СТАЦИОНАРНОЕ РАСПРЕДЕЛЕНИЕ СИСТЕМЫ ОБСЛУЖИВАНИЯ С РЕКУРРЕНТНЫМ ВХОДНЫМ ПОТОКОМ И ПОВТОРНЫМИ ВЫЗОВАМИ

В. Клименок

Белорусский государственный университет
Минск, Беларусь
vklimenok@yandex.ru

В статье рассмотрена однолинейная система массового обслуживания с рекуррентным входным потоком и повторными вызовами. Получено необходимое и достаточное условие существования стационарного режима в системе и формулы для вычисления стационарного распределения в моменты поступления заявок и в произвольные моменты времени.

Ключевые слова: система массового обслуживания, рекуррентный поток, повторные вызовы, постоянная интенсивность повторов, стационарное распределение.

1. ВВЕДЕНИЕ

Актуальным направлением в теории массового обслуживания является исследование систем с повторными вызовами, которые хорошо моделируют процессы повторной передачи информации в реальных телекоммуникационных сетях и системах. В таких системах заявки, заставшие обслуживающее устройство занятым, не уходят из системы и не становятся в очередь, а идут на так называемую орбиту - виртуальный буфер для таких заявок, откуда делают повторные попытки попасть на обслуживание через случайные интервалы времени. В большинстве ранних работ, посвященных системам с повторными вызовами, рассматриваются модели с наиболее простым при аналитическом анализе стационарным пуассоновским потоком (см, например, [1]).

В настоящее время существует ряд работ по системам с повторными вызовами и марковскими потоками, которые являются хорошими математическими моделями трафика в современных телекоммуникационных сетях. Обзор соответствующих работ можно найти в [2]. Вместе с тем, достаточно общая модель марковского потока не может быть использована при описании реальных потоков с произвольным распределением интервалов между поступлениями запросов. Примером могут служить детерминированное, логнормальное, равномерное распределения. Системы с повторными вызовами и такими потоками практически не исследованы в литературе. Насколько нам известно, существует только две статьи ([3] и [4]) по указанной теме. В [3] исследуется однолинейная система с экспоненциально распределенным временем обслуживания и постоянной интенсивностью повторных попыток. Однако автор данной статьи не смогла преодолеть известные математические трудности, возникающие при исследовании таких систем, совершив ошибку при выводе преобразования Лапласа-Стилтьеса распределения времени пребывания в системе. Кроме того, использованный в статье математический аппарат не позволил получить аналитические результаты для стационарного распределения вероятностей состояний системы в компактном, пригодном для вычислений виде. Указанные ошибки и недостатки были устранены в недавно появившейся работе [4], где применен подход к исследованию системы, основанный на переопределении понятия время обслуживания и применении матрично-аналитических методов.

В данной статье обощается модель, исследованная в [4], на случай более адекватного практике распределения времени обслуживания заявок. Вместо экспоненциального распределения рассматривается гораздо более общее распределение фазового типа. Для рассматриваемой системы найдено условие существования стационарного режима и получены формулы для вычисления стационарного распределения в моменты поступления заявок и в произвольные моменты времени.

2. ОПИСАНИЕ СИСТЕМЫ

Рассматривается однолинейная система массового обслуживания с рекуррентным входным потоком и повторными вызовами. Интервалы между поступлениями в систему первичных заявок являются независимыми случайными величинами с произвольной функцией распределения $A(t)$, преобразованием Лапласа-Стилтьеса $A^*(s) = \int\limits_0^\infty e^{-st}dA(t), Re\, s \geq 0$, и конечным первым моментом $a_1 = \int\limits_0^\infty tdA(t) < \infty$.

Времена обслуживания как первичных, так и повторных заявок есть независимые случайные величины, имеющие распределение фазового типа с M фазами и неприводимым представлением $(\boldsymbol{\beta}, S)$. Интенсивность обслуживания μ вычисляется по формуле $\mu = [\boldsymbol{\beta}(-S)^{-1}\mathbf{e}]^{-\mathbf{1}}$. Описание распределения фазового типа можно найти, например, в [5].

Система не имеет буфера. Если в момент поступления заявки обслуживающий прибор свободен, то заявка немедленно начинает обслуживаться. Если прибор занят, то заявка идет на орбиту неограниченного объема. Старейшая по времени пребывания на орбите заявка делает повторные попытки занять прибор через экспоненциально распределеное время с параметром γ.

