

УДК 681.324

В.В. Кирюхин

Севастопольский национальный технический университет.

Студгородок, г. Севастополь, Украина, 99053

E-mail: kvt@sevgtu.sebastopol.ua

ОПТИМАЛЬНОЕ РАСПРЕДЕЛЕНИЕ ПРОПУСКНЫХ СПОСОБНОСТЕЙ КАНАЛОВ В ЦЕНТРАЛИЗОВАННОЙ ЛВС С ПРИОРИТЕТАМИ

Рассматривается задача минимизации ожидаемых потерь в локальной вычислительной сети с центральным сервером, обслуживающим приоритетные потоки заявок ЛВС. Потери связаны с ожиданием передачи заявок по каналам и задержкой в очередях к процессору сервера. Задача формулируется в терминах математического программирования. Предложен алгоритм оптимального распределения пропускных способностей при ограниченной стоимости ЛВС.

Ряд проблем рациональной организации производственных, распределительных и управляющих систем связан с необходимостью обработки информации, поступающей от удалённых пользователей (станций) по каналам связи в центральный вычислитель – сервер. Эта система обработки обычно организована в виде централизованной локальной вычислительной сети звездообразной структуры, в центре которой расположен сервер. Станция i ЛВС связана с сервером каналом пропускной способности c_i и направляет в этот канал через его буфер многомерный поток сообщений известной интенсивности $\lambda_i = (\lambda_{i1}, \dots, \lambda_{ik}, \dots, \lambda_{in})$, где λ_{ik} – интенсивность потока сообщений k -го приоритета. Длина в битах сообщения k -го приоритета ($k = 1, \dots, n$) не зависит от номера i станции и является случайной величиной с известными средним l_k и вторым начальным моментом $l_k^{(2)}$. Каждое такое сообщение предварительно помещается в k -ю секцию буфера канала и выбирается из него для передачи в порядке FIFO. По приходу в центр оно направляется в k -ю секцию буфера сервера и оттуда, так же в порядке FIFO и в соответствии со своим приоритетом, поступает как заявка на обслуживание в процессор сервера. Время обслуживания этой заявки в общем – случайная величина с известным средним ϑ_k и вторым начальным моментом $\vartheta_k^{(2)}$, $k = 1, \dots, n$.

Порядок обслуживания (сканирования) буфера сервера и каналов в рассматриваемой ЛВС определяется фиксированной дисциплиной относительных приоритетов. Эффективность функционирования ЛВС оценивается функцией потерь вида

$$A_2 = \sum_{i=1}^m \sum_{k=1}^n \alpha_k \lambda_{ik} (w_k + w_{ik}), \quad (1)$$

где m – число станций ЛВС; α_k – потери (штраф) за единицу времени ожидания в очередях к каналу и к процессору одной заявки k -го приоритета; w_k – среднее время ожидания начала обслуживания этой заявки в процессоре; w_{ik} – среднее время её ожидания в буфере i -го канала.

Предполагается, что величины w_k и w_{ik} определены для стационарного режима функционирования как сервера, так и всех каналов ЛВС, т.е. режима, в котором загрузка каждого из указанных устройств ЛВС, рассматриваемых как одноканальные системы массового обслуживания, меньше единицы.

Предполагается также, что потери заявок и сообщений по причине нехватки памяти для их размещения в буферах пренебрежимо малы.

Для аренды каналов ЛВС в известном периоде времени выделена некоторая сумма D ; известны также величины d_i , $i = 1, \dots, m$, где d_i – стоимость аренды i -го канала единичной пропускной способности на этот период.

Проблема заключается в выборе пропускных способностей каналов, минимизирующем потери (1); выбор должен производиться в пределах выделенной на аренду каналов суммы.

Сформулируем эту задачу математически.

В указанных выше условиях время w_k ожидания заявки в буфере сервера не зависит от направления (канала) i , откуда она пришла. Каждый из m каналов может рассматриваться как система массового обслуживания типа $M/G/1$ с относительными приоритетами, причём среднее и второй начальный момент времени обслуживания в системе i равны соответственно I_k/C_i и $I_k^{(2)}/C_i$. Используя известные формулы для приоритетных систем, можно преобразовать функцию (1) к виду

$$A_2 = A_0 + A_1(C_1, \dots, C_m), \quad (2)$$

где

$$A_0 = \sum_{k=1}^n \alpha_k \lambda_k w_k, \quad \lambda_k = \lambda_{1k} + \dots + \lambda_{mk}; \quad (3)$$

$$w_k = \frac{1}{2(1 - \sum_{j=1}^k \lambda_j \vartheta_j)(1 - \sum_{j=1}^{k-1} \lambda_j \vartheta_j)} \cdot \sum_{k=1}^n \lambda_k \vartheta_k^{(2)}; \quad (4)$$

$$A_1(C_1, \dots, C_m) = \sum_{k=1}^n \alpha_k \sum_{i=1}^m \lambda_{ik} \frac{1}{2(C_i - \sum_{j=1}^k \lambda_{ij} I_j)(C_i - \sum_{j=1}^{k-1} \lambda_{ij} I_j)} \cdot \sum_{k=1}^n \lambda_{ik} I_k^{(2)}. \quad (5)$$

