

Система обслуживания с ветвящимися потоками вторичных требований

Alexander S. Baklashov

07 November, 2023

RUDN University, Moscow, Russian Federation

Введение

В работе “Система обслуживания с ветвящимися потоками вторичных требований” рассматривается однолинейная система обслуживания с простейшим первичным и ветвящимся вторичным потоком требований нескольких типов и произвольной длительностью обслуживания.

Цель исследования:

Оптимизация приоритетного обслуживания для эффективного управления разнообразными требованиями в системах

Основные задачи и актуальность

В данной работе рассматривается задача определения оптимальной дисциплины обслуживания в системе с несколькими типами требований и ветвящимися потоком вторичных требований.

Актуальность работы становится понятна при исследовании работы ЭВМ в различных режимах, исследовании информационно-поисковых и др. систем.

Примеры и прикладное значение

Значимость в реальных приложениях:

1. Пакетный режим работы компьютера:

Программа → стандартные программы, оперативная память, внешние устройства.

2. Поиск информации в массивах данных:

Анализ массивов для эффективного обнаружения информации.

Постановка задачи

Характеристики системы обслуживания:

- Однолинейная система обслуживания (СО) с r типами операций.
- Длительности операций - независимые случайные величины с функциями распределения, имеющими первые два момента b_i, b_{i2}
- Первичные требования - пуассоновский поток с интенсивностью $\lambda_i, \lambda_i \geq 0, i = \overline{1, r}$.

Особенности системы:

- Возможность появления вторичных требований с вероятностью $q_i(n)$ после выполнения операции типа i .
- Функция управления $u(l)$ выбирает требование на обслуживание в зависимости от числа требований в каждой очереди.
- Простои прибора при наличии требований не допускаются

Постановка задачи (3)

Задача и функционал потерь:

Задача: Стремление к минимизации потерь в единицу времени в стационарном режиме.

Функционал $J = \sum_{i=1}^r c_i L_i$ описывает средние потери,

где L_i - среднее число требований типа i в системе в стационарном режиме, c_i — стоимость единицы времени пребывания в системе требования типа r .

Вводится производящая функция (ПФ) числа вторичных требований, возникающих в результате выполнения операции типа i :

$$Q_i(z) = \sum_{n \geq 0} q_i(n) z^n \equiv \sum_{n_1 \geq 0, \dots, n_r \geq 0} q_i(n_1, \dots, n_r) z_1^{n_1} \dots z_r^{n_r}$$

Предположения

Предположение 1: Свойства распределения вторичных требований

1. Первые два момента распределения числа вторичных требований конечны при всех i :

$$q_{ij} = \frac{\partial}{\partial z_j} Q_i(z)|_{z=1} < \infty$$

$$q_{jk}^i = \frac{\partial^2}{\partial z_j \partial z_k} Q_i(z)|_{z=1} < \infty$$

2. Из теории Фробениуса известно, что у матрицы $Q = \|q_{ij}\|$ с неотрицательными элементами существует единственное собственное значение $\xi > 0$, такое, что модули всех остальных собственных значений не превосходят ξ .

Условия вырождения ветвящегося процесса:

- $\xi < 1$
- Ветвящийся процесс вторичных требований вырождается с вероятностью 1.

Предположение 2 (Часть 2: Доказательство)

1. Условия для определения вектора f :

- $\lambda = (\lambda_1, \dots, \lambda_r)$
- $b = (b_1, \dots, b_r)$

2. Доказательство: вектор $f = (I - Q)^{-1}b$ определен и имеет положительные компоненты $f > 0$.

- Из условия $\xi < 1$ и сходимости ряда $1 + x + x^2 + \dots$ на спектре матрицы Q следует, что матрица $I - Q$ обратима.
- Справедливо представление:

$$(I - Q)^{-1} = I + Q + Q^2 + \dots$$

- Вектор $f = (I - Q)^{-1}b$ положителен (покомпонентно), если положителен b .