3. ПРОЦЕСС ИЗМЕНЕНИЯ СОСТОЯНИЙ СИСТЕМЫ. ВЛОЖЕННАЯ ЦЕПЬ МАРКОВА

Определим состояние системы в момент времени t как (0), если система свободна в данный момент. В противном случае состояние системы определим как вектор (i, m), где i - количество заявок в системе (на орбите и на обслуживании), m есть индикатор состояния прибора, $m \in \{0, \dots M\}$. Этот индикатор $m = 0$, если прибор свободен в момент времени t, и $m = m'$, $m' = \overline{1, M}$, если прибор в этот момент находится на m'-й фазе обслуживания заявки. Нетрудно видеть, что в данной системе прибор может быть свободен в случае, если система пуста или на орбите присутствуют заявки, но повторная попытка попасть на обслуживание еще не произошла.

Состояние системы в произвольный момент времени t описывается процессом ζ_t, $t \geq 0$, с пространством состояний $X = \{0;\ (i,m),\ i \geq 1,\ m = 0, 1, \ldots, M\}$.

Очевидно, что процесс ζ_t не является марковским, если входной поток немарковский. В данной работе стационарное распределение этого процесса найдем через стационарное распределение вложенной цепи Маркова.

Пусть t_n - момент поступления n-й заявки. Процесс $\xi_n^- = \zeta_{t_n - 0}$, $n \geq 1$, с пространством состояний X является неприводимой непериодической цепью Маркова, вложенной в процесс ζ_t по моментам времени $t_n - 0$, $n \geq 1$. В дальнейшем используются обозначения $t_n - 0$ и $t_n + 0$ для моментов времени непосредственно до и после поступления заявки. Для того, чтобы построить вложенную цепь Маркова ξ_n^-, $n \geq 1$, необходимо знать распределение числа заявок (первичных и повторных), которые заканчивают обслуживание в промежутке времени между соседними моментами поступления. Для расчета такого распределения введем понятие обобщенного времени обслуживания "успешной" заявки с орбиты. Обобщенное время обслуживания состоит из распределенного по экспоненциальному закону с параметром γ интервала времени, в течение которого старейшая заявка на орбите достигает свободного прибора, и собственно времени обслуживания этой заявки. Распределение обобщенного времени обслуживания может быть описано как распределение фазового типа с представлением $(\widetilde{\backslash\beta}, \tilde{S})$, где $\widetilde{\backslash\beta} = (1, \mathbf{0}_M)$ и $\tilde{S} = \begin{pmatrix} -\gamma & \gamma\backslash\beta \\ 0 & S \end{pmatrix}$, $\mathbf{0}_M$ – вектор-строка порядка M. Матрица $\tilde{S}$ описывает интенсивности переходов между фазами обобщенного времени обслуживания, которые не приводят к завершению обслуживания. Вектор $\widetilde{\backslash S}_0 = -\tilde{S}\mathrm{e} = \begin{pmatrix} 0 \\ \mathbf{S}_0 \end{pmatrix}$ описывает интенсивности переходов, приводящих к завершению обслуживания. Вектор $\widetilde{\backslash\beta}$ указывает, что обобщенное обслуживание заявки с орбиты всегда начинается с первой фазы. В дальнейшем будет использоваться следующее обозначение: $G(z) = \tilde{S} + \widetilde{\backslash S}_0 \widetilde{\backslash\beta} z$, $|z| \leq 1$. Матрица $G(1)$ является инфинитезимальным генераторм цепи Маркова, которая описывает эволюцию фаз обобщенного времени обслуживания в произвольном интервале $(t_n + 0, t_{n+1} - 0)$ между поступлениями заявок.

Обозначим через $P_{i,j}$ матрицу порядка 2×2, элементами которой являются вероятности переходов цепи Маркова ξ_n^-, $n \geq 1$, из состояний, соответствующих i заявкам в системе, в состояния, соответствующие j заявкам, $i, j \geq 0$.

