СТАЦИОНАРНОЕ РАСПРЕДЕЛЕНИЕ СИСТЕМЫ МАССОВОГО ОБСЛУЖИВАНИЯ $MAP/PH/1$

1 Математическая модель

Рассматривается однолинейная система массового обслуживания с бесконечным буфером. Заявки поступают в систему в MAP (Markovian Arrival Process)-потоке, который является ординарным аналогом $BMAP$ (Batch Markovian Arrival Process). Поступление заявок в этом потоке возможно только в моменты скачков некоторой неприводимой цепи Маркова ν_t, $t \geq 0$, с непрерывным временем и конечным пространством состояний $\{0, 1, \dots, W\}$, которая называется управляющим процессом MAP-потока. Интенсивности переходов процесса ν_t, $t \geq 0$, сопровождающиеся поступлением заявки (не сопровождающиеся поступлением заявок), задаются элементами матрицы D_1 (недиагональными элементами матрицы D_0). При этом матрица $D_0 + D_1$ является неприводимым инфинитезимальным генератором цепи Маркова ν_t, $t \geq 0$. Вектор-строка $\boldsymbol{\theta}$ стационарного распределения этой цепи является решением системы уравнений $\boldsymbol{\theta}(D_0 + D_1) = \mathbf{0}$, $\boldsymbol{\theta}\mathbf{e} = 1$, где $\mathbf{e}$ – вектор-столбец, состоящий из единиц, $\mathbf{0}$-вектор-строка, состоящая из нулей. Интенсивность λ поступления заявок в стационарном MAP-потоке имеет вид $\lambda = \boldsymbol{\theta} D_1 \mathbf{e}$. Коэффициент вариации c_{var} длины интервала между моментами поступления последовательных групп определяется формулой $c_{var}^2 = 2\lambda\boldsymbol{\theta}(-D_0)^{-1}\mathbf{e} - 1$. Коэффициент корреляции c_{cor} длин двух соседних интервалов вычисляется следующим образом: $c_{cor} = (\lambda\boldsymbol{\theta}(-D_0)^{-1}D_1(-D_0)^{-1}\mathbf{e} - 1)/c_{var}^2$. Более подробное описание $BMAP$-потока можно найти в [1].

Время обслуживания заявки на приборе имеет PH (Phase type) распределение с неприводимым представлением $(\boldsymbol{\beta}, S)$ и управляющим процессом m_t, $t \geq 0$, с пространством состояний $\{1, \dots, M, M+1\}$, где состояние $M+1$ является поглощающим. Это означает следующее. Время обслуживания интерпретируется как время, за которое цепь Маркова m_t, $t \geq 0$, достигнет поглощающего состояния $M + 1$. Переходы цепи m_t, $t \geq 0$, в пространстве состояний $\{1, \dots, M\}$ задаются субгенератором S, а интенсивности переходов в поглощающее состояние задаются вектором $\boldsymbol{S}_0 = -S\mathbf{e}$. Когда обслуживание начинается, состояние процесса m_t, $t \geq 0$, выбирается в пространстве состояний $\{1, \dots, M\}$ на основании вероятностного вектора-строки $\boldsymbol{\beta}$. Полагаем, что матрица $S + \boldsymbol{S}_0\boldsymbol{\beta}$ неприводима. Интенсивность обслуживания задается как $\mu = -(\boldsymbol{\beta}S^{-1}\mathbf{e})^{-1}$, среднее время обслуживания $b_1 = \mu^{-1}$.

Дисциплина выбора из очереди на обслуживание FIFO.

Цель работы: вычислить стационарное распределение вероятностей состояний системы (векторы $\mathbf{p}_i, i \geq 0$) и основные характеристики производительности. Построить графики зависимостей характеристик от параметров системы.

2 Цепь Маркова, описывающая функционирование системы

Пусть в момент t

• i_t - число заявок в системе, $i_t \geq 0$,

• m_t - состояние управляющего процесса PH - обслуживания на занятом приборе, $m_t = \overline{1, M}$;

• ν_t - состояние управляющего процесса входящего MAP потока, $\nu_t = \overline{0, W}$.

