Расчет систем обслуживания с большим числом каналов[1]

Ю.И. Рыжиков

Институт информатики и автоматизации РАН,
Военно-космическая академия им. А.Ф. Можайского,
г. Санкт-Петербург, Россия

Постановка задачи

Современные методы расчета многоканальных немарковских систем обслуживания [1,2], особенно актуальные для вычислительных систем, медленно внедряются в учебный процесс вузов и практически неизвестны заказчикам и проектировщикам таких систем. Оценки средней длины очереди и среднего ожидания аппроксимации ([3–5]) весьма приблизительны и работают лишь для небольшого числа каналов n.

Известно, что при $n \to \infty$ стационарное распределение числа заявок в системе независимо от вида распределения обслуживания описывается распределением Пуассона

$$P_k = {(\lambda b)^k e^{-\lambda b}} \big/ {k!}, \quad k = 0,1,\ldots , \tag{1}$$

где λ – интенсивность входящего пуассонова потока и b – среднее время обслуживания. В связи с этим многие исследователи обращаются к идее применения для обсчета систем с большим числом каналов формулы Эрланга и соответствующих оценок для средней длины очереди.

В данной статье эта идея анализируется на количественной основе. В ней демонстрируется зависимость средней длины очереди от вариации распределения обслуживания, коэффициента загрузки и числа каналов.

Фазовые аппроксимации

В значительной доле реальных ситуаций входящий поток близок к простейшему, что определяет повышенную актуальность соответствующих моделей. Отметим, что происходящие при работе сетей обслуживания многократное случайное просеивание и суммирование потоков приближают результирующие потоки на входе узлов к простейшим.

---

[1] Работа поддержана государственным заданием 0073-2018-0003.

Расчет многоканальных систем практически возможен лишь при использовании фазовых аппроксимаций распределений длительностей обслуживания. В качестве таковых обычно используются эрлангова либо гиперэкспоненциальная второго порядка H_2. Диаграммы состояний и переходов между ними приводятся для модели $M/E_3/2$ в [6], а для $M/H_2/3$ – в [7]. Эрлангово распределение E_k применимо только для аппроксимации распределений с коэффициентом вариации, меньшим единицы. Критически влияющая на трудоемкость расчетов ширина диаграммы (максимальное количество микросостояний на ярусе) очень быстро растет по числу каналов n и порядку $k = 1/v^2$ (см. табл. 1).

Таблица 1

Количество микросостояний на ярусах модели $M/E_k/n$

Число каналов	Число фаз k				
	2	3	4	5	6
2	3	6	10	15	21
3	4	10	20	35	56
5	6	21	56	126	252
10	11	66	286	1001	3003

При обсчете моделей с меньшими v могут потребоваться значительно большие значения k с катастрофическим ростом размерности задачи при больших n.

Гиперэкспоненциальная аппроксимация H_2 позволяет выравнивать три момента произвольного распределения, что представляется необходимым и достаточным. Диаграмма переходов для $M/H_2/n$ имеет ширину $n+1$, что позволяет обсчитывать системы с очень большим числом каналов.

По указанным причинам наиболее актуальной представляется модель $M/H_2/n$, для которой и проводятся все последующие рассуждения.

Численный метод

Обсчет СМО методом фиктивных фаз основан на решении уравнений баланса переходов между микросостояниями системы, задаваемыми общим числом заявок в системе и распределением проходящих обслуживание заявок по его фазам.

