МАРКОВСКАЯ СИСТЕМА МАССОВОГО ОБСЛУЖИВАНИЯ КОНЕЧНОЙ ЕМКОСТИ С ОТРИЦАТЕЛЬНЫМИ ЗАЯВКАМИ В ДИСКРЕТНОМ ВРЕМЕНИ

П. Бочаров, Е. Вискова, Э. Надаев

Российский университет дружбы народов
Москва, Россия
pbocharov;eviskova@sci.pfu.edu.ru

Рассматривается однолинейная система массового обслуживания с конечным накопителем, марковским потоком заявок, потоком Бернулли "отрицательных" заявок и марковским обслуживанием, функционирующая в дискретном времени. Выведен рекуррентный матричный алгоритм для вычисления стационарных вероятностей состояний системы, а также получено выражение для вероятности потерь заявок из-за переполнения буфера.

Ключевые слова: марковский поток, марковское обслуживание, дискретное время, отрицательные заявки, конечный накопитель.

1. ОПИСАНИЕ СИСТЕМЫ

Рассмотрим однолинейную систему массового обслуживания (СМО) с накопителем ограниченной емкости r, $1 \le r < \infty$, функционирующую в дискретном времени. Все возможные изменения в СМО происходят в моменты времени nh, $n = 1, 2, \ldots$, где h – длина фиксированного интервала времени (такта), $h > 0$.

На систему поступает дискретный марковский поток заявок, который определяется матрицами A_0 и A_1 порядка l [1, 2]. Элементы матрицы A_0 задают вероятности изменения фазы генерации заявки без ее поступления, а элементы матрицы A_1 определяют также вероятности изменения фазы генерации заявки с одновременным ее поступлением. Будем считать, что матрица $A^* = A_0 + A_1$ неразложима, а матрица A_1 отлична от нулевой.

Процесс обслуживания заявок на приборе является дискретным марковским [2] и определяется только на периодах занятости системы. Рассматривается цепь Маркова (ЦМ) с конечным числом m состояний, которая управляет процессом обслуживания заявок. Тогда, если на некотором такте на приборе находится заявка и фаза ее обслуживания равна i, $i = \overline{1,m}$, то за такт с вероятностью $B_{0,ij}$ фаза изменится на j, $j = \overline{1,m}$, и заявка продолжит обслуживаться на этой фазе, и с вероятностью $B_{1,ij}$ фаза изменится на j, $j = \overline{1,m}$, но обслуживание заявки закончится, и она покинет систему. Начальная фаза обслуживания определяется одним из двух способов: 1) в качестве начальной фазы выбирается фаза, на которой оборвалась предыдущая ЦМ, управляющая процессом обслуживания (режим 1), и 2) начальная фаза выбирается с вероятностью f_j, $j = \overline{1,m}$, $\sum_{j=1}^{m} f_j = 1$, независимо от функционирования предыдущего процесса, управляющего обслуживанием (режим 2). Матрицы из элементов $B_{0,ij}$ и $B_{1,ij}$ будем обозначать через B_0 и B_1, соответственно. Матрица $B^* = B_0 + B_1$ предполагается неразложимой, а матрица B_1 – ненулевой.

Помимо обычных (положительных) заявок в систему поступает поток Бернулли отрицательных заявок [3, 4]. Отрицательная заявка с вероятностью a^- поступает в систему, а с вероятностью $\bar{a}^- = 1 - a^-$ нет. Поступившая за такт отрицательная заявка вытесняет из СМО последнюю ожидающую в накопителе положительную заявку и покидает систему, не получая никакого обслуживания; если в накопителе отсутствуют заявки, то отрицательная заявка уничтожает положительную заявку, находящуюся на приборе, и также покидает систему. В случае отсутствия в системе положительных заявок, как в очереди, так и на приборе, поступившая отрицательная заявка оставляет систему, не оказывая на нее никакого воздействия.

