Математика

ДЛИНА ОЧЕРЕДИ В ПРОСТЕЙШЕЙ МНОГОКАНАЛЬНОЙ СИСТЕМЕ МАССОВОГО ОБСЛУЖИВАНИЯ С МОНОТОННЫМИ ИНТЕНСИВНОСТЯМИ ОБСЛУЖИВАНИЯ И С ОЖИДАНИЕМ

Симонян А. Р., Симонян Р. А., Улитина Е. И.

QUEUE LENGTH IN ELEMENTARY MULTICHANNEL SYSTEM OF MASS SERVICE WITH MONOTONOUS INTENSITY OF SERVICE AND WAITING

Simonyan A. R., Simonyan R. A., Ulitina E. I.

The article studies multichannel system of mass service with several serving devices (channels) with Poisson calls input flow and with exponential function of calls service duration distribution.

В статье рассматривается многоканальная система массового обслуживания с несколькими обслуживающими приборами (каналами), с пуассоновским входящим потоком вызовов и с экспоненциальной функцией распределения длительностей обслуживания вызовов.

Keywords: Poisson flow, mass service theory, queue length.

Ключевые слова: пуассоновский поток, теория массового обслуживания, длина очереди.

УДК 519.21

Одной из первых систем массового обслуживания, позволивших свое изучение методами процессов размножения и гибели [1], является система $\boldsymbol{M|M|m|\infty}$, что по терминологии Кендалла-Башарина [1] означает: в m – канальную систему массового обслуживания с ожиданием – поступает пуассоновский поток вызовов с параметром $\lambda > 0$, а длительности обслуживания вызовов независимы в совокупности, не зависят от процесса поступления и имеют экспоненциальную функцию распределения:

$$1 - exp\{-\mu t\}, \mu > 0, t > 0.$$

В момент $t = 0$ система свободна от вызовов.

Поступивший первый вызов немедленно начинает свое обслуживание. Следующий поступивший вызов начинает обслуживание следующим прибором и так далее. Когда в момент поступления вызовов все приборы заняты, то поступившие вызовы становятся в очередь и ждут своего обслуживания в порядке поступления (FIFO).

Анализ такой системы – ее основных характеристик (длина очереди, время ожидания, период занятости) – содержится во многих работах по теории массового обслуживания. Отметим монографии Б.В. Гнеденко и И.Н. Коваленко [1], Г.П. Климова [2], Т.Л. Саати [3] и др.

На практике предположение о пуассоновости входящего потока не является сильным ограничением, поскольку, как показали интересные исследования С. Пальма [4], А. Рени [5], Г.А. Ососкова [6] и Б.И. Григелиониса [7], наложение многих «малых» потоков при достаточно широких условиях приводит к суммарному пуассоновскому.

В то же время использование предположения об экспоненциальности распределения длительностей обслуживания вызовов менее оправдано. Случаи экспоненциального распределения обслуживания редки на практике.

Но при разработке практических систем разработчики обычно интересуются не распределением длины очереди, времени ожидания, периода занятости, а их средними и иногда дисперсиями.

Оказывается, заменяя произвольное обслуживание на экспоненциальное с тем же средним и рассчитывая средние характеристики последней системы, обычно получают верхние оценки для тех же средних в исходной системе.

Таким образом, упрощающее предположение об экспоненциальности длительностей обслуживания все же дает разработчику верхние оценки характеристик исходной системы. Принимая также во внимание, что упрощающие предположения позволяют применять хорошо разработанный аппарат процессов размножения и гибели и приводят к простым конечным результатам, можно в первом приближении ограничиться изучением системы $\boldsymbol{M|M|m|\infty}$.

В конце 1980-х годов прошлого века внимание исследователей привлекали системы массового обслуживания, в которых длительности обслуживания вызовов зависят от длины очереди. И это естественно. Например, в магазине при увеличении очереди продавец начинает работать быстрее (хотя иногда менее качественно). Таких примеров можно привести много.

Обобщением системы $\boldsymbol{M|M|m|\infty}$ на случай зависимости длительностей обслуживания вызовов от длины очереди занимались T. Suzuki [8] и L. Takacs [9].

Мы же собираемся для более общего случая в стационарном режиме работы системы получить распределение длины очереди.

Рассмотрим систему $\boldsymbol{M|M|\downarrow m|\infty}$, которая от исходной системы отличается тем, что каждый прибор (канал) нумеруется и прибор с номером k $(k = 1,2,\dots,m)$ имеет свою интенсивность $\mu_k > 0$ обслуживания вызовов. Предполагается, без ограничения общности, что интенсивность обслуживания монотонно неубывающая: $\mu_1 \le \mu_2 \le \dots \le \mu_m$. Данная система называется простейшей многоканальной моделью с монотонными интенсивностями и с ожиданием.