При работе с большим числом каналов предпочтительно применение итерационного метода, восходящего к Такахаси и Таками [1, 8]. В нем последовательно пересчитываются векторы условных (нормированных в пределах яруса к единице) вероятностей микросостояний и отношения

$$x_j = p_{j+1}/p_j, \qquad j = 0,1,\ldots,m, \tag{2}$$

суммарных вероятностей ярусов. При стабилизации последних с разумной погрешностью (мы принимали допуск 10^{-4}) из условия баланса

$$\sum_{j=0}^{n-1}(n-j)p_j = n - \lambda b \tag{3}$$

определяется вероятность свободного состояния p_0 и далее через формулу (2) – вероятности $p_j, \quad j = 1,2,\ldots,m$. Конечным результатом каждого расчета считалась ожидаемая длина очереди

$$\bar{q} = \sum_{j=n+1}^{m}(j-n)p_j + \frac{p_m x}{1-x}\left[\frac{1}{1-x} + m - n\right]. \tag{4}$$

Здесь $x = p_m/p_{m-1}$. Последнее слагаемое отражает добавку от неучтенных членов бесконечной геометрической прогрессии со знаменателем x.

Машинный эксперимент

Целью машинного эксперимента являлось оценивание влияния загрузки системы и коэффициента вариации времени обслуживания на среднюю длину очереди q. Прокомментируем диапазон его условий.

1. Коэффициенты загрузки выбирались 0.7 и 0.9 (левая граница – из соображений экономичности системы, вторая – как предельно допустимая для сохранения удовлетворенности пользователей).

2. Коэффициент v_b вариации распределения длительности обслуживания рассматривался в разумном диапазоне $v \in [0,2]$. Частное значение $v_b = 0.577$ соответствует распределению Эрланга 3-го порядка и аппроксимируется с комплексными параметрами. Значение $v_b = 0.800$ выбрано из предположительно «опасного» интервала $(1/\sqrt{2},1)$, в котором H_2-аппроксимация имеет парадоксальные значения вероятностей выбора фазы – отрицательную и превышающую единицу.

3. Число каналов СМО выбиралось в пределах возможностей вычислительного метода и реализующих его Фортран-программ. Расчет заканчивался при получении очень малых длин очередей.

В табл. 2 и 3 представлены вычисленные средние длины очередей, в табл. 4 и 5 – их относительные значения по отношению к марковской модели ($v = 1$).

Таблица 2

Средняя длина очереди, ρ = 0.7

Число каналов	Коэффициенты вариации обслуживания						
	0.000	0.300	0.577	0.800	1.000	1.500	2.000
1	8.167e-1	8.902e-1	1.089	1.339	1.633	2.654	4.083
2	6.920e-1	7.514e-1	9.111e-1	1.111	1.345	2.151	3.270
3	6.027e-1	6.527e-1	7.867e-1	9.541e-1	1.149	1.816	2.740
5	4.760e-1	5.134e-1	6.136e-1	7.379e-1	8.816e-1	1.368	2.041
10	2.932e-1	3.143e-1	3.702e-1	4.389e-1	5.174e-1	7.799e-1	1.129
15	1.933e-1	2.063e-1	2.405e-1	2.822e-1	3.294e-1	4.839e-1	6.937e-1
20	1.318e-1	1.402e-1	1.621e-1	1.886e-1	2.183e-1	3.144e-1	4.434e-1
25	9.176e-2	9.729e-2	1.117e-1	1.290e-1	1.483e-1	2.098e-1	2.889e-1
30	6.476e-2	6.849e-2	7.816e-2	8.971e-2	1.025e-1	1.427e-1	1.936e-1
50	1.736e-2	1.822e-2	2.040e-2	2.296e-2	2.573e-2	3.406e-2	4.402e-2
70	4.956e-3	5.172e-3	–	–	7.009e-3	8.944e-3	1.112e-2

