УДК 519.872

О моделировании систем массового обслуживания с множественными ресурсами

В. А. Наумов*, **К. Е. Самуйлов**†

* *Исследовательский центр процессов обслуживания*
ул. Лённротинкату, д. 30 д, Хельсинки, Финляндия, 00180.
† *Кафедра прикладной информатики и теории вероятностей*
Российский университет дружбы народов
ул. Миклухо-Маклая, д. 6, Москва, Россия, 117198

Рассматриваются системы массового обслуживания, в которых для обслуживания заявок требуются некоторые ресурсы, освобождаемые после их ухода. Поступившие заявки теряются, если в системе недостаточно свободных ресурсов, необходимых для их обслуживания. Поскольку по завершении обслуживания занимаемые ресурсы должны быть освобождены, необходимо для каждой обслуживаемой заявки помнить вектор занимаемых ею ресурсов. Это существенно усложняет случайные процессы, описывающие поведение систем во времени.

Мы предлагаем вместо таких систем исследовать их упрощённый вариант. Упрощённая система функционирует аналогично исходной, за исключением того, что объёмы ресурсов, освобождаемых по завершении обслуживания, являются случайными и могут отличаются от тех, которые были выделены заявке в начале её обслуживания. При заданных суммарных объёмах занятых ресурсов и числе заявок в системе объёмы ресурсов, освобождаемых в момент завершения обслуживания, не зависят от поведения системы до этого момента и имеет функцию распределения, которую легко вычислить, используя формулу Байеса. Случайные процессы, описывающие поведение упрощённых систем, легче поддаются анализу, поскольку отпадает необходимость запоминания объёмов ресурсов, занимаемых каждой заявкой. Достаточно помнить суммарные объёмы занятых ресурсов. Результаты моделирования говорят, что характеристики исходной и упрощённой систем очень близки.

Ключевые слова: система массового обслуживания, ограниченные ресурсы, вероятность потери вызова, кусочно-линейчатый марковский процесс.

1. Введение

В настоящей работе рассматриваются системы, в которых на всё время обслуживания заявки занимают некоторые виды ресурсов. По завершении обслуживания должны быть освобождены в точности те объёмы ресурсов, которые были выделены обслуженной заявке. Поступившие заявки теряются, если в системе недостаточно свободных ресурсов, необходимых для их обслуживания. Предполагается, что случайные векторы, описывающие требования заявок к ресурсам, не зависят от процессов поступления и обслуживания заявок, независимы в совокупности и одинаково распределены. Многочисленные примеры таких, различных по сложности, систем можно найти в [1]. Поскольку по завершении обслуживания занимаемые ресурсы должны быть освобождены, необходимо для каждой обслуживаемой заявки помнить вектор занимаемых ею ресурсов. Это существенно усложняет случайные процессы, описывающие поведение систем во времени.

Для таких сложных систем мы предлагаем исследовать их упрощённый вариант. Упрощённая система функционирует аналогично исходной, за исключением того, что объёмы ресурсов, освобождаемых по завершении обслуживания, являются случайными и могут отличаются от тех, которые были выделены заявке в начале её обслуживания. При заданных суммарных объёмах занятых ресурсов и числе заявок в системе объёмы ресурсов, освобождаемых в момент завершения обслуживания, не зависят от поведения системы до этого момента и имеют функцию распределения, которую легко вычислить, используя формулу Байеса.

Статья поступила в редакцию 16 июня 2014 г.

Случайные процессы, описывающие поведение упрощённых систем, легче поддаются анализу, поскольку отпадает необходимость запоминания объёмов ресурсов, занимаемых каждой заявкой. Достаточно помнить суммарные объёмы занятых ресурсов. Мы иллюстрируем предложенный подход на примере одной системы массового обслуживания (СМО), описываемой кусочно-линейчатым марковским процессом [2].

