УДК 519.872

РАЗВИТИЕ И СОПОСТАВЛЕНИЕ МЕТОДОВ РАСЧЕТА МНОГОКАНАЛЬНЫХ СИСТЕМ ОБСЛУЖИВАНИЯ

Ю.И. Рыжиков

Санкт-Петербургский Институт информатики и автоматизации РАН

Россия, 199178, Санкт-Петербург, 14-я линия В.О., 39

E-mail: ryzhbox@yandex.ru

Ключевые слова: массовое обслуживание, теория очередей, итерационный метод, матрично-геометрическая прогрессия

Аннотация: Показана необходимость расчета систем обслуживания с немарковским обслуживанием и большим числом каналов. Дан обзор методов, основанных на фазовых аппроксимациях, и предложены их доработки. Приведены результаты численных расчетов, оценки их трудоемкости и рекомендации по применению обсуждаемых методов.

1. Введение

Вступление человечества в эру информационных технологий определяет растущий интерес к методам проектирования и оценки эффективности систем обработки и передачи данных. Основу таких методов составляет теория массового обслуживания (ТМО), она же — теория очередей. Особенностью текущего момента является приближение технологии производства СБИС, составляющих основу современных ЭВМ, к фундаментальным физическим ограничениям. По этой причине требуемые показатели производительности достигаются созданием многопроцессорных и многомашинных систем. Отметим, что многоканальное обслуживание имеет и другие применения (транспорт, медицина, промышленное производство, службы быстрого реагирования и т. п.). Однако методы расчета таких систем развиты недостаточно.

В данной работе предлагаются модификации схем расчета многоканальных систем обслуживания и дается анализ их вычислительной эффективности.

2. Фазовая аппроксимация многоканальных СМО

Эффективным методом марковизации сложных систем обслуживания является многофазное представление составляющих распределений с сохранением некоторого числа их начальных моментов, причем количество выравниваемых моментов может рассматриваться как порядок аппроксимации. Общий вариант (распределение Ньютса) сводится к представлению процесса прибытия и (или) обслуживания заявки цепью Маркова, определение параметров которой в литературе даже не обсуждается, а собственно расчет СМО проводится в терминах кронекеровых матричных операций и как правило — для одноканальных систем. Машинная реализация упомянутых операций крайне неэффективна из-за необходимости выполнения большого числа операций с заведомо нулевыми результатами. По этой причине практически реализуется «фольклорная» (по М. Ньютсу) концепция: для распределений с коэффициентом вариации, большим единицы, используют гиперэкспоненциальную (H_2-) аппроксимацию, а в остальных случаях — эрлангову.

Исследование апппроксимации Г-распределения с коэффициентом вариации v выявило, что

- случай $v > 1$ дает вещественные параметры аппроксимации;
- при $1 > v > 1/\sqrt{2}$ параметры вещественны, но парадоксальны: одна из вероятностей выбора экспоненты больше единицы, а вторая отрицательна;
- строгое равенство $v = 1/\sqrt{2}$ недопустимо (случай E_2, последовательная система двух фаз не может быть заменена параллельной);
- при $v < 1/\sqrt{2}$ имеем комплексные параметры H_2-аппроксимации.

Обширная серия численных экспериментов показала, что эта потенциальная патология проявляется только в промежуточных результатах — вероятностях «фиктивных» микросостояний, на которые расщепляются «физические». В «общий строй» можно поставить и распределение E_2 — если, учитывая неизбежную погрешность реальных данных, согласиться на небольшое отклонение дисперсии. Заметим к тому же, что ширина диаграмм (максимальное количество микросостояний на ярусе) для моделей с эрланговым обслуживанием быстро растет по числу каналов n и порядку k распределения обслуживания согласно формуле $\binom{n+k-1}{n}$ [13] — см. таблицу 1.

Таблица 1. Количество микросостояний на ярусах системы $M/E_k/n$

Число каналов n	Число фаз обслуживания 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
20	21	231	1771	10626	53130
30	31	496	5456	46376	324632

Эрланговы распределения позволяют строго выравнивать первый и лишь приближенно — второй моменты распределения обслуживания. Коэффициент вариации эрлангова распределения порядка k составляет $v = 1/\sqrt{k} < 1$. Для наибольшего из включенных в таблицу 1 значения $k = 6$ он равен 0.408. При обсчете моделей с ме́ньшими коэффициентами вариации могут потребоваться значительно бо́льшие значения k. С другой стороны, аппроксимация H_2 позволяет выравнивать три момента произвольного (исключая E_2) распределения, что представляется необходимым и достаточным. Поскольку диаграмма переходов для $M/H_2/n$ имеет ширину всего $n+1$, здесь можно рассчитывать системы с достаточно большим числом каналов. По указанным причинам распределение времени обслуживания целесообразно представлять гиперэкспонентой H_2.

Далее все рассуждения проводятся применительно к модели $M/H_2/n$, поскольку

- в значительной доле реальных ситуаций (в частности, при расчете *сетей обслуживания*, где потоки подвергаются многократным операциям случайного просеивания и суммирования) входящий поток близок к простейшему;
- гиперэкспоненциальная аппроксимация H_2 распределения длительности обслуживания позволяет выравнивать три момента произвольного распределения, что представляется необходимым и достаточным.

