

Однолинейная система массового обслуживания с фоновыми заявками в дискретном времени

П.П.Бочаров*, Л.О.Шлумпер**

*Российский университет дружбы народов, Москва, Россия

**Московский физико-технический институт, Москва, Россия

Поступила в редколлегию 09.08.2005

Аннотация—Исследуется однолинейная система массового обслуживания с потерями в дискретном времени. На систему поступают заявки двух типов с относительным приоритетом. Основные заявки образуют марковский поток и ожидают начала обслуживания в накопителе ограниченной емкости. Фоновые заявки, поступают из бункера, в котором их запас не ограничен. Обслуживание заявок обоих типов также является марковским. Фоновые заявки поступают на прибор лишь в случае, если система пуста. Выведен эффективный матричный алгоритм для расчёта стационарных вероятностей состояний данной системы и найдены некоторые показатели ее производительности в стационарном режиме.

1. ВВЕДЕНИЕ

Рассматривается однолинейная система массового обслуживания (СМО) с потерями, функционирующая в дискретном времени с фиксированной величиной такта h , $h > 0$. На систему поступают два потока заявок: основной поток заявок, которые будем называть 1-заявками, и фоновый поток заявок, называемых 2-заявками. 1-заявки ожидают начала обслуживания в накопителе ограниченной емкости r , $1 \leq r < \infty$.

Поток 1-заявок является дискретным марковским (discrete Markov arrival process — DMAP) [1], задаваемым матрицами A_0 и A_1 порядка l , элементы $(A_0)_{i,j}$ и $(A_1)_{i,j}$ которых задают переходы цепи Маркова (ЦМ), управляющей генерацией 1-заявок (в дальнейшем, для краткости, мы иногда будем говорить просто о процессе генерации 1-заявок), с фазы генерации i на фазу генерации j , не сопровождающиеся и сопровождающиеся поступлением новой заявки, соответственно.

Для дальнейших выкладок необходимо определить следующие характеристики входящего потока. Введём вектор $\mathbf{a}_1 = (a_{1,1}, \dots, a_{1,l})^T$, где $a_{1,i}$ — вероятность поступления новой 1-заявки на фазе генерации i . Также введём вектор $\mathbf{a}_0 = (a_{0,1}, \dots, a_{0,l})^T$, где $a_{0,i}$ — вероятность перехода процесса генерации основных заявок с фазы i на какую-либо другую фазу без поступления новой 1-заявки. Очевидно, что

$$\mathbf{a}_1 = A_1 \mathbf{1}_l, \quad \mathbf{a}_0 = \mathbf{1}_l - \mathbf{a}_1;$$

здесь и далее $\mathbf{1}_l$ — вектор размерности l , состоящий из единиц; если индекс у вектора $\mathbf{1}$ опущен, то его размерность определяется контекстом.

Обозначим через A^* матрицу вероятностей переходов управляющей ЦМ для процесса генерации основных заявок. Заметим, что $A^* = A_0 + A_1$. Тогда вектор \mathbf{p}^* стационарных вероятностей этой ЦМ определяется из системы уравнений равновесия (СУР)

$$\mathbf{p}^{*T} A^* = \mathbf{p}^{*T}$$

с условием нормировки

$$\mathbf{p}^{*T} \mathbf{1}_l = 1.$$

Заметим, что элемент p_i^* , $i = \overline{1, l}$, этого вектора совпадает со стационарной вероятностью пребывания процесса генерации основных заявок на фазе i .

Получив векторы \mathbf{a}_1 и \mathbf{p}^* , можно вычислить вероятность a поступления заявки за такт:

$$a = \mathbf{p}^{*T} \mathbf{a}_1.$$

Источником заявок фонового потока (2-заявок) является так называемый бункер, в котором запас заявок не ограничен (такой поток заявок иногда называют насыщенным). При этом фоновая заявка поступает на прибор лишь в случае, когда после завершения обслуживания заявки система свободна, т.е. накопитель пуст и не закончилась генерация новой основной заявки. Если после окончания обслуживания фоновой заявки система снова будет пуста, то на прибор немедленно поступает следующая фоновая заявка из бункера. Таким образом, процесс обслуживания фоновых заявок продолжается до тех пор, пока по завершению обслуживания очередной фоновой заявки в системе не окажется хотя бы одна основная заявка. При этом поступившая основная заявка не прерывает обслуживания фоновой заявки. Другими словами, основные заявки имеют относительный приоритет при обслуживании перед фоновыми заявками.

Процессы обслуживания i -заявок также, как и процесс поступления основных заявок, являются марковскими (discrete Markov service processes — DMSP) [1] и задаются матрицами $B_0^{(i)}$ и $B_1^{(i)}$ порядка m_i , $i = 1, 2$. Элементы $(B_0^{(i)})_{j,s}$ и $(B_1^{(i)})_{j,s}$ этих матриц задают вероятности перехода обрывающихся ЦМ, управляющих процессами обслуживания (основных и фоновых) заявок, с фазы обслуживания j на фазу обслуживания s без завершения обслуживания или с завершением обслуживания заявки на приборе, соответственно. Кроме того, задаются векторы $\mathbf{f}_i^T = (f_{i,1}, \dots, f_{i,m_i})$, $i = 1, 2$, элементы которых определяют вероятности выбора фазы обслуживания, с которой начинается процесс обслуживания новой заявки при смене типа заявки на приборе.