В качестве n-го такта рассмотрим полуинтервал $[(n-1)h, nh)$ и примем, что события на этом такте происходят в следующем порядке: сначала покидает систему заявка, обслуживание которой завершилось на n-м такте; затем из очереди выбирается заявка и сразу отправляется на прибор; после этого на систему оказывает воздействие отрицательная заявка, если таковая поступила за n-й такт; далее в очередь помещается обычная заявка, генерация которой завершилась на n-м такте; после чего фиксируется состояние системы и начинается новый $(n + 1)$-й такт.

Выбор заявок на обслуживание производится в порядке их поступления, т. е. в соответствии с дисциплиной FCFS.

Положительные и отрицательные заявки, заставшие в момент их поступления в систему занятыми все места в накопителе, теряются, при этом потерянные заявки более не возобновляются и не оказывают никакого влияния на систему и входящий поток.

Функционирование рассматриваемой системы можно описать однородной ЦМ $\{\gamma_n,\ n \geq 0\}$ по моментам времени nh с множеством состояний

$$X = \bigcup_{k=0}^{R} X_k,$$

где:

для режима 1 $\quad X_k = \{(i,j,k),\ \ i = \overline{1,l},\ \ j = \overline{1,m}\},\ k = \overline{0,R};$

для режима 2 $\quad X_0 = \{(i,0),\ \ i = \overline{1,l}\},\ \ X_k = \{(i,j,k),\ \ i = \overline{1,l},\ \ j = \overline{1,m}\},\ k = \overline{1,R}$, а $R = r + 1$.

Для режимов 1 и 2 состояние (i,j,k) соответствует такой ситуации, когда в СМО имеется k заявок, новая заявка генерируется на фазе i, на приборе заявка обслуживается на фазе j. Для режима 2 состояние $(i,0)$ означает, что система пуста и новая заявка генерируется на фазе i.

2. РЕШЕНИЕ СИСТЕМЫ УРАВНЕНИЙ РАВНОВЕСИЯ

При сделанных предположениях легко показать, что ЦМ $\{\gamma_n, \ n \geq 0\}$ неприводима, следовательно, финальное распределение

$$p_x = \lim_{n\to\infty} P\{\gamma_n = x\}, \quad x \in X,$$

существует, не зависит от начального распределения и совпадает со стационарным распределением вероятностей.

Положим

$$\boldsymbol{p}^T = (\boldsymbol{p}_0^T, \boldsymbol{p}_1^T, ..., \boldsymbol{p}_R^T), \tag{1}$$

где $\boldsymbol{p}_k^T = (p_{11k}, p_{12k}, ..., p_{1mk}, p_{21k}, ..., p_{2mk}, ..., p_{l1k}, ..., p_{lmk})$, $k = \overline{0,R}$ для режима 1 и $k = \overline{1,R}$ для режима 2, и дополнительно для режима 2 введем вектор $\boldsymbol{p}_0^T = (p_{10}, p_{20}, ..., p_{l0})$.

Стационарные вероятности $\boldsymbol{p}_k$, $k = \overline{0,R}$, образуют единственное решение системы уравнений равновесия:

для режима 1

$$\boldsymbol{p}_0^T = \boldsymbol{p}_0^T(A_0 \otimes I_m) + \boldsymbol{p}_1^T(A_0 \otimes B_1 + a^- A_0 \otimes B_0) + \boldsymbol{p}_2^T(a^- A_0 \otimes B_1), \tag{2}$$

$$\boldsymbol{p}_1^T = \boldsymbol{p}_0^T(A_1 \otimes I_m) + \boldsymbol{p}_1^T(\bar{a}^- A_0 \otimes B_0 + a^- A_1 \otimes B_0 + A_1 \otimes B_1) + \boldsymbol{p}_2^T(a^- A_0 \otimes B_0 +$$
$$+ \bar{a}^- A_0 \otimes B_1 + a^- A_1 \otimes B_1) + \boldsymbol{p}_3^T(a^- A_0 \otimes B_1); \tag{3}$$