Таблица 3

Средняя длина очереди, ρ = 0.9

Число каналов	Коэффициенты вариации обслуживания						
	0.000	0.300	0.577	0.800	1.000	1.500	2.000
1	4.050	4.414	5.400	6.642	8.100	13.162	20.250
2	3.868	4.211	5.139	6.306	7.674	12.394	19.013
3	3.727	4.055	4.940	6.053	7.354	11.844	18.086
5	3.506	3.811	4.632	5.661	6.862	11.011	16.780
10	3.120	3.385	4.099	4.985	6.019	9.585	14.529
15	2.843	3.080	3.719	4.505	5.424	8.587	12.938
20	2.621	2.838	3.418	4.127	4.957	7.783	11.680
25	2.437	2.636	3.165	3.813	4.571	7.122	10.620
30	2.278	2.461	2.948	3.567	4.243	6.555	9.732
50	1.800	1.938	–	2.765	3.275	4.954	7.300
70	–	–	–	–	2.623	3.904	5.699
100	–	–	–	–	1.952	2.841	4.092
150	–	–	–	–	1.263	1.780	2.505
200	–	–	–	–	0.850	1.165	1.608
250	–	–	–	–	0.587	0.785	1.063

Таблица 4

Относительная длина очереди, ρ = 0.7

Число каналов	Коэффициенты вариации обслуживания						
	0.000	0.300	0.577	0.800	1.000	1.500	2.000
1	0.5000	0.5451	0.6669	0.8200	1.0000	1.6252	2.5003
2	0.5145	0.5587	0.6774	0.8261	1.0000	1.5992	2.4312
3	0.5245	0.5680	0.6847	0.8304	1.0000	1.5805	2.3847
5	0.5399	0.5824	0.6960	0.8370	1.0000	1.5517	2.3151
10	0.5667	0.6075	0.7154	0.8483	1.0000	1.5073	2.1821
15	0.5868	0.6263	0.7298	0.8561	1.0000	1.4702	2.1060
20	0.6038	0.6422	0.7426	0.8639	1.0000	1.4402	2.0311
25	0.6187	0.6560	0.7532	0.8699	1.0000	1.4147	1.9480
30	0.6318	0.6682	0.7625	0.8752	1.0000	1.3922	1.8888
50	0.6747	0.7081	0.7928	0.8923	1.0000	1.3237	1.7108
70	0.7071	0.7379	–	–	1.0000	1.2761	1.5865

Таблица 5

Относительная длина очереди, ρ = 0.9

Число каналов	Коэффициенты вариации обслуживания						
	0.000	0.300	0.577	0.800	1.000	1.500	2.000
1	0.5000	0.5449	0.6667	0.8200	1.0000	1.6249	2.5000
2	0.5040	0.5487	0.6697	0.8217	1.0000	1.6151	2.4776
3	0.5068	0.5514	0.6717	0.8231	1.0000	1.6106	2.4593
5	0.5109	0.5554	0.6750	0.8250	1.0000	1.6046	2.4453
10	0.5184	0.5624	0.6810	0.8282	1.0000	1.5924	2.4139
15	0.5241	0.5678	0.6857	0.8306	1.0000	1.5831	2.3850
20	0.5287	0.5725	0.6895	0.8326	1.0000	1.5701	2.3526
25	0.5331	0.5767	0.6924	0.8342	1.0000	1.5581	2.3233
30	0.5369	0.5800	0.6948	0.8407	1.0000	1.5449	2.2937
50	0.5496	0.5918	–	0.8443	1.0000	1.5127	2.2290
70	–	–	–	–	1.0000	1.4884	2.1727
100	–	–	–	–	1.0000	1.4570	2.0963
150	–	–	–	–	1.0000	1.4093	1.9834
200	–	–	–	–	1.0000	1.3703	1.8913
250	–	–	–	–	1.0000	1.3385	1.8118

По поводу этих результатов и способов их получения отметим следующее:

1. Расчет системы $M/D/n$ предполагалось выполнить методом Кроммелина [9]. Однако при $n \geq 25$ матрица коэффициентов системы линейных уравнений для вероятностей $\{p_j\}$, $j = \overline{0, n-1}$, становится

вырожденной, что блокирует продолжение вычислений. Поэтому все системы независимо от коэффициента вариации обслуживания обсчитывались общей программой $M/H_2/n$. Прочерками отмечены случаи расходимости итераций.