В дальнейшем будем предполагать, что матрицы A^* и $B^{*(i)} = B_0^{(i)} + B_1^{(i)}$, $i = 1, 2$, неразложимы, а матрицы A_1 и $B_1^{(i)}$, $i = 1, 2$, отличны от нулевых. Также в дальнейшем будет использоваться следствие этих условий — невырожденность матриц $I_l - A_1$ и $I_{m_i} - B_1^{(i)}$, $i = 1, 2$.

Для построения корректной математической модели нам необходимо ввести порядок следования событий, происходящих за один такт. Будем считать, что все события в системе происходят в моменты времени nh , $n = 1, 2, \dots$, кратные такту. При этом n -м тактом будем называть промежуток времени $[(n-1)h, nh]$ и примем следующий порядок событий на такте n :

в момент $nh - 0$ обслуженная на такте n заявка покидает систему;

в момент nh , если прибор не занят и очередь не пуста, то основная заявка из накопителя, выбранная в соответствии с дисциплиной FCFS, поступает на прибор;

в момент $nh + 0$ в систему поступает 1-заявка (если на такте n закончилась её генерация) и помещается в накопитель, если прибор занят, или на прибор, если прибор свободен; если на такте n не окончилась генерация новой 1-заявки, и прибор в момент $nh + 0$ не занят, то на прибор в этот момент поступает фоновая заявка.

Описанную СМО в обозначениях Кендалла будем классифицировать как $(DMAP, ST)/(DMSP, DMSP)/1/r$ (символ “ST” — сокращение от английского saturated). Данная система является дискретным аналогом системы $(MAP, ST)/(MSP, MSP)/1/r$, рассмотренной в [2].

Как и её непрерывный аналог, система $(DMAP, ST)/(DMSP, DMSP)/1/r$ может быть эффективно использована для моделирования широкого спектра процессов в современных сетевых технологиях. СМО с фоновыми заявками описывают также поведение автоматизированной информационной системы (АИС) с оперативной реорганизацией баз данных (БД) [3, 4]. Такие системы управления БД поддерживают совместное обслуживание запросов и реорганизацию БД с относительным приоритетом запросов перед задачами реорганизации. Следует заметить, что в виду дискретности процессов обработки данных в ЭВМ СМО, функционирующие в дискретном времени, более адекватно описывают поведение таких систем.

Поскольку в рассматриваемую СМО фоновые заявки поступают только при её освобождении от основных заявок, то данную СМО можно отнести к системам с прогулками — системам с отключением прибора при их опустошении от заявок основного потока (exhaustive service discipline).

Остановимся теперь на наиболее близких по тематике (рассматриваемой СМО) работах. В [5] исследовалась СМО $(PH, ST)/(PH, PH)/1/r$, функционирующая в непрерывном времени, с потоком основных заявок фазового типа и PH -распределениями длительностей обслуживания как основных, так и фоновых заявок. Для этой СМО в [5] получены замкнутые матричные формулы для вычисления стационарного распределения вероятностей её состояний. В [2] рассмотрена СМО $(MAP, ST)/(MSP, MSP)/1/r$, обобщающая СМО, изученную в [5], на случай, когда процесс поступления основных заявок, а также процессы обслуживания как основных, так и фоновых заявок являются марковскими. В [2] получен устойчивый рекуррентно-матричный алгоритм для расчета стационарных вероятностей состояний системы $(MAP, ST)/(MSP, MSP)/1/r$, рассматриваемой как в произвольные моменты времени, так и в моменты поступления основных заявок, а также выведены выражения для ряда стационарных показателей производительности СМО. Кроме того, в [2] получено выражение ПЛС для времени ожидания начала обслуживания основных заявок. Более обширный обзор библиографии по СМО в непрерывном времени с прогулками при опустошении системы приводится в [2].

В данной работе анализируется дискретный аналог СМО $(MAP, ST)/(MSP, MSP)/1/r$ — система $(DMAP, ST)/(DMSP, DMSP)/1/r$, функционирующая в дискретном времени. Для цепи Маркова, описывающей функционирование данной СМО, выведен рекуррентно-матричный алгоритм для вычисления стационарных вероятностей ее состояний. Получены также выражения для стационарных вероятностей состояний СМО, рассматриваемой в моменты поступления основных заявок при условии незавершения обслуживания на такте.