2. Исходная СМО с ограниченными ресурсами

Рассмотрим некоторую многолинейную систему массового обслуживания. Предположим, что поступающий поток заявок является рекуррентным. Длительности обслуживания заявок независимы между собой, не зависят от поступающего потока и одинаково распределены. Система располагает ограниченным объёмом ресурсов нескольких типов и функционирует следующим образом:

1. Для обслуживания каждой заявки требуется некоторое количество ресурса каждого типа.
2. Поступившая заявка теряется, если в момент поступления количество требуемого ей ресурса превышает количество свободного ресурса этого типа.
3. В момент начала обслуживания заявки суммарный объём занятого ресурса каждого типа увеличивается на величину ресурса, выделенного этой заявке.
4. В момент окончания обслуживания заявки суммарный объём занятого ресурса каждого типа уменьшается на величину ресурса, выделенного этой заявке.

Обозначим через M число типов ресурсов, R_m — общий объём ресурса типа m, и $\mathbf{r}_j = (r_{j1}, r_{j2}, \ldots, r_{jM})$ — вектор объёмов ресурсов, необходимых для обслуживания j-й заявки, $j = 1, 2, \ldots$. Будем считать, что случайные векторы $\mathbf{r}_j$ не зависят от процессов поступления и обслуживания заявок, независимы в совокупности и одинаково распределены с функцией распределения $F(\mathbf{x})$.

В общем случае состояние системы в момент t можно описать непрерывным слева кусочно-линейчатым марковским процессом $X(t) = (\xi(t), \alpha(t), \beta(t), \gamma(t))$. Здесь $\xi(t)$ — число заявок в системе; $\alpha(t)$ — время, оставшееся до ближайшего момента поступления заявки; $\beta(t) = (\beta_1(t), \beta_2(t), \ldots, \beta_{\xi(t)}(t))$, где $\beta_i(t)$ — оставшееся время обслуживания i-й обслуживаемой заявки; $\gamma(t) = (\boldsymbol{\gamma}_1(t), \boldsymbol{\gamma}_2(t), \ldots, \boldsymbol{\gamma}_{\xi(t)}(t))$, где $\boldsymbol{\gamma}_i(t)$ — вектор объёмов ресурсов, занимаемых i-й обслуживаемой заявкой.

Находящиеся на обслуживании заявки имеют свою внутреннюю нумерацию, и при поступлении на обслуживание очередной заявки они перенумеровываются так, чтобы элементы последовательности $\beta(t)$ оставались всё время упорядоченными по убыванию. На интервалах времени, в течение которых нет ни поступлений, ни уходов заявок из системы, последовательность $\gamma(t)$ остаётся неизменной, а процесс $\alpha(t)$ и каждый элемент последовательности $\beta(t)$ убывает с единичной скоростью. Если система свободна, то обе последовательности $\beta(t)$ и $\gamma(t)$ являются пустыми.

Процесс $X(t)$ имеет довольно сложное пространство состояний X. Так, его подмножество X_k, включающее состояния системы с k заявками, состоит из всех векторов вида $(k, a, b_1, \ldots, b_k, c_{11}, \ldots c_{1M}, \ldots, c_{k1}, \ldots c_{kM})$ длины $k(M+1)+2$, удовлетворяющих условиям

$$a \geqslant 0, \quad b_1 \geqslant b_2 \geqslant \ldots \geqslant b_k \geqslant 0, \quad c_{11}, \ldots, c_{1M}, \ldots, c_{k1}, \ldots, c_{kM} \geqslant 0,$$

$$\sum_{i=1}^{k} c_{im} \leqslant R_m, \quad m = 1, 2, \ldots, M.$$

3. Упрощённая СМО с ограниченными ресурсами

Состояние упрощённой системы в момент t можно описать непрерывным слева кусочно-линейчатым марковским процессом $X^*(t) = (\xi^*(t), \alpha^*(t), \beta^*(t), \boldsymbol{\delta}^*(t))$, в котором $\xi^*(t)$ — число заявок в системе; $\alpha^*(t)$ — время, оставшееся до ближайшего момента поступления заявки; $\beta^*(t) = (\beta_1^*(t), \beta_2^*(t), \ldots, \beta_{\xi^*(t)}^*(t))$, где $\beta_i^*(t)$ — оставшееся время обслуживания i-й обслуживаемой заявки; $\boldsymbol{\delta}^*(t) = (\delta_1^*(t), \ldots, \delta_M^*(t))$ — вектор суммарных объёмов занятых ресурсов.