для режима 2

$$\boldsymbol{p}_0^T = \boldsymbol{p}_0^T A_0 + \boldsymbol{p}_1^T(A_0 \otimes \boldsymbol{b}_1 + a^- A_0 \otimes \boldsymbol{b}_0) + \boldsymbol{p}_2^T(a^- A_0 \otimes \boldsymbol{b}_1), \tag{2'}$$

$$\boldsymbol{p}_1^T = \boldsymbol{p}_0^T(A_1 \otimes \boldsymbol{f}^T) + \boldsymbol{p}_1^T(\bar{a}^- A_0 \otimes B_0 + a^- A_1 \otimes B_0 + A_1 \otimes B_1) + \boldsymbol{p}_2^T(a^- A_0 \otimes B_0 +$$
$$+ \bar{a}^- A_0 \otimes B_1 + a^- A_1 \otimes B_1) + \boldsymbol{p}_3^T(a^- A_0 \otimes B_1); \tag{3'}$$

для обоих режимов

$$\boldsymbol{p}_k^T = \boldsymbol{p}_{k-1}^T(\bar{a}^- A_1 \otimes B_0) + \boldsymbol{p}_k^T(\bar{a}^- A_0 \otimes B_0 + a^- A_1 \otimes B_0 + \bar{a}^- A_1 \otimes B_1) +$$

$$+ \boldsymbol{p}_{k+1}^T(a^- A_0 \otimes B_0 + \bar{a}^- A_0 \otimes B_1 + a^- A_1 \otimes B_1) + u(r-k)\boldsymbol{p}_{k+2}^T(a^- A_0 \otimes B_1), \quad k = \overline{2,r}, \tag{4}$$

$$\boldsymbol{p}_R^T = \boldsymbol{p}_r^T(\bar{a}^- A_1 \otimes B_0) + \boldsymbol{p}_R^T(\bar{a}^- A_0 \otimes B_0 + A_1 \otimes B_0 + \bar{a}^- A_1 \otimes B_1) \tag{5}$$

с условием нормировки

$$\sum_{k=0}^{R} \boldsymbol{p}_k^T \mathbf{1} = 1. \tag{6}$$

Здесь $\boldsymbol{b}_1 = B_1\mathbf{1}$, I_m – единичная матрица порядка m, $u(x)$ – функция Хевисайда, $V \otimes W$ – кронекерово произведение матриц V и W, а $\mathbf{1}$ – вектор из единиц соответствующей контексту размерности.

Для решения СУР (2)–(5) будем использовать метод "урезания" множества состояний с последующим склеиванием оставшейся траектории исходной цепи Маркова [5]. Исходя из СУР (2)–(5), матрица $P = (P_{kq})_{k,q=\overline{1,R}}$ переходных вероятностей ЦМ

$\{\gamma_n, \quad n \geq 0\}$ представляется в блочном виде и имеет четыре диагонали ненулевых элементов:

$$P = \begin{pmatrix} P_{00} & P_{01} & 0 & 0 & \dots & 0 & 0 \\ P_{10} & P_{11} & P_{12} & 0 & \dots & 0 & 0 \\ P_{20} & P_{21} & P_{22} & P_{23} & \dots & 0 & 0 \\ 0 & P_{31} & P_{32} & P_{33} & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & P_{rr} & P_{rR} \\ 0 & 0 & 0 & 0 & \dots & P_{Rr} & P_{RR} \end{pmatrix},$$

где:

для режима 1

$$P_{00} = A_0 \otimes I_m, \quad P_{01} = A_1 \otimes I_m, \quad P_{10} = A_0 \otimes B_1 + a^- A_0 \otimes B_0, \quad P_{20} = a^- A_0 \otimes B_1;$$