2. СТОХАСТИЧЕСКАЯ МОДЕЛЬ СИСТЕМЫ

Для построения стохастической модели исследуемой СМО рассмотрим однородную ЦМ $\{X_s, s = 0, 1, 2, \dots\}$, по моментам времени sh , определенную в множестве состояний

$$\mathcal{X} = \bigcup_{i=1}^2 \bigcup_{k=2-i}^{r+u(2-i)} \mathcal{X}_{ki},$$

где

$$\mathcal{X}_{ki} = \{(j, n, k, i), j = \overline{1, l}, n = \overline{1, m_i}, k = \overline{2-i, r+u(2-i)}, i = 1, 2\};$$

здесь $u(y)$ —единичная функция Хевисайда.

Т.о. состояние (j, n, k, i) рассматриваемой цепи соответствует состоянию исследуемой СМО следующим образом: процесс генерации 1-заявок находится на фазе j , на приборе обслуживается i -заявка, в системе находится k 1-заявок, и процесс обслуживания i -заявки находится на фазе n .

Заметим, что общее число 1-заявок в системе следующим образом зависит от типа заявки, находящейся на приборе. Если обслуживается 1-заявка, то максимальное число 1-заявок, которые одновременно могут находиться в системе, равно R , где $R = r + 1$; если же на приборе находится 2-заявка, то в системе могут находиться не более r 1-заявок.

При этом общее количество $k(i)$ заявок в системе, находящейся в состоянии (j, n, k, i) , определяется соотношением

$$k(i) = \begin{cases} k, & i = 1, \\ k + 1, & i = 2, \end{cases}$$

при этом $k(i) = \overline{1, R}$, $i = 1, 2$.

Множество \mathcal{X} состояний ЦМ $\{X_s, s = 0, 1, 2, \dots\}$ конечно, и его мощность $|\mathcal{X}|$ равна

$$|\mathcal{X}| = lR(m_1 + m_2).$$

Кроме того, все состояния данной ЦМ сообщаются, т.е. цепь неприводима, следовательно, финальные вероятности $\lim_{s \rightarrow \infty} \mathcal{P}\{X_s = x\} = p_x$, $x \in \mathcal{X}$, цепи $\{X_s, s = 0, 1, 2, \dots\}$ существуют, строго положительны, не зависят от начального состояния X_0 ЦМ и совпадают со стационарными вероятностями.

3. СИСТЕМА УРАВНЕНИЙ РАВНОВЕСИЯ

Для удобства записи вектор стационарных вероятностей представим в виде

$$\begin{aligned} \mathbf{p}_{ki}^T &= (p_{1,1,k,i}, \dots, p_{1,m_i,k,i}, \dots, p_{l,1,k,i}, \dots, p_{l,m_i,k,i}), \\ \mathbf{p}^T &= (\mathbf{p}_{0,2}^T, \mathbf{p}_{1,1}^T, \mathbf{p}_{1,2}^T, \mathbf{p}_{2,1}^T, \mathbf{p}_{2,2}^T, \dots, \mathbf{p}_{r,2}^T, \mathbf{p}_{R,1}^T). \end{aligned} \tag{1}$$

Вектор \mathbf{p} стационарных вероятностей состояний СМО является решением системы уравнений равновесия

$$\mathbf{p}^T P = \mathbf{0}^T, \tag{2}$$

где P — матрица интенсивностей переходов ЦМ $\{X_s, s = 0, 1, 2, \dots\}$.

С учетом порядка состояний СМО, определяемого согласно (1) векторами \mathbf{p}_{ki} и \mathbf{p} , запишем СУР (2) в развернутой форме:

$$\begin{aligned} \mathbf{p}_{0,2}^T (A_0 \otimes B_0^{(2)} + A_0 \otimes B_1^{(2)}) + \mathbf{p}_{1,1}^T (A_0 \otimes \beta_1 \mathbf{f}_2^T) &= \mathbf{p}_{0,2}^T, \\ \mathbf{p}_{1,1}^T (A_0 \otimes B_0^{(1)} + A_1 \otimes B_1^{(1)}) + \mathbf{p}_{2,1}^T (A_0 \otimes B_1^{(1)}) + \mathbf{p}_{1,2}^T (A_0 \otimes \beta_2 \mathbf{f}_1^T) + \mathbf{p}_{0,2}^T (A_1 \otimes \beta_2 \mathbf{f}_1^T) &= \mathbf{p}_{1,1}^T, \\ \mathbf{p}_{k-1,1}^T (A_1 \otimes B_0^{(1)}) + \mathbf{p}_{k,1}^T (A_0 \otimes B_0^{(1)} + A_1 \otimes B_1^{(1)} + u(k-r)(A_1 \otimes B_0^{(1)})) & \\ + u(R-k) (\mathbf{p}_{k+1,1}^T (A_0 \otimes B_1^{(1)}) + \mathbf{p}_{k-1,2}^T (A_1 \otimes \beta_2 \mathbf{f}_1^T)) + \mathbf{p}_{k,2}^T (A_0 \otimes \beta_2 \mathbf{f}_1^T) &= \mathbf{p}_{k,1}^T, \quad k = \overline{2, R}, \\ \mathbf{p}_{k-1,2}^T (A_1 \otimes B_0^{(2)}) + \mathbf{p}_{k,2}^T (A_0 \otimes B_0^{(2)} + u(k-r+1)(A_1 \otimes B_0^{(2)})) &= \mathbf{p}_{k,2}^T, \quad k = \overline{1, r}, \end{aligned} \tag{3}$$

где $U \otimes V$ — кронекерово произведение матриц U и V , $U \oplus V = U \otimes I_n + I_m \otimes V$, m и n — порядки матриц U и V , соответственно, $\beta_i = B_1^{(i)} \mathbf{1}$.