2. Наблюдалось очень быстрое возрастание средней длины очереди по коэффициенту вариации v (для $n=1$ в соответствии с формулой Полячека – Хинчина – в 5 раз) и по коэффициенту загрузки пропорционально отношению $\rho^2/(1-\rho)$ – в 4.96 раза (близость этих результатов – случайное следствие подбора исходных данных).

3. При $\rho = 0.9$ и коэффициенте вариации $v > 1$ средняя длина очереди $\overline{q}$ значима по крайней мере до $n = 30$.

4. При прочих равных условиях увеличение числа каналов приближает результаты к получаемым согласно марковской модели: для $v_b < 1$ растет по n, при $v_b > 1$ – убывает. Это приближение происходит «медленнее, чем хотелось бы». При граничных исходных данных и $\rho = 0.9$ получаем соответственно 0.55 и 1.80 – отклонения более чем значимые.

5. Ожидаемое число свободных каналов, значимое для оценки экономичности системы и возможности параллельного решения вспомогательных задач, на основе условия (3) баланса заявок не зависит от v. То же можно сказать о среднем числе занятых каналов.

Заключение

1. В связи с расширением применения многомашинных и многопроцессорных вычислительных систем возрастает актуальность разработки методов расчета показателей их работы и прежде всего – средних длин очередей.

2. При увеличении числа каналов средняя длина очереди приближается (в зависимости от значения v_b – снизу или сверху) к полученным на основе марковской модели, однако разница в исследованном диапазоне условий остается весьма значительной и указанная аппроксимация – недопустимой.

3. В связи с выявленным ограничением применимости метода Такахаси – Таками по числу каналов необходимо углубленное исследование сходимости итерационных методов решения систем линейных алгебраических уравнений. Это касается матриц с комплексными коэффици-

ентами и общей проблемы появления при больших n «квазисобственных» значений [10], замедляющих или исключающих сходимость итераций.

ЛИТЕРАТУРА

1. *Рыжиков Ю.И.* Развитие и сопоставление методов расчета многоканальных систем обслуживания // Труды XII Всероссийского совещания по проблемам управления. М.: ИПУ РАН, 2014. С. 5208–5219.
2. *Яшков С.Ф.* Столетие теории очередей: предисловие к тематическому выпуску // Автоматика и телемеханика. 2009. № 12. С. 3–8.
3. *Bhat U.N.* An Introduction to Queueing Theory. Springer, 2015. 268 p.
4. *Bolch G., Greiner S., de Meer H., Trivedi K.S.* Queueing Networks and Markov Chains. Modelling and Preformance Evaluation. Wiley and Sons, 2006. 878 p.
5. *Medhi J.* Stochastic Models in Queueing Theory. Elsevier, 2003. 450 p.
6. *Рыжиков Ю.И.* Алгоритм расчета многоканальной системы с эрланговским обслуживанием // Автоматика и телемеханика. 1980. № 5. С. 30–37.
7. *Рыжиков Ю.И., Хомоненко А.Д.* Итеративный метод расчета многоканальных систем с произвольным распределением времени обслуживания // Проблемы управления и теории информации. 1980. № 3. С. 32–38.
8. *Takahashi Y., Takami Y.* A numerical method for the steady-tate probabilities of a GI/G/c queueing system in a general class // J. Operat. Res. Soc. of Japan. 1976. V. 19. No. 2. P. 147–157.
9. *Рыжиков Ю.И.* Алгоритмический подход к задачам массового обслуживания: Монография. СПб.: ВКА им. А.Ф. Можайского, 2013. 496 с.
10. *Бахвалов Н.С., Жидков Н.П., Кобельков Г.М.* Численные методы. М.: Лаборатория Базовых Знаний, 2000. 624 с.
11. *Neuts M.F.* Matrix-Analytic Methods in the Queueing Theory // Eur. J. Operational Research. 1984. V. 5. P. 2–12.

Рыжиков Юрий Иванович, д.т.н., профессор; ryzhbox@yandex.ru