для режима 2

$$P_{00} = A_0, \quad P_{01} = A_1 \otimes \boldsymbol{f}^T, \quad P_{10} = A_0 \otimes \boldsymbol{b}_1 + a^- A_0 \otimes \boldsymbol{b}_0, \quad P_{20} = a^- A_0 \otimes \boldsymbol{b}_1;$$

для обоих режимов

$$P_{11} = \bar{a}^- A_0 \otimes B_0 + a^- A_1 \otimes B_0 + A_1 \otimes B_1,$$

$$P_{kk} = \bar{a}^- A_0 \otimes B_0 + a^- A_1 \otimes B_0 + \bar{a}^- A_1 \otimes B_1, \quad k = \overline{2, r},$$

$$P_{k,k-1} = a^- A_0 \otimes B_0 + \bar{a}^- A_0 \otimes B_1 + a^- A_1 \otimes B_1, \quad k = \overline{2, R},$$

$$P_{k,k-2} = a^- A_0 \otimes B_1, \quad k = \overline{3, R},$$

$$P_{k,k+1} = \bar{a}^- A_1 \otimes B_0, \quad k = \overline{1, r},$$

$$P_{RR} = \bar{a}^- A_0 \otimes B_0 + A_1 \otimes B_0 + \bar{a}^- A_1 \otimes B_1.$$

Согласно методу, изложенному в [5], сначала с точностью до нормировочной константы, находится решение СУР:

$$\boldsymbol{p}_0^T = \boldsymbol{p}_0^T P_{00}^{(R)}, \tag{7}$$

где матрица $P_{00}^{(R)}$ находится с помощью рекуррентного матричного алгоритма:

$$P_{kk}^{(r-k+1)} = (I - P_{kk}^{(r-k)})^{-1}, \quad P_{k,k-1}^{(r-k+1)} = P_{kk}^{(r-k+1)} P_{k,k-1}, P_{k-1,k-1}^{(r-k+1)} = P_{k-1,k-1} + P_{k-1,k}^{(r-k)} \times$$

$$\times P_{k-1,k}^{(r-k+1)}, \quad P_{k-2,k-1}^{(r-k+1)} = P_{k-2,k-1}, P_{-1}^{(r-k+1)} \equiv P_{-2}^{(r-k+1)} \equiv 0, \quad k = \overline{0, R}.$$

Затем последовательно (также с точностью до нормировочной константы) вычисляются $\boldsymbol{p}_k$, по формуле

$$\boldsymbol{p}_k^T = (\sum_{s=0}^{k-1} \boldsymbol{p}_s^T P_{sk}^{(R-k)}) P_{kk}^{(R-k+1)}, \quad k = \overline{1, R}. \tag{8}$$

Наконец, на последнем этапе производится нормировка полученных вероятностей в соответствии с условием нормировки (6).

3. ВЕРОЯТНОСТЬ ПОТЕРИ ЗАЯВОК ИЗ-ЗА ПЕРЕПОЛНЕНИЯ НАКОПИТЕЛЯ

Для расчета стационарной вероятности потери заявки в результате переполнения накопителя найдем стационарные вероятности состояний СМО в моменты непосредственно перед завершением генерации положительных заявок и их поступлением в систему, исключая те моменты, когда за рассматриваемый такт произойдет окончание обслуживания заявки на приборе или поступление отрицательной заявки.

Обозначим через p^-_{x,A^+} стационарную вероятность состояния $x \in X^-_{A^+}$, где множество $X^-_{A^+}$ состоит из тех же состояний, что и множество X, однако индексы i, j и k в состояниях из множества $X^-_{A^+}$ обозначают соответственно фазу генерации, фазу обслуживания и число заявок в системе непосредственно перед поступлением положительной заявки в систему.