Теорема 1. *Матрица вероятностей переходов цепи Маркова* ξ_n^-, $n \geq 1$, *имеет следующую блочную структуру*

$$P^- = (P_{i,j})_{i,j \geq 0} = \begin{pmatrix} B_0 & C_0 & 0 & 0 & \ldots \\ B_1 & A_1 & A_0 & 0 & \ldots \\ B_2 & A_2 & A_1 & A_0 & \ldots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix},$$

где матрицы $A_k,\ B_k,\ C_k\ k \geq 0,$ *определяются как*

$$C_0 = (0, \backslash\beta) \int\limits_0^\infty P(k,t)dA(t),\ B_0 = 1 - \backslash\beta A^*(-S)\mathbf{e},$$

$$A_k = \mathcal{B} \int\limits_0^\infty P(k,t)dA(t),\ B_k = \mathcal{B} \sum_{l=k+1}^\infty \int\limits_0^\infty P(l,t)dA(t)\mathbf{e},\ k > 0.$$

Здесь $\mathcal{B} = \begin{pmatrix} 0 & \backslash\beta \\ \mathbf{0}_M^T & I_M \end{pmatrix}$, *а матрицы* $P(k,t),\ k \geq 0,$ *определяются как коэффициенты матричного ряда* $\sum\limits_{k=0}^\infty P(k,t)z^k = e^{G(z)t},\ |z| \leq 1.$

Анализируя структуру матрицы P^- вероятностей переходов, легко видеть, что рассматриваемая цепь $\xi_n^-,\ n \geq 1$ принадлежит классу цепей Маркова типа $GI/M/1$ (см. [5]). Из общего условия эргодичности для такого рода цепей мы выводим в следующем разделе условие существования стационарного режима в рассматриваемой системе в терминах функций и параметров, определяющих ее функционирование.

4. УСЛОВИЕ СУЩЕСТВОВАНИЯ СТАЦИОНАРНОГО РЕЖИМА. ВЫЧИСЛЕНИЕ СТАЦИОНАРНОГО РАСПРЕДЕЛЕНИЯ

Условие существования стационарного режима в системе совпадает с условием эргодичности вложенной цепи Маркова $\xi_n^-,\ n \geq 1,$ и формулируется следующим образом.

Теорема 2. *Стационарный режим существует тогда и только тогда, когда выполняется неравенство*

$$[\backslash\beta + \hat{\mathbf{x}}(I - B)]\{S(\gamma B - S)^{-2}[A^*(\gamma B - S) - I]\}\mathbf{S}_0 + \frac{a_1\mu\gamma}{\mu + \gamma} > 1, \qquad (1)$$

где

$$\hat{\mathbf{x}} = \backslash\beta(I - \Phi)[I - (I - B)(I - \Phi)]^{-1},$$

$$\Phi \stackrel{def}{=} S(\gamma B - S)^{-1}[A^*(\gamma B - S) - I],\ B \stackrel{def}{=} \mathbf{e}\backslash\beta.$$

Далее будем предполагать, что условие (1) выполняется. Обозначим стационарные вероятности цепи Маркова ξ_n^- как $\pi_0^-,\ \pi^-(i,m),\ i > 0,\ m = 0, 1, \ldots, M.$ Пусть также $\backslash\pi_i^- = (\pi^-(i,0),\ \pi^-(i,1), \ldots, \pi^-(i,M)),\ i > 0,$ - вектор-строки этих вероятностей.

Теорема 3. *Стационарное распределение цепи Маркова* $\xi_n^-,\ n \geq 1,$ *вычисляется следующим образом:*

$$\boldsymbol{\pi}_i^- = \boldsymbol{\pi}_1^- \mathcal{R}^{i-1},\ i \geq 1,$$

где матрица $\mathcal{R}$ является минимальным неотрицательным решением матричного уравнения $\quad \mathcal{R} = \sum\limits_{j=0}^{\infty} \mathcal{R}^j A_j, \quad$ *а вектор* $(\pi_0^-, \boldsymbol{\pi}_1^-)$ *определяется как единственное решение следующей системы линейных алгебраических уравнений*

$$(\pi_0^-, \boldsymbol{\pi}_1^-) = (\pi_0^-, \boldsymbol{\pi}_1^-) \begin{pmatrix} B_0 & C_0 \\ \sum\limits_{j=1}^{\infty} \mathcal{R}^{j-1} B_j & \sum\limits_{j=1}^{\infty} \mathcal{R}^{j-1} A_j \end{pmatrix}, \quad \pi_0^- + \boldsymbol{\pi}_1^- (I - \mathcal{R})^{-1} \mathbf{e} = 1.$$

Рассмотрим теперь цепь Маркова $\xi_n^+,\ n \geq 1$, вложенную в процесс ζ_t по моментам времени $t_n + 0,\ n \geq 1$. Эта цепь имеет пространство состояний $\{(i, m),\ i \geq 1,\ m = 1, \ldots, M\}$. Очевидно, что условие эргодичности цепи совпадает с соответствующим условием для цепи $\xi_n^-,\ n \geq 1$, и определяется в теореме 2. Обозначим через $\backslash\pi_i^+ = (\pi^+(i, 1), \ldots, \pi^+(i, M)),\ i \geq 1$, вектор-строки стационарных вероятностей процесса $\xi_n^+,\ n \geq 1,$. Эти векторы можно вычислить через найденное выше стационарное распределение цепи $\xi_n^-,\ n \geq 1$.