Эта система массового обслуживания аналогична предыдущей системе во всём, за исключением того, что вместо сформулированного выше правила 4) уменьшение суммарных объёмов занятых ресурсов в момент ухода заявки подчиняется другому правилу:

4^*) В момент окончания обслуживания заявки t_i вектор суммарных объёмов занятых ресурсов уменьшается на случайную величину $\boldsymbol{\nu}_i = (\nu_{i1}, \nu_{i2} \ldots, \nu_{iM})$, которая, при заданном числе k заявок в системе и векторе $\mathbf{y}$ суммарных объёмов занятых ресурсов, не зависит от поведения системы в прошлом и имеет следующую функцию распределения:

$$P(\boldsymbol{\nu}_i \leqslant \mathbf{x} | \xi^*(t_i) = k \ ; \ \boldsymbol{\delta}^*(t_i) = \mathbf{y}); \ X^*(t), \ 0 \leqslant t \leqslant t_i) = F_k(\mathbf{x}|\mathbf{y}),$$

где

$$F_k(\mathbf{x}|\mathbf{y}) = P(\mathbf{r}_k \leqslant \mathbf{x} | \mathbf{r}_1 + \mathbf{r}_2 + \ldots + \mathbf{r}_k = \mathbf{y}), \quad 0 \leqslant \mathbf{x} \leqslant \mathbf{y}.$$

Здесь $\mathbf{r}_i$, $i = 1, 2, \ldots$ это те же, что и для первой системы, векторы объёмов ресурсов, необходимых для обслуживания заявок.

Пусть, например, функция распределения $F(\mathbf{x})$ объёмов ресурсов, необходимых для обслуживания заявок, имеет плотность $f(\mathbf{x})$. В этом случае функция распределения $F_k(\mathbf{x}|\mathbf{y})$ также имеет плотность $f_k(\mathbf{x}|\mathbf{y})$, которую можно вычислить используя формулу, Байеса. В результате получим:

$$f_k(\mathbf{x}|\mathbf{y}) = f(\mathbf{x}) \frac{f^{(k-1)}(\mathbf{y} - \mathbf{x})}{f^{(k)}(\mathbf{y})}, \quad 0 \leqslant \mathbf{x} \leqslant \mathbf{y},$$

где $f^{(n)}(\mathbf{x})$ есть n-кратная свёртка плотности $f(\mathbf{x})$.

Пространство состояний X^*процесса $X^*(t)$ значительно проще, чем пространство состояний исходного процесса $X(t)$. Подмножество $X_k^* \subset X^*$ возможных состояний упрощённой системы обслуживающей k заявок состоит из векторов $(k, a, b_1, \ldots, b_k, c_1, \ldots, c_M)$ длины $k + M + 2$, удовлетворяющих условиям $a \geqslant 0$, $b_1 \geqslant b_2 \geqslant \ldots \geqslant b_k \geqslant 0$, $0 \leqslant c_m \leqslant R_m$, $m = 1, 2, \ldots, M$.

Заметим, что при неизменных объёмах необходимых заявкам ресурсов, когда $\mathbf{r}_i = \mathbf{d}$ для всех $i = 1, 2, \ldots$, упрощённая система функционирует совершенно так же, как и исходная система.

4. Пример

Для примера рассмотрим систему с одним типом ресурса и предположим, что функция распределения необходимого заявкам ресурса имеет гамма-распределение с плотностью

$$f(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}, \quad x \geqslant 0.$$

Тогда функция $f_k(x|y)$ совпадает с плотностью бета-распределения

$$f_k(x|y) = \frac{\Gamma(k\alpha)}{\Gamma(\alpha)\,\Gamma(k\alpha - \alpha)} \left(\frac{x}{y}\right)^{\alpha-1} \left(1 - \frac{x}{y}\right)^{k\alpha-\alpha-1}, \quad 0 \leqslant x \leqslant y.$$

Варьируя параметры гамма-распределения, можно промоделировать поведение исходной и упрощённой систем при различных средних значениях $m = \alpha/\beta$ и коэффициенте вариации $c = 1/\sqrt{\alpha}$ объёма необходимого заявкам ресурса.