Для замкнутости системы уравнений требуется наложить на вектор стационарных вероятностей условие нормировки

$$\mathbf{p}^T \mathbf{1} = 1. \tag{4}$$

Вывод СУР (3) удобно пояснять при помощи схемы переходов цепи X_s (см. рисунок) за такт. Обоснуем сначала вывод третьего уравнения системы для фиксированного $k = \overline{2, r}$ с помощью схемы (с).

В подмножество состояний $\mathcal{X}_{k,1}$ система может попасть из подмножества $\mathcal{X}_{k-1,1}$ за счёт поступления 1-заявки с учетом того, что не будет закончено обслуживание 1-заявки, занимающей прибор. Вероятность поступления новой 1-заявки определяется матрицей A_1 , при этом матрица A_1 также определяет и фазу, с которой начнётся генерация следующей 1-заявки. Вероятность изменения фазы обслуживания 1-заявки на приборе (без завершения ее обслуживания) задаётся матрицей $B_0^{(1)}$.

Кроме того, в подмножество состояний $\mathcal{X}_{k,1}$ система может попасть из подмножества $\mathcal{X}_{k+1,1}$ за счёт окончания обслуживания 1-заявки и выбора из накопителя очередной 1-заявки, которая поступит на прибор, с учетом того, что на данном такте не завершится генерация новой 1-заявки на входе системы. Вероятность завершения обслуживания 1-заявки определяется матрицей $B_1^{(1)}$, при этом матрица $B_1^{(1)}$ также определяет и фазу, с которой начнётся обслуживание следующей 1-заявки. Вероятность изменения фазы генерации (без поступления новой 1-заявки) задается матрицей A_0 .

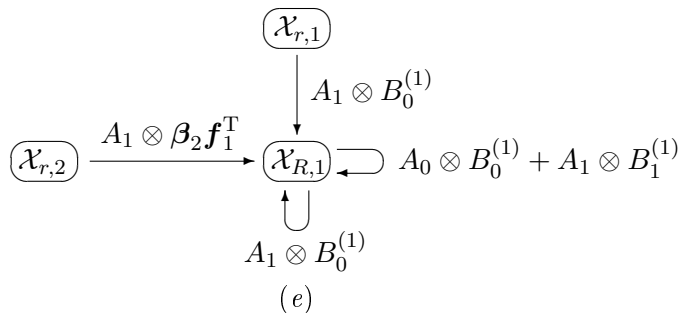
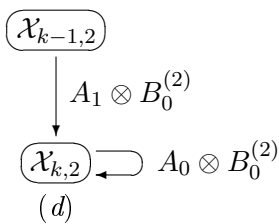
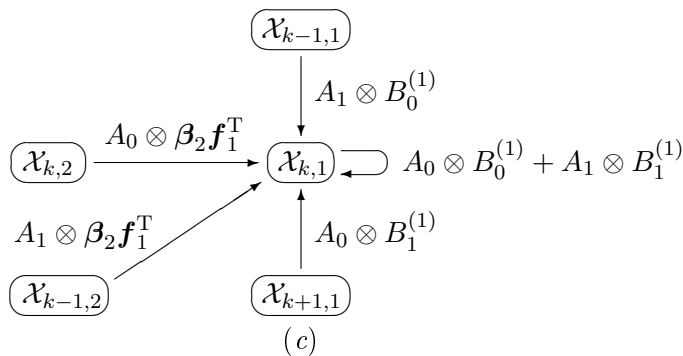
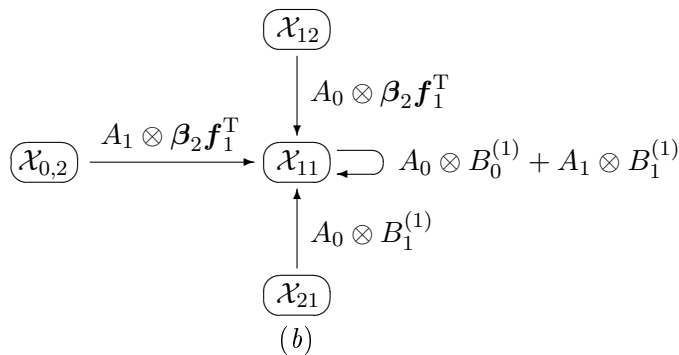
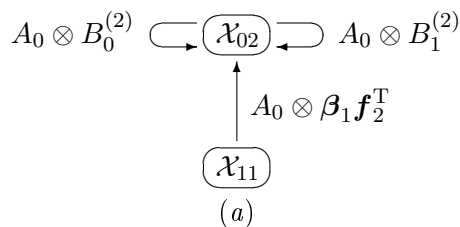
Возможен также переход системы в подмножество состояний $\mathcal{X}_{k,1}$ из подмножества $\mathcal{X}_{k,2}$ за счёт окончания обслуживания 2-заявки с учетом того, что за данный такт в систему не поступит новая 1-заявка. Вероятность завершения обслуживания 2-заявки определяется вектором β_2 при этом фаза, с которой начнётся обслуживание следующей 1-заявки, определяется в соответствии с вектором начальных вероятностей f_1 , а вероятность изменения фазы генерации 1-заявки на входе системы задается матрицей A_0 .