Предположение 3: Условия эргодичности процесса $L(t)$

Определение ρ и условия эргодичности:

- $\rho < 1$
- $\rho = \lambda' f$, где ρ - загрузка системы.

Предположение 3: Доказательство эргодичности процесса $L(t)$ (1)

Доказательство эргодичности процесса $L(t)$:

1. Ограничение на точки регенерации типа 0:

- Использование теоремы Смита.
- Периоды регенерации - циклы занятости с абсолютно непрерывной ФР.

2. Утверждение в предположениях 1, 2, 3:

- а) ПЛС ФР периода занятости удовлетворяет системе уравнений.
- б) Система имеет единственное решение, где каждая $\pi_i(s)$ - ПЛС собственной ФР.

3. Уравнение для первых моментов π_i :

- Дифференцирование и решение уравнения $A\pi = b$.
- Существование и единственность конечного решения.

4. Вектор вторых моментов π_2 :

- Уравнение $A\pi_2 = \varphi$, где φ определено.

5. Существование двух первых моментов цикла занятости и эргодичность процесса $L(t)$:

- Доказательство вытекает из предположений.

6. Смысл ρ :

- ρ имеет смысл загрузки системы.
- Матрица Q может быть как неразложимой, так и разложимой.

В каждом диагональном блоке матрицы Q найдется такой индекс i , что $\lambda_i > 0$. Это предположение обеспечивает возможность появления в системе требований всех типов.

Основные соотношения

- Для удобства использования аппарата регенерирующих процессов.
- Поведение процесса после t_n зависит от состояния $L(t_n)$ и значения функции переключения.
- Моменты регенерации типа i : t_n такие, что $u(L(t_n)) = i$.

- Вероятности не зависят от начального состояния ($L(0) = 0$).
- Введены функции $H_0(t)$, $H_i(t, z)$ с рядами (??).
- Связь с производящей функцией $L(t)$: $P(t, z) = E\{z^{L(t)}\}$.

1. Связь между функциями и производящей функцией $L(t)$.

$$\cdot p(s, z), \chi_0(s), \chi_i(s, z).$$

2. Соотношения:

$$\cdot p(s, z) = \sum_{i=1}^r \left(\frac{\lambda_i}{s+\lambda_0} \chi_0(s) + \chi_i(s, z) \right) z_i \frac{1-\beta_i(s+\lambda_0-\lambda'z)}{s+\lambda_0-\lambda'z} : p(s, z)$$

$$\cdot \chi_0(s) + \sum_{i=1}^r z_i \chi_i(s, z) = \sum_{i=1}^r \left(\frac{\lambda_i}{s+\lambda_0} \chi_0(s) + \chi_i(s, z) \right) \times \beta_i(s+\lambda_0-\lambda'z) Q_i(z) : \chi_0(s) + \sum_{i=1}^r z_i \chi_i(s, z)$$

3. Пределы:

$$\cdot \lim_{s \rightarrow 0} sp(s, z) = P(z)$$

$$\cdot \lim_{s \rightarrow 0} s\chi_i(s, z) = \chi_i(z)$$

- Существование пределов следствие эргодичности процесса $L(t)$.
- Соотношения:

$$P(z) = \sum_{i=1}^r [\lambda_i(1 - \rho) + \chi_i(z)] \frac{1 - \beta_i(\lambda_0 - \lambda'z)}{\lambda_0 - \lambda'z}$$

$$(1 - \rho) \sum_{i=1}^r \lambda_i(z_i - 1) = \sum_{i=1}^r [z_i - b_i(z)][\lambda_i(1 - \rho) + \chi_i(z)]$$

- $P(z)$ и характеристики процесса $L(t)$.
- Задача оптимизации в виде задачи линейного программирования.

Задача линейного программирования

- Связь между функционалом и функцией управления через переменные x_{ij} .

$$x_{ij} = \frac{\partial}{\partial z_j} \chi_i(z)|_{z=1} = \sum_{l: u(l)=i} l_i h_i(l)$$

- Важное свойство: $x_{ij} = 0$ тогда и только тогда, когда требования типа j имеют приоритет перед требованиями типа i .