Процесс функционирования системы описывается неприводимой цепью Маркова ξ_t с пространством состояний

$$X = \{(i, \nu),\, i = 0, \nu = \overline{0, W}\} \bigcup$$
$$\{(i, \nu, m),\, i \geq 1, \nu = \overline{0, W}, m = \overline{1, M}\}.$$

Далее будем предполагать, что состояния цепи $\xi_t, t \geq 0$, упорядочены в лексикографическом порядке. Обозначим через $Q_{i,j}$ матрицу интенсивностей переходов цепи из состояний, соответствующих значению i первой (счетной) компоненты в состояния, соответствующие значению j этой компоненты, $i, j \geq 0$.

Лемма 1. Рассматриваемая цепь Маркова ξ_t, $t \geq 0$, является векторным процессом гибели и размножения (QBD). Инфинитезимальный генератор Q этой цепи имеет блочную трехдиагональную структуру

$$Q = \begin{pmatrix} Q_{0,0} & Q_{0,1} & O & O & \dots \\ Q_{1,0} & Q_1 & Q_2 & O & \dots \\ O & Q_0 & Q_1 & Q_2 & \dots \\ O & O & Q_0 & Q_1 & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix},$$

где ненулевые блоки имеют следующий вид:

$$Q_{0,0} = D_0, \quad Q_{0,1} = D_1,$$
$$Q_{1,0} = I_{\bar{W}} \otimes \boldsymbol{S}_0,$$
$$Q_0 = I_{\bar{W}} \otimes \boldsymbol{S}_0 \boldsymbol{\beta}, \qquad Q_1 = D_0 \oplus S, \qquad Q_2 = D_1 \otimes I_M,$$

где $\otimes$, $\oplus$ - символы кронекерова произведения и суммы матриц соответственно (см. [2]), $\bar{W} = W + 1$.

Следствие 1. Цепь Маркова $\xi_t, t \geq 0$, принадлежит классу квазитеплицевых цепей с непрерывным временем (КТЦМ), см. [3].

Доказательство следует из вида генератора, заданного леммой 1, и определения КТЦМ, данного в [3].

3 Условие эргодичности

Теорема 1. Цепь Маркова ξ_t эргодична тогда и только тогда, когда выполняется неравенство

$$\rho = \frac{\lambda}{\mu} < 1.$$

Здесь
$$\lambda = \boldsymbol{\theta} D_1 \mathbf{e},$$
где $\boldsymbol{\theta}$–вектор-строка порядка $\bar{W}$, который находится как единственное решение СЛАУ
$$\boldsymbol{\theta}(D_0 + D_1) = \mathbf{0},\ \boldsymbol{\theta}\mathbf{e} = 1,$$

$$\mu = -(\boldsymbol{\beta} S^{-1}\mathbf{e})^{-1}.$$

4 Стационарное распределение

Предполагаем, что условие эргодичности, заданное теоремой 1 выполняется, что гарантирует, что в рассматриваемой системе не накапливается бесконечной очереди и существуют стационарные вероятности состояний системы, задаваемые как

$$p(0,\nu) = \lim_{t\to\infty} P\{i_t = 0, \nu_t = \nu\},\ n = 0,1,\ \nu = \overline{0,W};$$

$$p(i,\nu,m) = \lim_{t\to\infty} P\{i_t = i,\ \nu_t = \nu, m_t = m\},\ i \geq 1,\ n = 0,1, \nu = \overline{0,W},\ m = \overline{1,M}.$$

Упорядочим вероятности в лексикографическом порядке компонент и сформируем векторы этих вероятностей $\mathbf{p}_i,\ i \geq 0$.

Порядки этих векторов равны:

$\mathbf{p}_i \sim\ \bar{W}(1+M),\ i \geq 1.$

Векторы стационарных вероятностей $\mathbf{p}_i, i \geq 0$, вычисляются по следующему алгоритму (см. [3]).

АЛГОРИТМ

1) Находим матрицу G как единственное минимальное неотрицательное решение уравнения
$$Q_0 + Q_1 G + Q_2 G^2 = O.$$

   Замечание 1. Это матричное уравнение можно решать методом итераций
$$G^{(n+1)} = (-Q_1)^{-1}[Q_0 + Q_2(G^{(n)})^2],$$
   где $G^{(0)} = I_{\bar{W}(1+M)}$. Останавливаемся, когда становится $\|G^{(n+1)} - G^{(n)})\| < \varepsilon_G$.