Обозначим через $P_k(t)$ вероятность того, что в момент t в системе находятся k вызовов $k = 1,2,\dots$.

Доказано (см. [1]), что система с ожиданием в случае пуассоновского входящего потока и экспоненциального времени обслуживания представляет собой случайный процесс Маркова. Это облегчает дальнейшие рассуждения.

Нас интересуют стационарные вероятности

$$p_k = \lim_{t\to\infty} P_k(t), k = 1,2,\dots . \quad (1)$$

Существование этих пределов следует из теоремы Феллера [10].

Основной результат работы формулируется следующим образом:

<u>Теорема 1.</u> В модели $\boldsymbol{M|M|\downarrow m|\infty}$ при $\rho < 1$, для вероятностей p_k, $k = 0,1,2,\cdots$ имеют место формулы

$$\begin{cases} p_k = \prod_{i=0}^{k} \rho_i \; p_0 & \text{при } 1 \le k < m, \\ p_k = \rho^{k-m} \cdot \prod_{i=0}^{m} \rho_i \, p_0 & \text{при } k \ge m, \end{cases}$$

где $p_0 = \left(\sum_{k=0}^{m} \prod_{i=0}^{k} \rho_i + \frac{\rho}{1-\rho} \cdot \prod_{i=0}^{m} \rho_i \right)^{-1}$, $\rho_0 = 1$, $\rho_k = \frac{\lambda}{\overline{\mu_k}}$,

$$\overline{\mu_k} = \sum_{i=1}^{k} \mu_i, k = 1,2,\dots m, \rho = \rho_m.$$

<u>Доказательство.</u> Задача в первую очередь состоит в том, чтобы найти те уравнения, которым удовлетворяют вероятности $P_k(t)$, после чего, применяя (1), получим p_k. Одно из уравнений очевидно, а именно для любого t

$$\sum_{k=0}^{\infty} P_k(t) = 1. \quad (2)$$

В первую очередь найдем вероятность того, что в момент $t + h$ все приборы свободны от вызовов. Это может произойти следующими способами:

- в момент t все приборы были свободны, и за время h новых вызовов не поступило;

- в момент t первый прибор был занят обслуживанием вызова, все остальные приборы свободны, и за время h обслуживание вызова было завершено, и новых вызовов не поступило.

Остальные возможности из-за свойства ординарности (см. [1]) входящего потока имеют вероятность $o(h)$.

Вероятность первого из указанных событий равна

$$P_0(t)e^{-\lambda t} = P_0(t)\big(1-\lambda h + o(h)\big),$$

Вероятность второго события

$$P_1(t)e^{-\lambda t}(1-e^{-\mu_1 h}) = P_1(t)\mu_1 h + o(h).$$

Таким образом,

$$P_0(t+h) = P_0(t)(1-\lambda h) + P_1(t)\mu_1 h + o(h).$$

Отсюда нетрудно прийти к уравнению (см. [1] и [3])

$$P_0'(t) = -\lambda P_0(t) + \mu_1 P_1(t). \quad (3)$$

Перейдем теперь к составлению уравнений для $P_k(t)\sum_{i=1}^{k+1}\mu_i h$ при $k \geq 1$. Рассмотрим два случая: $1 \leq k < m$ и $k \geq m$.

Пусть $1 \leq k < m$. Перечислим все состояния, из которых можно попасть в состояние E_k (система находится в состоянии E_k, если в системе находятся k штук вызовов) в момент $t+h$. Эти состояния следующие:

- в момент t система находилась в состоянии E_k, а за время h новых вызовов не было, и ни один прибор не завершил обслуживание. Эта вероятность вычисляется следующим образом

$$P_k(t)e^{-\lambda h}\prod_{i=1}^{k} e^{-\mu_i h} = P_k(t)\big(1-\lambda h - \textstyle\sum_{i=1}^{k}\mu_i h\big) + o(h);$$

- в момент t система находилась в состоянии E_{k-1}, за время h поступил один вызов, но ни один прибор не завершил обслуживание. Эта вероятность вычисляется следующим образом

$$P_{k-1}(t)(1-e^{-\lambda h})\prod_{i=1}^{k-1} e^{-\mu_i h} = \lambda h P_{k-1}(t) + o(h);$$

- в момент t система находилась в состоянии E_{k+1}, за время h новых вызовов не было, но один вызов был обслужен. Эта вероятность вычисляется следующим образом

$$P_{k+1}(t)e^{-\lambda h}\sum_{i=1}^{k+1}\big((1-e^{-\mu_i h})\prod_{j\epsilon\theta_i} e^{-\mu_i h}\big) = P_{k+1}(t)\sum_{i=1}^{k+1}\mu_i h + o(h),$$

где $\theta_i = \{1,2,\cdots,k\}\backslash\{i\}$.

