

УДК 621.658.012.531

В.В. Кирюхин, доцент, канд. физ.-мат. наук,**М.И. Логинов***Севастопольский национальный технический университет**Студгородок, г. Севастополь, Украина, 99053**E-mail: kvt@sevgtu.sebastopol.ua***ОПТИМИЗАЦИЯ РАЗМЕЩЕНИЯ ФАЙЛОВ ДАННЫХ ПО КРИТЕРИЮ
МИНИМУМА СТОИМОСТИ НА НАКОПИТЕЛЯХ
С ПОСТОЯННЫМ ВРЕМЕНЕМ ДОСТУПА**

Рассматривается задача выбора оптимального размещения файлов данных по накопителям с постоянным временем доступа, минимизирующего общую стоимость размещения, при этом учитываются ограничения на среднее время доступа и на равномерность размещения данных. Задача формулируется и решается в терминах линейного программирования.

На серверах компьютерной сети, обслуживающих интенсивные потоки мультимедийных данных и располагающих внешней памятью большой емкости, возникает нетривиальная проблема рационального в заданном смысле распределения этих данных по накопителям памяти. Эта проблема в общем виде упоминается, как одна из важных, в монографии [1]. В работе [2], в частности, решается задача оптимального размещения данных по файловым серверам, минимизирующего среднее время доступа для случая, когда стоимость размещения и хранения единицы данных является возрастающей (линейно) функцией от их объема. Однако результаты статистических исследований [3] показывают, что в достаточно широком диапазоне объемов эту стоимость для современных накопителей можно считать не зависящей от указанного объема.

Цель данной работы – решение соответствующей оптимизационной задачи, которая математически формулируется следующим образом.

В памяти, состоящей из N накопителей известной емкости $V_1, \dots, V_N$, необходимо разместить некоторые данные, общий объем которых равен V , так, чтобы минимизировать общую стоимость S размещения и при этом удовлетворить ограничениям как по предельно допустимому среднему времени T доступа, так и по равномерности загрузки накопителей. Предполагается, что характеристики i -го накопителя: стоимость S_i размещения на нем единицы данных и время a_i доступа – известны, причем a_i не зависят от объема размещаемых данных, по крайней мере, в границах, определяемых некоторыми коэффициентами A и B , $0 < A < B < 1$.

Пусть $x = (x_1, \dots, x_n)$, где x_i – количество данных, размещаемых на i -ом накопителе. Тогда общая стоимость $S(x_1, \dots, x_n)$ размещения выразится как

$$S(x_1, \dots, x_n) = \sum_{i=1}^n S_i \cdot x_i. \quad (1)$$

Требуется минимизировать функцию (1) при ограничениях

$$\sum_{i=1}^n \frac{x_i}{V} \cdot a_i \leq T, \quad (2)$$

$$A \cdot V_i \leq x_i \leq B \cdot V_i, \quad (3)$$

$$\sum_{i=1}^n x_i = V, \quad i = 1, \dots, N. \quad (4)$$

Вектор $x = (x_1, \dots, x_n)$, удовлетворяющий условиям (2) – (4), назовем *допустимым размещением*. Допустимое размещение, минимизирующее функцию (1), будем называть *оптимальным размещением*.

Поскольку сформулированная задача является задачей линейного программирования, то уместно применить симплекс метод для ее решения. Следуя этому методу, получим каноническую задачу линейного программирования:

минимизировать

$$S(x_1, \dots, x_n) = \sum_{i=1}^n S_i \cdot x_i + M \cdot \sum_{j=1}^{n+1} Y_j \quad (5)$$

при условиях

$$\left. \begin{aligned}
 &\frac{a_1}{V} \cdot x_1 + \dots + \frac{a_n}{V} \cdot x_n + x_{n+1} = T \\
 &x_1 + x_{n+2} = B \cdot V_1 \\
 &\dots\dots\dots \\
 &x_n + x_{2n+1} = B \cdot V_n \\
 &x_1 - x_{2n+2} + y_1 = A \cdot V_1 \\
 &\dots\dots\dots \\
 &x_n - x_{3n+1} + y_n = A \cdot V_n \\
 &\sum_{i=1}^N x_i + y_{n+1} = V
 \end{aligned} \right\} \quad (6)$$

$$x_i \geq 0, \quad i = 1, \dots, 3n+1; \quad y_j \geq 0, \quad j = 1, \dots, n+1.$$

Здесь M – достаточно большое положительное число.

При поиске оптимального размещения необходимо сначала убедиться в том, что совокупность заданных ограничений допускает хотя бы одно решение.

В системе (6) ограничений необходимо проверить границы размещения файлов на накопителях. Нижняя граница, определяемая коэффициентом A , всегда должна быть меньше верхней границы, определяемой коэффициентом B .

Необходимо также определить суммарный максимальный объем V_{sum} данных, который могут содержать накопители:

$$V_{sum} = B \cdot \sum_{i=1}^n V_i.$$

При этом, если V_{sum} меньше общего объема размещаемых данных, $V_{sum} < V$, то задача не имеет решения.