2) Находим матрицу G_0 из уравнения обратной рекурсии:
$$G_0 = -\left(Q_1 + Q_2 G\right)^{-1} Q_{1,0}.$$

3) Вычисляем матрицы $\bar{Q}_{i,j}$ по формулам

$$\bar{Q}_{i,i} = \begin{cases} Q_{0,0} + Q_{01}G_0, \ i = 0, \\ Q_1 + Q_2 G, \ i \geq 1, \end{cases}$$

$$\bar{Q}_{i,i+1} = \begin{cases} Q_{0,1}, \ i = 0, \\ Q_2, \ i \geq 1. \end{cases}$$

4) Находим матрицы F_i из рекуррентных соотношений:

$$F_0 = I, F_i = F_{i-1}\bar{Q}_{i-1,i}\left(-\bar{Q}_{i,i}\right)^{-1}, i \geq 1.$$

5) Вычисляем вектор $\boldsymbol{p}_0$ как единственное решение СЛАУ:

$$\boldsymbol{p}_0(-\bar{Q}_{0,0}) = 0, \ \boldsymbol{p}_0 \sum_{i=0}^{\infty} F_i \mathbf{e} = 1.$$

6) Вычисляем векторы $\boldsymbol{p}_i$ по формулам $\boldsymbol{p}_i = \boldsymbol{p}_0 F_i, i \geq 0$.

Замечание. При выполнении шагов 3-6 как-то нужно выбрать значение i, при котором мы заканчиваем счет. Чтобы это сделать, мы должны понимать, что при возрастании i норма матриц F_i убывает. Нам надо, чтобы эта норма была очень малой. Поэтому в качестве предельного значения i, при котором мы заканчиваем счет, берем такое, при котором уже будет выполняться неравенство $\|F_i\| < \epsilon_F$.

Преимуществом этого алгоритма является отсутствие операции вычитания в рекурсиях. Все матрицы, фигурирующие в алгоритме, являются неотрицательными. Это обеспечивает устойчивость вычислений при компьютерной реализации алгоритма.

5 Характеристики производительности системы

Вычислив векторы стационарных вероятностей $\mathbf{p}_i$, $i \geq 1$, можно вычислить также различные характеристики производительности системы.

- Вероятность того, что в системе находится i заявок $\quad p_i = \mathbf{p}_i \mathbf{e}, i \geq 0$.
- Среднее число заявок в системе $\quad L = \sum\limits_{i=1}^{\infty} i p_i$.
- Вероятность того, что поступившая заявка застанет i заявок в системе

$$p_i^{(arrival)} = \frac{1}{\lambda} \mathbf{p}_i D_1 \mathbf{e}.$$

- Среднее время пребывания

$$\bar{w} = -\sum_{i=1}^{\infty} \mathbf{p}_i (\mathbf{e}_{\bar{W}} \otimes I_M) S^{-1} \mathbf{e}_M + L b_1.$$

6 Численные примеры

Исходные данные должны вводится с экрана. К ним относятся:

1. W;
2. Матрицы D_0, D_1;
3. Число M;
4. Вектор-строка $\boldsymbol{\beta}$, матрица S.
5. Точность ε_G.

Для отладки взять

1. $W = 1$;
2. $D_0 = \begin{pmatrix} -8 & 1 \\ 1 & -11 \end{pmatrix}$, $D_1 = \begin{pmatrix} 2 & 5 \\ 4 & 6 \end{pmatrix}$.
3. $M = 2$;
4. $\boldsymbol{\beta} = (1, 0)$, $S = \begin{pmatrix} -30 & 30 \\ 0 & -30 \end{pmatrix}$.
5. $\varepsilon_G = 10^{-8}$.

Список литературы

[1] Lucantoni D.M. (1991). New results on the single server queue with a batch Markovian arrival process. *Communications in Statistics-Stochastic Models*. Vol. **7**, pp. 1-46.

[2] Graham A. Kronecker Products and Matrix Calculus with Applications. Cichester: Ellis Horwood, 1981.

[3] Klimenok V.I., Dudin A.N. (2006). Multi-dimensional asymptotically quasi-Toeplitz Markov chains and their application in queueing theory. *Queueing Systems*. Vol. **54**, pp. 245-259.