Все остальные возможные переходы в состояние E_k за промежуток времени h имеют вероятность $o(h)$. Таким образом, мы получили следующую формулу для $P_k(t+h)$ при $1 \leq k < m$

$$P_k(t+h) =$$

$$= P_k(t)\big(1-\lambda h - \textstyle\sum_{i=1}^{k}\mu_i h\big) + \lambda h P_{k-1}(t) + P_{k+1}(t)\sum_{i=1}^{k+1}\mu_i h + o(h).$$

Несложные преобразования (см. [1] и [3]) приводят к следующему уравнению при $1 \leq k < m$

$$P_k'(t) = -\big(\lambda h + \textstyle\sum_{i=1}^{k}\mu_i h\big)P_k(t) + \lambda h P_{k-1}(t) + \big(\textstyle\sum_{i=1}^{k+1}\mu_i h\big)P_{k+1}(t). \quad (4)$$

Аналогичными рассуждениями для $k \geq m$ получаем

$$P'_k(t) = -(\lambda h + \sum_{i=1}^{m} \mu_i h) P_k(t) + \lambda h P_{k-1}(t) + (\sum_{i=1}^{m} \mu_i h) P_{k+1}(t). \quad (5)$$

Таким образом, вероятности $P_k(t)$ являются решением бесконечной системы дифференциальных уравнений (2) – (5). Решение этой системы представляет технические трудности.

В теории массового обслуживания обычно изучают только решение системы (2)-(5) при $t \to +\infty$. Как следует из [8]

$$\lim_{t \to +\infty} P'_k(t) = 0. \quad (6)$$

Из (1) и (6) приходим к следующим уравнениям

$$\begin{cases} -\lambda p_0 + \mu_1 p_1 = 0, \\ \lambda p_{k-1} - (\lambda + \sum_{i=1}^{k} \mu_i) p_k + p_{k+1} \sum_{i=1}^{k+1} \mu_i = 0 \quad \text{при } 1 \leq k < m, \\ \lambda p_{k-1} - (\lambda + \sum_{i=1}^{m} \mu_i) p_k + p_{k+1} \sum_{i=1}^{m} \mu_i = 0 \quad \text{при } k \geq m. \end{cases} \quad (7)$$

Пользуясь обозначением $\overline{\mu_k} = \sum_{i=1}^{k} \mu_i$, из (7) получаем

$$\begin{cases} -\lambda p_0 + \overline{\mu_1} p_1 = 0, \\ \lambda p_{k-1} - (\lambda + \overline{\mu_k}) p_k + p_{k+1} \overline{\mu_{k+1}} = 0 \quad \text{при } 1 \leq k < m, \\ \lambda p_{k-1} - (\lambda + \overline{\mu_m}) p_k + p_{k+1} \overline{\mu_m} = 0 \quad \text{при } k \geq m. \end{cases} \quad (8)$$

Добавляя к (8) условие

$$\sum_{k=0}^{\infty} p_k = 1 \quad (9)$$

и применяя обозначения

$$\begin{cases} z_k = \lambda p_{k-1} - \overline{\mu_k} p_k \quad \text{при } 1 \leq k < m, \\ z_k = \lambda p_{k-1} - \overline{\mu_m} p_k \quad \text{при } \quad k \geq m, \end{cases} \quad (10)$$

Из (8) и (10) получаем

$$z_1 = z_2 = \cdots = z_k = 0, \quad (11)$$

для любого $k = 1, 2, \ldots$.

Пусть $\rho = \rho_m$ загрузка системы, тогда из (11), пользуясь обозначениями $\rho_k = \frac{\lambda}{\overline{\mu_k}}, k = 1, 2, \ldots m$, получаем решение системы (8):

$$\begin{cases} p_k = \prod_{i=0}^{k} \rho_i \ p_0 \quad \text{при } 1 \leq k < m, \\ p_k = \rho^{k-m} \cdot \prod_{i=0}^{m} \rho_i \, p_0 \quad \text{при } k \geq m, \end{cases} \quad (12)$$

где $p_0 = \left(\sum_{k=0}^{m} \prod_{i=0}^{k} \rho_i + \frac{\rho}{1-\rho} \cdot \prod_{i=0}^{m} \rho_i \right)^{-1}$, $\rho_0 = 1$ получается из (9).