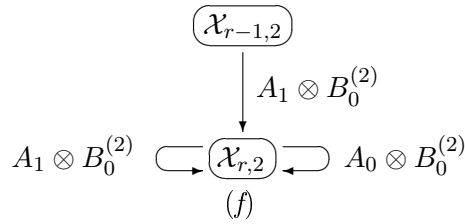
Аналогично описывается переход из макросостояния $\mathcal{X}_{k-1,2}$ в $\mathcal{X}_{k,1}$ за счёт окончания обслуживания 2-заявки и поступления новой 1-заявки. Этот переход отличается только тем, что вероятность изменения фазы генерации задается матрицей A_1 .

Кроме того, система может остаться в подмножестве состояний $\mathcal{X}_{k,1}$ на данном такте, что может осуществиться двояким образом. В первом случае за такт не закончится обслуживание 1-заявки на приборе и в систему не поступит новая 1-заявка. Вероятности изменения фазы генерации заявки на входе системы и фазы обслуживания заявки на приборе при этом задаются матрицами A_0 и $B_0^{(1)}$ соответственно. Во втором случае, наоборот, за такт происходит и окончание обслуживания 1-заявки, находившейся на приборе, и поступление новой 1-заявки; вероятности изменения фаз процессов обслуживания и генерации заявки определяются матрицами $B_1^{(1)}$ и A_1 соответственно.

Аналогичное уравнение для $k = R$ получается согласно схеме (е). При этом отличие от случая $k = \overline{2, r}$ заключается в том, что приход новой 1-заявки, если за данный такт не завершилось обслуживание заявки на приборе, также оставляет систему в исходном макросостоянии (так как поступившая 1-заявка теряется из-за переполнения накопителя). Заметим, что во время обслуживания 1-заявки в системе не может находиться более R 1-заявок, а во время обслуживания 2-заявки — не более r 1-заявок, поэтому система может перейти в макросостояние $\mathcal{X}_{R,1}$ только за счет поступления новой 1-заявки из макросостояний $\mathcal{X}_{r,1}$ (без завершения обслуживания 1-заявки, находившейся на приборе) и $\mathcal{X}_{r,2}$ (без завершения обслуживания фоновой заявки, находящейся на приборе).

Путём аналогичных рассуждений на основе схем (а) и (б) обосновываются соответственно первое и второе уравнения СУР. В свою очередь, схемы (д) и (ф) служат для пояснения вывода четвертого уравнения СУР.





4. РЕШЕНИЕ УРАВНЕНИЙ РАВНОВЕСИЯ

Для решения СУР (2) запишем матрицу P в явной форме:

$$P = \begin{matrix} & \mathcal{X}_{0,2} & \mathcal{X}_{1,1} & \mathcal{X}_{1,2} & \mathcal{X}_{2,1} & \mathcal{X}_{2,2} & \cdots & \mathcal{X}_{r,1} & \mathcal{X}_{r,2} & \mathcal{X}_{r+1,1} \\ \begin{matrix} \mathcal{X}_{0,2} \\ \mathcal{X}_{1,1} \\ \mathcal{X}_{1,2} \\ \mathcal{X}_{2,1} \\ \mathcal{X}_{2,2} \\ \vdots \\ \mathcal{X}_{r,1} \\ \mathcal{X}_{r,2} \\ \mathcal{X}_{r+1,1} \end{matrix} & \begin{pmatrix} P_{em} & H & B_2 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 \\ C_0 & P_{od} & 0 & B_1 & 0 & \cdots & 0 & 0 & 0 & 0 \\ 0 & C & P_{ev} & H & B_2 & \cdots & 0 & 0 & 0 & 0 \\ 0 & D & 0 & P_{od} & 0 & \cdots & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & C & P_{ev} & \cdots & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \cdots & P_{od} & 0 & B_1 & 0 \\ 0 & 0 & 0 & 0 & 0 & \cdots & C & P_{2r} & H & 0 \\ 0 & 0 & 0 & 0 & 0 & \cdots & D & 0 & P_{2r+1} & 0 \end{pmatrix} & \end{matrix} ; \tag{5}$$