Проверка ограничения на время доступа осуществляется согласно следующему алгоритму

1. Накопители сортируются по возрастанию времени доступа, т.е. индексы i упорядочиваются в соответствии с неравенствами $a_1 < a_2 < \dots < a_n$.
2. На накопителе с наименьшим временем доступа (1-й в списке) размещается объем данных, равный BV_1 (V_1 – емкость 1-го накопителя); если $V > BV_1$, иначе размещается объем V .
3. На втором накопителе размещается BV_2 данных, если оставшийся объем больше BV_2 , иначе размещается весь объем неразмещенных данных. И так далее, пока все данные не будут размещены.
4. Полученное размещение подставляется в левую часть неравенства (2), вычисляется среднее время доступа и сравнивается с заданным T ; если T меньше полученного значения, то задача решения не имеет. Иначе решается задача отыскания оптимального размещения как классическая задача (5) – (6) линейного программирования.

С учетом вышесказанного нами была составлена компьютерная программа, реализующая алгоритм симплекс-метода применительно к канонической задаче (5) – (6).

В программу вводятся исходные данные в следующем порядке: количество накопителей N , их емкости V_i , стоимости размещения единицы данных S_i на каждом из накопителей, суммарный объем размещаемых данных V_{sum} , предельное время доступа T , а также коэффициенты A и B . В результате программа выводит соответствующее сообщение об ошибке, если заданы противоречивые ограничения, иначе – объем данных для каждого накопителя (оптимальное размещение) и значение целевой функции, т.е. минимальную стоимость.

Приведем несколько примеров, иллюстрирующих работу программы.

1. Пусть требуется разместить $V=380$ Тбайт данных на $N=5$ накопителях при предельно допустимом среднем времени T доступа 0,002 с. Коэффициенты A и B равны соответственно 0,3 и 0,8. Значения V_i в Тбайт, удельные стоимости S_i в у.е./Тбайт и времена доступа a_i в миллисекундах – следующие: $V_1=100$; $V_2=150$; $V_3=75$; $V_4=125$; $V_5=50$; $S_1=4$; $S_2=2$; $S_3=5$; $S_4=3$; $S_5=7$; $a_1=1,7$; $a_2=1,8$; $a_3=1,3$; $a_4=1,7$; $a_5=1$.

Задача линейного программирования в стандартной математической формулировке минимизировать

$$S = 4x_1 + 2x_2 + 5x_3 + 3x_4 + 7x_5,$$

при условиях:

$$\begin{cases} 4,47x_1 + 4,73x_2 + 3,42x_3 + 4,47x_4 + 2,63x_5 \leq 2 \cdot 10^3 \\ x_1 \leq 80 \\ x_2 \leq 120 \\ x_3 \leq 60 \\ x_4 \leq 100 \\ x_5 \leq 40 \\ x_1 \geq 30 \\ x_2 \geq 45 \\ x_3 \geq 22 \\ x_4 \geq 38 \\ x_5 \geq 15 \end{cases}$$

$$x_1 + x_2 + x_3 + x_4 + x_5 = 380.$$

На 12-й итерации программа выдает следующее оптимальное размещение: $X = (80, 120, 60, 100, 20)$, $S_{\min} = 1300$ у.е.

2. Изменим значения параметров A и B распределения файлов: $A = 0,5$; $B = 0,8$.

Оптимальное размещение, полученное при этом на 13-й итерации: $X = (80, 120, 55, 100, 25)$, $S_{\min} = 1310$ у.е.

3. Приведем примеры, демонстрирующие отсутствие области допустимых решений. Изменим в первом примере суммарный объем размещаемой информации с 380 на 410 Тбайт.

Программа выдает сообщение:

“При заданных коэффициентах A и B и объемах накопителей задача решения не имеет. Необходимо уменьшить объем размещаемой информации”.

Аналогично, при изменении предельного среднего времени доступа к данным для предыдущего примера с 0,002 до 0,001 с. Получим сообщение, выданное программой:

“При заданном ограничении на время доступа задача решения не имеет. Необходимо увеличить предельное время доступа”.

В перспективе предполагается постановка и решение более общей, нежели рассмотренная, задачи, в которой функция цели $S(x_1, \dots, x_n)$ может быть представлена кусочно-линейной аппроксимацией зависимости, полученной в работе [3].

Библиографический список

1. Столлингс В. Компьютерные системы передачи данных / В. Столлингс. — М.; СПб.; К.: Вильямс, 2002. — 920 с.
2. Гуревич И.М. Об оптимальном размещении данных по файловым серверам вычислительной системы / И.М. Гуревич, В.В. Кирюхин, А.В. Герус // Материалы междунар. науч.-практ. конф. “ИНФОТЕХ-2007” 10–16 сентября 2007 г. — Севастополь, 2007. — Ч. 1. — С. 13–16.
3. Кирюхин В.В. Оценка характеристик дисковых накопителей для серверов компьютерных сетей / В.В. Кирюхин, Ф.Г. Цой // Тез. докл. ВСНТК “Изучаю компьютер – познаю вселенную” 19–23 марта 2007 года. — Севастополь, 2007. — С. 21.

Поступила в редакцию 9.07.2007 г.