В табл. 1 приведены вероятности потерь для системы с пуассоновским входящим потоком интенсивности λ и экспоненциальным обслуживанием с параметром μ, полученные с помощью имитационного моделирования. Были использованы датчики случайных чисел гамма и бета распределений из популярной библиотеки random.f90, основанной на алгоритмах, изложенных в [3]. Относительная ошибка результатов моделирования с вероятностью 0,95 не превосходит 1%.

Таблица 1

Вероятность потери вызова при $m = 10$

c^2	λ/μ	Исходная СМО	Упрощённая СМО
0.5	2	0,180	0,180
	4	0,433	0,433
	6	0,570	0,570
	8	0,654	0,654
	10	0,711	0,711
1,0	2	0,162	0,162
	4	0,404	0,403
	6	0,540	0,541
	8	0,625	0,625
	10	0,682	0,683
2.0	2	0,111	0,111
	4	0,306	0,305
	6	0,441	0,441
	8	0,530	0,530
	10	0,593	0,592
4.0	2	0,052	0,052
	4	0,146	0,145
	6	0,260	0,259
	8	0,354	0,352
	10	0,429	0,427

Из таблицы видно, что при указанных входных параметрах обе системы имеют если и не совпадающие, то очень близкие вероятности потерь.

5. Выводы

В статье предложен новый подход к анализу СМО с множественными ресурсами. Результаты моделирования обнадёживают, однако требуются дополнительные исследования области его применимости.

Литература

1. *Наумов В. А., Самуйлов К. Е., Яркина Н. В.* Теория телетрафика мультисервисных сетей. — Москва: РУДН, 2007. — 192 с.
2. *Гнеденко Б. В., Коваленко И. Н.* Введение в теорию массового обслуживания. — Москва: Наука, 1966. — 432 с.
3. *Dagpunar J.* Principles of Random Variate Generation. — Oxford University Press, 1988. — 248 p.

UDC 519.872

On the Modeling of Queueing Systems with Multiple Resources

V. A. Naumov[*], K. E. Samuylov[†]

[*] *Palveluinnovaatioiden kehityskeskus (Service Innovation Research Institute)*
30, Lönnrotinkatu, Helsinki, Finland, 00180

[†] *Department of Applied Informatics and Probability Theory*
Peoples' Friendship University of Russia
6, Miklukho-Maklaya str., Moscow, Russian Federation, 117198

We consider queueing systems, in which customers occupy some resources that are released after customer departure. Arriving customers are lost if there is not enough free resources required for their servicing. In such systems for each customer it is necessary to record vector of occupied resources until its departure. This greatly complicates the stochastic processes describing the behavior of systems in time.

Instead of systems of this type we propose to investigate their simplified analogy. Simplified system operates similarly to the original, except that the amount of resources released upon completion of service, is random and may differ from those that have been allocated to the customer. For given total amount of resources employed and the number of applications in the system, the amount of resources released at the completion of service, does not depend on the behavior of the system up to this point and has a distribution function, which can be easily computed using Bayes' formula. Random processes describing the behavior of simplified systems are easier to analyze, because there is no need to memorize the volume of resources held by each customer. It is enough to record the total amount of occupied resources. The simulation results say that the characteristics of the original and simplified systems are very close.

Key words and phrases: queuing systems, limited resources, loss probability, piecewise linear Markov process.

References

1. V. F. Naumov, K. E. Samuylov, N. V. Yarkina, Teletraffic Theory for Multiservice Networks, RUDN, Moscow, 2007, in Russian.
2. B. V. Gnedenko, I. N. Kovalenko, Introduction to Queueing Theory, 2d ed., Birkhäuser, Boston, 1998.
3. J. Dagpunar, Principles of Random Variate Generation, Oxford University Press, 1988.