здесь используются следующие обозначения:

$$\begin{aligned}
 P_{od} &= A_0 \otimes B_0^{(1)} + A_1 \otimes B_1^{(1)}, \quad P_{ev} = A_0 \otimes B_0^{(2)}, \quad P_{em} = P_{ev} + A_0 \otimes B_1^{(2)} = A_0 \otimes I_{m_1}, \\
 P_{2r} &= P_{ev} + A_1 \otimes B_0^{(2)} = I_l \otimes B_0^{(2)}, \quad P_{2r+1} = P_{od} + A_1 \otimes B_0^{(1)} = A_0 \otimes B_0^{(1)} + A_1 \otimes I_{m_1}, \quad B_i = A_1 \otimes B_0^{(i)}, \\
 C &= A_0 \otimes \beta_2 f_1^T, \quad C_0 = A_0 \otimes \beta_1 f_2^T, \quad D = A_1 \otimes B_1^{(1)}, \quad H = A_1 \otimes \beta_2 f_1^T.
 \end{aligned}$$

Далее для нахождения вектора стационарных вероятностей \mathbf{p} применим подход, изложенный в [6, гл. 1, § 6] и использованный в [2]. Идея этого подхода заключается в последовательном “урезании” множества состояний \mathcal{X} ЦМ $\{X_s, s = 0, 1, 2, \dots\}$ с последующим “склеиванием” частей её траектории, оставшихся после очередного “урезания” и соответствующим пересчетом переходных вероятностей.

Этот метод особенно эффективен для СУР следующего вида:

$$\mathbf{q}^T = \mathbf{q}^T Q, \tag{6}$$

где вектор \mathbf{q} имеет вид:

$$\mathbf{q}^T = (\mathbf{q}_1^T, \mathbf{q}_2^T, \dots, \mathbf{q}_n^T), \tag{7}$$

где \mathbf{q}_i^T — вектор длины l_i , $i = \overline{0, n}$, а матрица Q имеет вид

$$Q = (Q_{ij})_{i,j=\overline{0,n}}, \tag{8}$$

где Q_{ij} — $(l_i \times l_j)$ -матрица, является стохастической матрицей, для которой выполнено условие $Q_{ij} = 0$ при $|i - j| > 1$.

Применение метода “урезания” к СУР такого вида даёт рекуррентный алгоритм вычисления стационарных вероятностей, приведённый в следующей лемме [2].

Лемма 1. Пусть вектор \mathbf{q} имеет вид (7) и матрица Q имеет вид (8). Тогда решение СУР (6) определяется следующим рекуррентным алгоритмом:

на первом шаге находится (с точностью до нормировочного множителя) вектор \mathbf{q}_1 из системы уравнений

$$\mathbf{q}_1^T = \mathbf{q}_1^T Q_{11}^{(n-1)},$$

где матрица $Q_{11}^{(n-1)}$ находится с помощью рекуррентного матричного алгоритма:

$$\begin{aligned} Q_{n-k+1, n-k+1}^{(k)} &= (I - Q_{n-k+1, n-k+1}^{(k-1)})^{-1}, & Q_{n-k+1, n-k}^{(k)} &= Q_{n-k+1, n-k+1}^{(k)} Q_{n-k+1, n-k}, \\ Q_{n-k, n-k}^{(k)} &= Q_{n-k, n-k} + Q_{n-k, n-k+1} Q_{n-k+1, n-k}^{(k)}, \\ & & k &= \overline{1, n-1}; \end{aligned}$$

на втором шаге последовательно (также с точностью до нормировочной константы) вычисляются \mathbf{q}_i , $i = \overline{2, n}$, по формуле

$$\mathbf{q}_i^T = \mathbf{q}_{i-1}^T Q_{i-1, i} Q_{ii}^{(n-i+1)};$$

на последнем шаге полученные ранее вероятности нормируются в соответствии с условием

$$\mathbf{q}^T \mathbf{1} = 1.$$

Для того, чтобы применить к СУР (2) результаты леммы 1, требуется ввести альтернативное разбиение множества состояний \mathcal{X} на макросостояния:

$$\mathcal{X} = \bigcup_{q=1}^R \mathcal{X}_q,$$

где

$$\mathcal{X}_q = \mathcal{X}_{q-1,2} \cup \mathcal{X}_{q,1}.$$

Заметим, что подмножество состояний \mathcal{X}_q содержит все состояния ЦМ $\{X_s, s = 0, 1, 2, \dots\}$ с q заявками в системе без учета их типа.

