Для обчислення значення m, при якому E досягає мінімуму, дістаємо рівняння

$$2^m=1+\left(\frac{b_0}{d_0}+m\right)\ln 2 .$$

3. Якщо ймовірності звертання до сторінок розподілені за законом Зіпфа, то аналогічно як в [4]

$$E=\frac{1}{H_N}\left(\left(H_N+n-\frac{1}{2}\ln n-\frac{1}{2}\ln 2\pi\right)a_0+Nt_0\right).$$

де

$$C_1=\frac{1}{2}\ln 2\pi .$$

Для визначення параметра n, при якому функція E досягає мінімуму, одержуємо рівняння

$$\left(2n-1\right)n=\frac{Nd_0}{b_0}\left(2H_N-2C_1+1-\ln n\right).$$

4. Нехай розподіл ймовірностей звертання до сторінок задовольняє узагальнений закон розподілу. Тоді аналогічно як в [4]

$$E=\frac{1}{H_N^{(c)}}\left(\left(H_N^{(c)}-\frac{N^{1-c}}{1-c}\left(\frac{c-1}{2-c}n+\frac{\alpha^{(c)}(n)}{n^{1-c}}\right)\right)\left(b_0+\frac{d_0N}{n}\right)+H_N^{(c-1)}t_0\right)$$

Якщо похідну від $\alpha^{(c)}(n)$ замінити різницею $\alpha^{(c)}(n+1)-\alpha^{(c)}(n)$, то для наближеного обчислення значення n, при якому E досягає мінімуму, дістаємо рівняння

$$n^{3-c}+\left(2-c\right)\left(n+\frac{2-c}{1-c}\frac{d_0}{b_0}N\right)\alpha^{(c)}(n)=\left(2-c\right)\frac{d_0}{b_0}N^cn^{1-c}H_N^{(c)}+$$

$$+\frac{2-c}{1-c}n\left(n+\frac{d_0}{b_0}N\right)\left(\alpha^{(c)}(n+1)-\alpha^{(c)}(n)\right).$$

Висновки

У статті приведено математичну модель визначення оптимальної кількості файлів, розподілених серед вузлів обчислювальної мережі, яка зводиться до узагальненої задачі про призначення з нелінійною цільовою функцією. Запропонований евристичний метод реалізації побудованої математичної моделі, який дає змогу за скінченну кількість кроків знайти оптимальний або майже оптимальний розв'язок задачі.

Запропоновано підхід до побудови математичних моделей оптимального розподілу задач серед вузлів обчислювальної мережі. За критерій оптимальності вибирається сумарний час розв'язування задач. Математичною моделлю задачі є задача дискретного програмування, для розв'язування якої запропоновано евристичний алгоритм. Цей алгоритм дає змогу за скінченну кількість кроків знайти оптимальний або майже оптимальний розподіл задач серед вузлів ОМ. Недоліком алгоритму є те, що він не дає гарантії отримання оптимального розподілу в разі розв'язування задачі.

Для розв'язування отриманих задач дискретного програмування використано генетичний алгоритм. Запропоновано кодувати оптимізуючі параметри в двійкові рядки у специфічний спосіб, щоб автоматично задовольнити умови задачі. Це дало змогу значно скоротити витрати часу і ресурсів комп'ютера при розв'язувані задач цим алгоритмом.

Побудовано математичні моделі оптимального доступу до розбитої на сторінки інформації серверу з боку користувача. За критерій оптимальності прийнято математичне сподівання загального часу, необхідного для пошуку сторінки. Математичні моделі враховують ймовірності звертання до сторінок, час читання блоку сторінок і середній час перегляду однієї сторінки користувачем.

Знайдено явний вираз математичного сподівання для різних законів розподілу ймовірностей звертання до сторінок і виведені співвідношення для знаходження параметрів, за яких математичне сподівання досягає мінімуму.

Література

1. Демидович О.В. Математичні моделі оптимального розподілу інформаційних ресурсів серед вузлів обчислювальних мереж і методи їх реалізації: Автореф. дис. канд. техн. наук. Львів, 2001. – 16с.
2. Кнут Д. Искусство программирования для ЭВМ. Т.3: Сортировка и поиск. – М.:Издательский дом “ Вильямс ”, 2000. – 840с.
3. Пелещишин А.М. Методи та алгоритми моделювання Web-систем: Дис. ... канд. техн. наук: 01.05.03; Львів, 2001. – 150с.
4. Цегелик Г.Г. Организация и поиск информации данных. – Львов: Свит, 1990. – 186с.
5. Цегелик Г.Г., Тичковський Р.О. Математичне моделювання оптимального доступу до інформації серверів зі сторони користувачів // Відбір і обробка інформації: Зб. наук. праць. – Львів, 2004. Вип 21. с.196-200.
6. Baeza-Yates R., Castilio C. Relating Web Structure and User Search Behavior. – Center for Web Research // Department of Computer Science, University of Chile, 2002. – 24p.