Теорема 4. *Стационарные вероятности цепи Маркова* $\xi_n^+,\ n \geq 1$, *вычисляются через стационарные вероятности цепи Маркова* $\xi_n^-,\ n \geq 1$, *следующим образом:*

$$\backslash\pi_1^+ = \pi_0^- \backslash\beta, \quad \backslash\pi_i^+ = \backslash\pi_{i-1}^- \begin{pmatrix} \backslash\beta \\ I_M \end{pmatrix} = \backslash\pi_1^- \mathcal{R}^{i-2} \begin{pmatrix} \backslash\beta \\ I_M \end{pmatrix},\ i > 1.$$

Используя полученные результаты для вложенных цепей Маркова, найдем стационарное распределение немарковского процесса $\zeta_t,\ t \geq 0$. Искомое распределение может быть выражено через стационарное распределение цепи ξ_n^+ либо цепи ξ_n^-. Оба случая эквивалентны с точки зрения математической сложности. Мы будем рассматривать цепь $\xi_n^+,\ n \geq 1$.

Пусть

$$p_0 = \lim_{t \to \infty} P\{\zeta_t = 0\}, \quad p(i, m) = \lim_{t \to \infty} P\{\zeta_t = (i, m)\},\ i \geq 1, m = 0, 1, \ldots, M, \qquad (2)$$

стационарные вероятности процесса ζ_t. Пусть также $\backslash p_i = (p(i, 0), p(i, 1), \ldots, p(i, M))$, $i \geq 1$.

Пределы (2) существуют, если неравенство (1) выполняется и $a_1 < \infty$.

Теорема 5. *Стационарное распределение процесса* $\zeta_t,\ t \geq 0$, *вычисляется через стационарное распределение вложенной цепи Маркова* $\xi_n^+,\ n \geq 1$, *следующим образом:*

$$\boldsymbol{p}_i = \sum_{j=i}^{\infty} \boldsymbol{\pi}_j^+ \Omega_{j-i},\ i \geq 1,\ \ p_0 = 1 - \sum_{i=0}^{\infty} \boldsymbol{p}_i \mathbf{e}.$$

где

$$\Omega_k = \lambda \hat{I} \int\limits_0^{\infty} P(k, t)(1 - A(t))dt,\ k \geq 0,\ \ \hat{I} = (\mathbf{0}_M^T \,|\, I_M).$$

5. ЗАКЛЮЧЕНИЕ

В статье рассмотрена однолинейная система массового обслуживания с рекуррентным входным потоком и постоянной интенсивностью повторов. Функционирование системы описано немарковским двумерным процессом и двумя вложенными по моментам поступления цепями Маркова, найдено стационарное распределение состояний системы в моменты поступления заявок и в произвольные моменты времени. Используя эти распределения, нетрудно найти ряд важных характеристик производительности системы. В процессе будущих исследований планируется также рассмотреть задачу нахождения такой важной характеристики как распределение времени пребывания в системе. Заметим, что все представленные в статье результаты согласуются с результатами для классической системы $GI/PH/1$. Это можно проверить, перейдя к пределу $\gamma \to \infty$ в соответствующих формулах.

ЛИТЕРАТУРА

1. *Falin G. I., Templeton J. G. C.* Retrial Queues.London: Chapman & Hall, 1997.
2. *Artalejo, J. R., Comez-Corral A.* Retrial queueing systems: a computational approach. Berlin-Heidelberg: Springer, 2008.
3. *Lillo R. E.* A $G/M/1$ - Queue with Exponential Retrial // Top. 1966. V. 4. P. 99–120.
4. *Kim C. S., Klimenok V., Dudin A.* A G/M/1 retrial queue with constant retrial rate // Top. DOI: 10.1007/s11750-012-0267-3.
5. *Neuts M.* Matrix-geometric solutions in stochastic models. Baltimore: Johns Hopkins University Press, 1981.