Далее введем векторы

$$\mathbf{p}_q^T = (\mathbf{p}_{q-1,2}^T, \mathbf{p}_{q,1}^T).$$

Тогда вектор \mathbf{p} можно представить в виде

$$\mathbf{p}^T = (\mathbf{p}_1^T, \mathbf{p}_2^T, \dots, \mathbf{p}_r^T, \mathbf{p}_R^T).$$

Следуя разбиению множества состояний \mathcal{X} , определяемому вектором \mathbf{p} , запишем матрицу P в следующем виде:

$$P = \begin{matrix} & \mathcal{X}_1 & \mathcal{X}_2 & \mathcal{X}_3 & \dots & \mathcal{X}_r & \mathcal{X}_R \\ \begin{matrix} \mathcal{X}_1 \\ \mathcal{X}_2 \\ \mathcal{X}_3 \\ \vdots \\ \mathcal{X}_r \\ \mathcal{X}_R \end{matrix} & \begin{pmatrix} F_1 & B & 0 & \dots & 0 & 0 \\ G & F & B & \dots & 0 & 0 \\ 0 & G & F & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & F & B \\ 0 & 0 & 0 & \dots & G & F_R \end{pmatrix} & \end{matrix}, \tag{9}$$

где

$$\begin{aligned} F_1 &= \begin{pmatrix} P_{em} & H \\ C_0 & P_{od} \end{pmatrix}, \quad F = \begin{pmatrix} P_{ev} & H \\ 0 & P_{od} \end{pmatrix}, \quad F_R = \begin{pmatrix} P_{2r} & H \\ 0 & P_{2r+1} \end{pmatrix}, \\ B &= \begin{pmatrix} B_2 & 0 \\ 0 & B_1 \end{pmatrix}, \quad G = \begin{pmatrix} 0 & C \\ 0 & D \end{pmatrix}. \end{aligned} \quad (10)$$

Сформулируем теперь основной результат, определяющий решение СУР (2).

Теорема 1. Пусть вектор \mathbf{p} имеет вид (1) и матрица P имеет вид (5). Тогда решение СУР (2) определяется следующим рекуррентным алгоритмом:

на первом шаге с точностью до нормировочного коэффициента находится вектор $\mathbf{p}_{0,2}$ из системы уравнений

$$\mathbf{p}_{0,2}^T P_0 = \mathbf{p}_{0,2}^T,$$

где матрица P_0 находится с помощью рекуррентного матричного алгоритма:

$$\begin{aligned} Q_{2r-1} &= B_2 (I - P_{2r})^{-1}, \\ Q &= B_2 (I - P_{ev})^{-1}, \\ \tilde{P}_{2n+1} &= (I - P_{2n+1})^{-1}, \quad n = \overline{r, 0}, \\ P_{2n-1} &= P_{od} + B_1 \tilde{P}_{2n+1} D, \quad n = \overline{r, 1}, \\ C_{2r-1} &= C + H \tilde{P}_{2r+1} D, \\ H_{2r-1} &= H + Q_{2r-1} C_{2r-1}, \\ C_{2n-1} &= C + H_{2n+1} \tilde{P}_{2n+1} D, \quad n = \overline{r-1, 1}, \\ H_{2n-1} &= H + Q C_{2n-1}, \quad n = \overline{r-1, 1}, \\ P_0 &= P_{em} + H_1 \tilde{P}_1 C_0. \end{aligned}$$

на втором шаге последовательно (также с точностью до нормировочной константы) вычисляются векторы \mathbf{p}_i , $i = \overline{2, n}$, по рекуррентным формулам

$$\begin{aligned} \mathbf{p}_{n,2}^T &= \mathbf{p}_{n-1,2}^T Q, \quad n = \overline{1, r-1}, \\ \mathbf{p}_{r,2}^T &= \mathbf{p}_{r-1,2}^T Q_{2r-1}, \\ \mathbf{p}_{1,1}^T &= \mathbf{p}_{1,2}^T H_1 \tilde{P}_1, \\ \mathbf{p}_{n,1}^T &= (\mathbf{p}_{n-1,1}^T B_1 + \mathbf{p}_{n,2}^T H_{2n-1}) \tilde{P}_{2n-1}, \quad n = \overline{2, r}, \\ \mathbf{p}_{r+1,1}^T &= \mathbf{p}_{r,1}^T B_1 \tilde{P}_{2r+1}, \end{aligned}$$

на последнем этапе производится нормировка полученных вероятностей в соответствии с условием (4).

Доказательство. Для доказательства теоремы достаточно подставить выражения (9) и (10) в результаты леммы 1 и применить несложные алгебраические преобразования, которые здесь не приводим.

Заметим, что алгоритм, приведенный в теореме 1, в применении к рассматриваемой СУР является более эффективным в смысле вычислительных затрат, чем изложенный в лемме 1, так как первый оперирует с матричными блоками меньшего размера $l \cdot t_i$, $i = 1, 2$, чем последний, который оперирует с блоками размера $l(t_1 + t_2)$.