Поскольку, как это следует из (3) и (4), A_0 не зависит от C_i , то при выборе C_i можно ограничиться минимизацией функции (5) при следующих условиях:

$$\sum_{i=1}^m d_i C_i \leq D; \quad (6)$$

$$C_i \geq 0, \quad i = 1, \dots, m. \quad (7)$$

Таким образом, имеем задачу математического программирования: минимизировать функцию (5) при ограничениях (6) и (7). Следуя установившейся терминологии, будем называть вектор $C = (C_1, \dots, C_m)$, удовлетворяющий условиям (6), (7), допустимым распределением пропускных способностей по каналам ЛВС. Допустимое распределение, минимизирующее потери (5), назовём оптимальным распределением.

Поскольку переменные C_i в (5)-(7) непрерывны, задачу в принципе можно решить, используя известную теорему Куна-Таккера [1]. Однако специфика задачи позволяет избежать применения этого довольно громоздкого аппарата и обойтись более элементарными средствами.

Наличие стационарного режима в ЛВС предполагает одновременное выполнение следующих m неравенств для всех каналов, рассматриваемых как системы массового обслуживания:

$$\rho_i < 1; \quad i = 1, \dots, m,$$

где ρ_i – загрузка i -го канала; $\rho_i = (\lambda_{i1}l_1 + \dots + \lambda_{in}l_n) / C_i$, или

$$C_i > \sum_{k=1}^n \lambda_{ik} l_k > 0, \quad i = 1, \dots, m. \quad (8)$$

Поскольку неравенства (8) имплицируют неравенства (7), последние можно опустить.

Таким образом, задача (5)-(7) эквивалентна следующей: найти вектор $C = (C_1, \dots, C_m)$, минимизирующий функцию

$$A(C_1, \dots, C_m) = \sum_{k=1}^n \alpha_k \sum_{i=1}^m \frac{\lambda_{ik} r_i}{(C_i - a_{ik})(C_i - a_{i,k-1})} \quad (9)$$

при условии

$$\sum_{i=1}^m \alpha_i C_i = D. \quad (10)$$

Здесь $A(C_1, \dots, C_m) = 2A_1(C_1, \dots, C_m)$,

$$r_i = \sum_{k=1}^n \lambda_{ik} l_k^{(2)}, \quad i = 1, \dots, m; \quad (11)$$

$$a_{ik} = \sum_{j=1}^k \lambda_{ij} l_j, \quad i = 1, \dots, m; \quad k = 1, \dots, n. \quad (12)$$

Задача (9)-(10) – это классическая задача Лагранжа отыскания минимума функции (9) при ограничении-равенстве (10). Применяв соответствующий аппарат, получим следующую систему уравнений для оптимального распределения $(C_1, \dots, C_m)$:

$$\left. \begin{aligned} \sum_{k=1}^n \alpha_k \lambda_{ik} r_i \cdot \frac{a_{ik} + a_{i,k-1} - 2C_i}{(C_i - a_{ik})^2 (C_i - a_{i,k-1})^2} + \gamma d_i &= 0; \\ i &= 1, \dots, m; \\ \sum_{i=1}^m d_i C_i - D &= 0, \end{aligned} \right\} \quad (13)$$

где γ – неопределённый множитель Лагранжа, а константы r_i и a_{ik} определяются формулами (11) и (12) соответственно.

Систему (13) можно представить в виде:

$$\left. \begin{aligned} f_i(C_i) &= \gamma; \\ i &= 1, \dots, m; \\ \sum_{i=1}^m d_i C_i &= D, \end{aligned} \right\} \quad (14)$$

где

$$f_i(C_i) = \frac{1}{d_i} \sum_{k=1}^n \alpha_k \lambda_{ik} r_i \cdot \frac{2C_i - a_{ik} - a_{i,k-1}}{(C_i - a_{ik})^2 (C_i - a_{i,k-1})^2} \quad (15)$$