- Задача минимизации функционала потерь сводится к задаче линейного программирования.

$$J = \sum_{i,j=1}^r b_i c_i x_{ij} \rightarrow \min$$

- Ограничения задачи линейного программирования.

$$\sum_{i=1}^r (a_{ij}x_{ik} + a_{ik}x_{ij}) = \gamma_{ik}, \quad j, k = 1, r$$

- Вывод формул $J = \sum_{i,j=1}^r b_i c_i x_{ij} \Rightarrow \min$

и

$$\sum_{i=1}^r (a_{ij} x_{ik} + a_{ik} x_{ij}) = \gamma_{ik}, \quad j, k = 1, r.$$

- Обозначим интенсивность выполнения операций типа i
 $R_i = \lambda_i(1 - \rho) + \chi_i(1).$

- Существует оптимальное управление системой, которое осуществляется с помощью приоритетной дисциплины обслуживания.

Алгоритм назначения приоритетов

- Строится последовательность индексов i_1, \dots, i_r .

- На каждом шаге $m, m = \overline{0, r-2}$ вычисляется вектор f^{r-m} из уравнения

$$(I - Q_{r-m})f^{r-m} = b^{r-m} \quad (1)$$

где Q_{r-m} и b_{r-m} составлены из строк и столбцов Q и b с индексами, не совпадающими с i_r, \dots, i_{r-m+1} .

- Вычисление c_i^m по формуле:

$$c_i^m = c_i^{m-1} - \frac{f_i^{r-m+1}}{f_{i_{r-m+1}}^{r-m+1}} c_{i_{r-m+1}}^{m-1}, \quad i \neq i_r, \dots, i_{r-m+1} \quad (c_i^0 = c_i)$$

(2)

- Выбор индекса i_{r-m} из условия:

$$\frac{c_{i_{r-m}}^m}{f_{i_{r-m}}^{r-m}} \leq \frac{c_i^m}{f_i^{r-m}}, \quad i \neq i_r, \dots, i_{r-m+1} \quad (3)$$

- Переход на шаг $m + 1$.

- Оптимальной является дисциплина, при которой $i_1 \succ \dots \succ i_r$.

Обсуждение и примеры

- В зависимости от вида матрицы Q , рассмотренная модель включает в себя многофазные системы и системы с обратной связью [7].
- Решение задачи, поставленной в [12], также может быть получено с использованием данного алгоритма.

Пример 1: $Q = 0$

- Если $Q = 0$, на каждом шаге алгоритма имеем $f^m = b^m$.
- Условие оптимальности дисциплины: $i_1 \succ i_2 \succ \dots \succ i_r$, что соответствует хорошо известному результату [5].

Пример 2: Модельный пример

- Однопроцессорная система в пакетном режиме.
- 3 операции: ввод пакета и его обработка, счет по каждой программе, выдача результатов счета.
- Матрица Q :

$$Q = \begin{bmatrix} 0 & 10 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

- Весовые коэффициенты: $c_1 = 1, c_2 = c_3 = 10$.
- Расчет по алгоритму требует всего 2 шага, оптимальная дисциплина: $3 \succ 2 \succ 1$.

- Условия оптимальности зависят от b_i, c_i и Q , но также необходимо существование вторых моментов b_{i2}, q_{ij}^k и выполнение условия $\rho < 1$.

- Разработана программа по предложенному алгоритму, подготовленная для передачи в фонд алгоритмов и программ по ТМО.

Вывод

- Модель массового обслуживания успешно оптимизирована с использованием линейного программирования.
- Разработан эффективный алгоритм назначения приоритетов для оптимизации систем обслуживания.
- Примеры и обсуждение подчеркивают практическую применимость методов.
- Результаты обещающи и подтверждаются готовой программой для передачи в фонд алгоритмов по теории массового обслуживания.