Введем векторы $\boldsymbol{p}^-_{k,A^+}$, совпадающие по размерности и структуре с векторами $\boldsymbol{p}_k$, $k = \overline{0,R}$. Далее определим вероятность $a^+_{(\bar{s})}$ поступления положительной заявки в систему за такт при незавершенном обслуживании и непоступлении отрицательной заявки на этом такте:

для режима 1

$$a^+_{(\bar{s})} = \boldsymbol{p}_0^T \bar{a}^-(\boldsymbol{a}_1 \otimes \mathbf{1}) + \sum_{k=1}^{R} \boldsymbol{p}_k^T \bar{a}^-(\boldsymbol{a}_1 \otimes \boldsymbol{b}_0); \tag{9}$$

для режима 2

$$a_{(\bar{s})} = \boldsymbol{p}_0^T \bar{a}^- \boldsymbol{a}_1 + \sum_{k=1}^{R} \boldsymbol{p}_k^T \bar{a}^-(\boldsymbol{a}_1 \otimes \boldsymbol{b}_0). \tag{9$'$}$$

Найдем теперь стационарные вероятности состояний СМО в моменты непосредственно перед поступлением заявки в систему при незавершенном обслуживании и непоступлении отрицательной заявки на данном такте.

Заметим, что поступающая в СМО заявка, генерация которой завершилась на n-м такте, застанет в ней в момент непосредственно перед поступлением k других заявок, если на n-м такте в СМО находилось k заявок и за n-й такт не закончилось обслуживание заявки на приборе и не поступила отрицательная заявка. Кроме того, заметим, что стационарная вероятность p^-_{x,A^+} представляет собой условную вероятность состояния $x \in X^-_{A^+}$ системы при условии, что в данный момент поступит новая заявка и на предыдущем такте не завершится обслуживание заявки, находящейся на приборе, и не поступит отрицательная заявка. С учетом этого, а также принимая во внимание (9) и (9′), получим:

для режима 1

$$\boldsymbol{p}^{T-}_{0,A^+} = \frac{1}{a^+_{(\bar{s})}} \boldsymbol{p}_0^T \bar{a}^-(A_1 \otimes I_m); \tag{10}$$

для режима 2

$$\boldsymbol{p}^{T-}_{0,A^+} = \frac{1}{a^+_{(\bar{s})}} \boldsymbol{p}_0^T \bar{a}^- A_1; \tag{10$'$}$$

для обоих режимов

$$p^{T-}_{k,A^+} = \frac{1}{a^+_{(\bar{s})}} p^T_k \bar{a}^-(A_1 \otimes B_0), \quad k = \overline{1,R}. \tag{11}$$

Рассмотрим механизм потери заявки. Если в момент $nh+0$ СМО находится в одном из состояний множества X_R, то поступившая на $(n+1)$-м такте заявка потеряется лишь при условии, что находящаяся на приборе на этом такте заявка не обслужится и не поступит отрицательная заявка. Таким образом, стационарная вероятность π потери заявки из-за переполнения накопителя для обоих режимов определяется следующим выражением:

$$\pi = \frac{1}{a^+_{(\bar{s})}} p^T_R \bar{a}^-(\boldsymbol{a}_1 \otimes \boldsymbol{b}_0), \tag{12}$$

где $a^+_{(\bar{s})}$ определяется формулами (10), (10′) и (11).

ЛИТЕРАТУРА

1. *Blondia C.* A discrete-time batch Markovian arrival process as B-ISDN traffic model // Belgian J. of Operations Research, Statistics and Computer Science. 1993. V. 32 (3,4). P. 3–23.
2. *Bocharov P. P., Viskova E. V.* The DMAP/DMSP/1/r queueing system in discrete time // Trans. of XXIV Int. Seminar on Stability Problems for Stochastic Models, Latvia, 2004. P. 103–108.
3. *Gelenbe E.* Queueing networks with negative and positive customers // J. Appl. Prob. 1991. V. 28. P. 656–663.
4. *Gelenbe E., Glynn P., Sigman K.* Queues with negative arrivals // J. Appl. Prob. 1991. V. 28. P. 245–250.
5. *Bocharov P. P., D'Apice C., Pechinkin A. V., Salerno S.* Queueing theory. Ultrecht-Boston: VSP, 2004.