– положительные функции, поскольку имеют место неравенства (8). В силу тех же неравенств (8) производные от всех f_i , определяемых формулами (15),

$$\frac{df_i(C_i)}{dC_i} = \frac{1}{d_i} \sum_{k=1}^n \alpha_k \lambda_{ik} r_i \cdot \frac{2r_i C_i (a_{ik}^2 a_{i,k-1} + a_{ik} a_{i,k-1}^2 + C_i^3 - 3C_i a_{ik} a_{i,k-1})}{(C_i - a_{ik})^3 (C_i - a_{i,k-1})^3},$$

отрицательны. Это означает, что все функции $f_i(C_i)$ монотонны, и поэтому каждое из m первых уравнений системы (14) имеет единственное решение

$$C_i = f_i^{-1}(\gamma), \quad (16)$$

где $i = 1, \dots, m$; f_i^{-1} – функция, обратная функции f_i ; в силу свойств обратных функций, она так же положительна и монотонна. Подставив (16) в последнее уравнение системы (14), получим

$$F(\gamma) = \sum_{i=1}^m d_i f_i^{-1}(\gamma) = D. \quad (17)$$

Функция $F(\gamma)$ как взвешенная с положительными коэффициентами сумма монотонных и положительных функций, также монотонна и положительна. Следовательно, уравнение (17) относительно γ имеет единственное решение

$$\gamma = F^{-1}(D) > 0, \quad (18)$$

где F^{-1} – функция, обратная F .

Подставляя (18) в (16), получаем оптимальное распределение:

$$C_i = f_i^{-1}(F^{-1}(D)), \quad i = 1, \dots, m. \quad (19)$$

Наконец, подставляя (19) в (9), получаем выражение для минимальных потерь в каналах ЛВС.

В заключение остановимся на двух моментах, обеспечивающих корректное и достаточно быстрое отыскание оптимального распределения. Уравнение (17) имеет решение лишь при соблюдении условий (8). Умножая обе части неравенств (8) на d_i и суммируя, получаем, с учётом (10)

$$D > D_0, \quad (20)$$

где

$$D_0 = \sum_{i=1}^m d_i \sum_{k=1}^n \lambda_{ik} l_k \quad (21)$$

– величина, которая имеет смысл минимума арендной платы, обеспечивающей стационарный режим в ЛВС.

Система (14) по необходимости должна решаться численно, поскольку аналитических выражений не существует. Каждое из первых m уравнений можно решить, например, методом Ньютона, взяв в качестве исходной точку с координатами

$$C_i^0 = \sum_{k=1}^n \lambda_{ik} l_k + \frac{D - D_0}{\sum_{i=1}^m d_i}, \quad (22)$$

где $i = 1, \dots, m$, и D_0 вычисляется по формуле (21). Легко проверить непосредственной подстановкой в (10), что это – допустимое распределение.

Систематизируя результаты, можно предложить следующий алгоритм.

Начало

1. Ввод исходных данных:

$$(\lambda_{ik})_m^n, (l_1, \dots, l_n), (l_1^{(2)}, \dots, l_n^{(2)}), (\vartheta_1, \dots, \vartheta_n), (\vartheta_1^{(2)}, \dots, \vartheta_n^{(2)}), (d_1, \dots, d_m), D.$$

2. Проверка условия (20). Если оно не выполняется, то – выдача сообщения «РЕШЕНИЯ НЕТ» и переход на «Конец».

3. Численное решение системы (14) с начальной точкой (21) и функциями $f_i(C_i)$ вида (15).

4. Подсчёт ожидаемых потерь A_2 по формуле (2) с учётом (3) и (5).

5. Вывод оптимального распределения $(C_1, \dots, C_n)$ и ожидаемых потерь A_2 .

Конец

Рассмотрена задача оптимального распределения пропускных способностей каналов ЛВС, минимизирующего потери от ожидания сообщениями начала передачи по этим каналам в направлении сервера. Задача сформулирована в терминах математического программирования. Предложено достаточно простое её решение, учитывающее специфику проблемы. Сформулирован алгоритм оптимального распределения. Реализующая этот алгоритм компьютерная программа используется в составе пакета программной поддержки системотехнического проектирования локальных сетей.

Отметим, что предложенная постановка обобщает известную проблему [2] на случай произвольного распределения длины сообщений и наличия системы приоритетов в каналах.

В перспективе предусматривается постановка задачи комплексной оптимизации как по распределению приоритетов в сервере и каналах, так и по распределению пропускных способностей при большем числе ограничений.

Библиографический список

1. Хедли Дж. Нелинейное и динамическое программирование / Дж. Хедли. — М.: Мир, 1967. — 506 с.
2. Клейнрок Л. Вычислительные системы с очередями / Л. Клейнрок. — М.: Мир, 1979. — 600 с.

Поступила в редакцию 17.03.2003 г.