5. СТАЦИОНАРНЫЕ ВЕРОЯТНОСТИ СОСТОЯНИЙ СИСТЕМЫ ПО МОМЕНТАМ ПОСТУПЛЕНИЯ ОСНОВНЫХ ЗАЯВОК

В этом разделе получим соотношения, связывающие стационарное распределение \mathbf{p} вероятностей состояний СМО, наблюдаемой в произвольные моменты времени, со стационарными вероятностями состояний системы, рассматриваемой в моменты непосредственно перед поступлениями 1-заявок $\{\pi_A^-(x), x \in \mathcal{X}\}$ и сразу после них $\{\pi_A^+(x), x \in \mathcal{X}\}$. При этом исключим из рассмотрения те 1-заявки, непосредственно перед поступлением которых (на одном такте) завершилось обслуживание заявки, находившейся на приборе. Таким образом, учитываться будут только такие события поступления 1-заявок, которые были единственными событиями за такт. Для записи выражений, описывающих эти вероятности, будем использовать векторы $\pi_A^\pm(k, i)$, аналогичные по структуре векторам $\mathbf{p}_{k,i}$.

Прежде, чем переходить к вычислению этих величин, определим вероятность $a_{(s)}$ поступления за такт 1-заявки при условии незавершения обслуживания заявки на приборе:

$$a_{(s)} = \sum_{i=1}^2 \sum_{k=u(2-i)}^{r+u(2-i)} \mathbf{p}_{k,i}^T (A_1 \otimes B_0^{(i)}) (\mathbf{1}_l \otimes \mathbf{1}_{m_i}).$$

Теорема 2. Пусть \mathbf{p} — вектор стационарных вероятностей состояний СМО $(DMAP, ST)/(DMSP, DMSP)/1/r$, удовлетворяющий СУР (2). Тогда стационарные вероятности $\{\pi_A^-(x), x \in \mathcal{X}\}$ состояний системы непосредственно перед поступлением 1-заявок при незавершенном обслуживании заявки на приборе определяются соотношениями

$$\begin{aligned} \pi_A^{T-}(k, 1) &= \frac{1}{a_{(s)}} \mathbf{p}_{k,1}^T (A_1 \otimes B_0^{(1)}), \quad k = \overline{1, R}, \\ \pi_A^{T-}(k, 2) &= \frac{1}{a_{(s)}} \mathbf{p}_{k,2}^T (A_1 \otimes B_0^{(2)}), \quad k = \overline{0, r}. \end{aligned}$$

Теорема 3. Пусть \mathbf{p} — вектор стационарных вероятностей состояний СМО $(DMAP, ST)/(DMSP, DMSP)/1/r$, удовлетворяющий СУР (2). Тогда стационарные вероятности $\{\pi_A^+(x), x \in \mathcal{X}\}$ состояний системы непосредственно после поступления 1-заявок при условии завершения обслуживания заявки на приборе определяются соотношениями

$$\begin{aligned} \pi_A^{T+}(1, 1) &= 0, \\ \pi_A^{T+}(k, 1) &= \pi_A^{T-}(k-1, 1) + u(k-r) \pi_A^{T-}(R, 1), \quad k = \overline{2, R}, \\ \pi_A^{T+}(0, 2) &= 0, \\ \pi_A^{T+}(k, 2) &= \pi_A^{T-}(k-1, 2) + u(k-r+1) \pi_A^{T-}(r, 2), \quad k = \overline{1, r}. \end{aligned}$$

Доказательства теорем 2 и 3 производятся с помощью вероятностных рассуждений, аналогичных используемым в секции 7 работы [1].

6. ЗАКЛЮЧЕНИЕ

Исследуется стохастическая модель однолинейной СМО с двумя типами заявок, имеющих относительный приоритет обслуживания. Поток основных заявок является марковским, а фоновых - насыщенным, т.е. потоком из источника, в котором запас заявок не ограничен. Процессы обслуживания обоих типов заявок также являются марковскими. Для основных заявок имеется накопитель конечной емкости. Для данной СМО получен эффективный матричный алгоритм для вычисления стационарных вероятностей её состояний в произвольные дискретные моменты времени. Также на их основе вычисляются стационарные вероятности состояний системы, наблюдаемой в моменты поступления заявок.

СПИСОК ЛИТЕРАТУРЫ

1. Бочаров П.П., Вискова Е.В. Однолинейная система массового обслуживания конечной емкости с марковскими потоком и обслуживанием в дискретном времени, *Автоматика и телемеханика*, 2005, № 2, стр. 73–91.
2. Бочаров П.П., Шлумпер Л.О. Однолинейная система массового обслуживания с фоновыми заявками, *Автоматика и телемеханика*, 2005, № 6, стр. 74–88.
3. Литвин В.Г., Лысяков П.К. О влиянии реорганизации базы данных на время ответа информационной системы. *Управляющие системы и машины*, 1981, № 2, стр. 105–109.
4. Литвин В.Г., Лысяков П.К. Дискретно-непрерывная модель информационной системы с совместным обслуживанием запросов и реорганизацией баз данных. *Автоматика и телемеханика*, 1983, № 5, стр. 149–158.
5. Бочаров П.П. Об обслуживании двух типов заявок с относительным приоритетом и распределениями фазового типа. В кн.: *Теория телеграфика в системах информатики*, М.: Наука, 1989, стр. 45–50.
6. Bocharov P.P., D'Apice C, Pechinkin A.V., Salerno S. *Queueing Theory*. Utrecht–Boston: VSP, 2004.