Пределы (1) и решения (12) существуют в стационарном режиме работы системы, т.е. при $\rho < 1$ (см. [1]). Теорема 1 доказана.

Таким образом, получено распределение числа вызовов в системе.

Обозначим через ξ – число вызовов в очереди, тогда ξ является случайной величиной и принимает значения:

$$\xi = \begin{cases} 0 \quad \text{при } \zeta \leq m; \\ \zeta - m \quad \text{при } \zeta > m, \end{cases}$$

где ζ – число вызовов в системе.

Пусть $L_{ср}$ – средняя длина очереди в системе, т.е. среднее число вызовов. Тогда имеет место следующая теорема:

Теорема 2. В модели $\boldsymbol{M|M|\downarrow m|\infty}$ при $\rho < 1$, для средней длины очереди имеет место формула

$$L_{ср} = \frac{\rho}{(1-\rho)^2} \prod_{i=0}^{m} \rho_i \, p_0, \quad (13)$$

где ρ, ρ_i и p_0 определены в теореме 1.

Доказательство. Обозначим через $q_k, k = 1,2,\cdots$ стационарную вероятность того, то в очереди k вызовов. Нетрудно убедится, что

$$q_k = \begin{cases} \sum_{i=0}^{m} p_i, & \text{при } k = 0 \\ p_{m+k}, & \text{при } k \geq 1. \end{cases} \quad (14)$$

Так как $L_{ср}$ математическое ожидание дискретной случайной величины ξ, с вероятностями q_k, то

$$L_{ср} = M\xi = \sum_{k=0}^{\infty} k q_k. \quad (15)$$

Подставляя значения q_k из (14) в (15) и применяя формулы (12), мы получаем (13). Сходимость ряда (15) обеспечивает условие стационарности $\rho < 1$. Теорема 2 доказана.

Таким образом, для модели $\boldsymbol{M|M|\downarrow m|\infty}$ получены условие стационарности, распределение длины очереди и выведена формула вычисления средней длины очереди. Все полученные результаты могут быть использованы в решениях таких задач, как построение имитационных моделей в экономике, в вычислительных сетях и системах, в туриндустрии, а также в таком актуальном на сегодня направлении, как моделирование транспортных потоков с целью избавления от транспортных заторов. Ведь основными характеристиками транспортной инфраструктуры, которые определяют возникновение транспортных заторов, являются загрузка и длина очереди.

Литература

1. Гнеденко Б.В., Коваленко И.Н. Введение в теорию массового обслуживания. М.: Наука, 1987. С. 336.
2. Климов Г.П. Стохастические системы обслуживания. М.: Наука, 1966. С. 243.
3. Саати Т.Л. Элементы теории массового обслуживания и ее применения. М.: Сов. радио, 1965. С. 520.
4. Palm. C. Intensitatsschwankungen in fernsprechverkehr // Ericson Technics. 1943. V.44. No. 1. P. 1–189.
5. Renyi A. Poisson-folyamat egy jemllemzese // Тр. Мат. ин-та АН Венгрии. 1956. Т. 1. № 4. С. 519–527.
6. Ососков Г.А. Одна предельная теорема для потоков однородных событий // Теория вероятностей и ее применение. 1956. Т. 1. №2. С. 274–282.
7. Гигелионис Б.И. О точности приближения композиции процессов восстановления пуассоновским процессом // Литовский математический сборник. 1962, Т. 2. №2. С. 135-143.
8. Suzuki T. A queueing system with service depending on queue length // Journal of Operations Research Society of Japan. 1962, №4. P. 147-169.
9. Tacacs L. Investigation of waiting time problems by reduction to Markov processes. 1955, 6, №1–2. P. 101–129.
10. Феллер В. Введение в теорию вероятностей и ее приложения. М., 1984. Т. 2. С. 751.

Сведения об авторах:

1. Симонян Арсен Рафикович, канд. физ.-мат. наук, доцент, зав. каф. общей математики СГУТиКД (Сочи).
E-mail: oppm@mail.ru

2. Симонян Рафик Арсенович, студент 5-го курса факультета математики и компьютерных наук КубГУ (Краснодар).
E-mail: raf55@list.ru

3. Улитина Елена Ивановна, канд. физ.-мат. наук, доцент СГУТиКД (Сочи).
E-mail: ulitinaelena@mail.ru