Работа системы $M/H_2/n$ может быть интерпретирована как обслуживание потока заявок двух типов, причем выбор типа определяет параметр показательно распределенного обслуживания — см. диаграммы переходов в [5, 6, 11]. На этих диаграммах «ключ» микросостояния указывает количество находящихся в каналах обслуживания заявок каждого типа. Прибывающая с интенсивностью λ (или выбираемая из очереди) заявка с вероятностью y_i относится к i-му типу. Параметр потока обслуживаний заявок i-го типа равен $m_i\mu_i$, где m_i — содержимое i-й позиции ключа. Нетривиальна часть диаграммы переходов по обслуживанию *при наличии очереди*: его завершение в зависимости от типа выбранной из очереди заявки с вероятностями $\{y_i\}$ приводит в различные микросостояния вышележащего яруса.

Обозначим через S_j множество всех возможных микросостояний системы, современных методов расчета многоканальных СМО, при которых на обслуживании находится ровно j заявок, а через σ_j — количество элементов в S_j. Далее в соответствии с диаграммой переходов построим матрицы интенсивностей инфинитезимальных переходов:

$A_j[\sigma_j \times \sigma_{j+1}]$ — в S_{j+1} (прибытие заявки),
$B_j[\sigma_j \times \sigma_{j-1}]$ — в S_{j-1} (полное завершение обслуживания заявки),
$D_j[\sigma_j \times \sigma_j]$ — ухода из состояний яруса j
(в квадратных скобках здесь и далее указывается размер матриц).

Введем векторы-строки $\gamma_j = \{\gamma_{j,1}, \gamma_{j,2}, \dots, \gamma_{j,\sigma_j}\}$ нахождения СМО в состоянии (j,i), $j = 0, 1, \dots$ Теперь можно записать векторно-матричные уравнения баланса переходов между состояниями

$$\begin{aligned} \gamma_0 D_0 &= \gamma_1 B_1, \\ \gamma_j D_j &= \gamma_{j-1} A_{j-1} + \gamma_{j+1} B_{j+1}, \qquad j = 1, 2, \dots \end{aligned} \tag{1}$$

Систему (1), дополненную условием нормировки, практически приходится расписывать покомпонентно. Даже для моделей с ограниченной очередью она характеризуется чрезвычайно высокой размерностью, и стандартные методы решения систем линейных алгебраических уравнений применительно к ней оказываются малоэффективными.

3. Итерационный метод

Основная расчетная схема. Такахаси и Таками [13] предложили алгоритм итерационного расчета подобных систем, центральной идеей которого является переход к расчету условных (нормированных к единице) вероятностей микросостояний системы $\{\gamma_{j,i}^{(m)}\}$ для фиксированного числа заявок в системе (ярус диаграммы) и параллельно — вычисления отношений $\{x_j\}$ суммарных безусловных вероятностей на смежных ярусах, $x_j = p_{j+1}/p_j$, $j = 0, 1, \ldots$. Расчет выполняется для ограниченного числа ярусов $j = \overline{0, N}$. В итерации номер m вектор условных вероятностей $\gamma_j^{(m)}$ для каждого яруса при прогонке сверху вниз выражается через $\gamma_{j-1}^{(m)}$ и $\gamma_{j+1}^{(m-1)}$. При обсчете последнего яруса используется замыкающее систему уравнений *приближенное* равенство

$$\gamma_{j+1}^{(m-1)} \approx \gamma_{j-1}^{(m)}. \tag{2}$$

В [13], однако, отсутствовал ряд ключевых деталей алгоритма (расчет $\{x_j\}$) и не обсуждался выбор начальных приближений к векторам $\{\gamma_j\}$. Дадим ответы на эти вопросы, развивая изложенные в [4, 5] подходы.

Положим $t_j = \gamma_j/p_j$, где p_j — суммарная вероятность наличия в системе ровно j заявок, и обозначим

$$x_j = p_{j+1}/p_j, \qquad z_j = p_{j-1}/p_j. \tag{3}$$

Тогда систему (1) можно переписать относительно векторов условных вероятностей $\{t_j\}$, нормированных к единице в пределах яруса:

$$\begin{array}{rcl} t_0 D_0 & = & x_0 t_1 B_1, \\ t_j D_j & = & z_j t_{j-1} A_{j-1} + x_j t_{j+1} B_{j+1}, \qquad j = 1, 2, \ldots \end{array} \tag{4}$$

С помощью векторов-столбцов $j = \{1, 1, \ldots, 1\}^T$ размера σ_j для всех j могут быть записаны дополняющие систему (4) условия нормировки

$$t_j j = 1 \tag{5}$$

и баланса суммарных интенсивностей переходов между смежными ярусами

$$t_j A_j j + 1 = x_j t_{j+1} B_{j+1} j. \tag{6}$$

Алгоритм расчета набора векторов $\{t_j\}$ и чисел $\{x_j\}$ и $\{z_j\}$, удовлетворяющих соотношениям (4)–(6), в случае разомкнутой системы с неограниченной очередью опирается на существование предельного вектора условных вероятностей $t = \lim_{j\to\infty} t_j$, которое является следствием стабилизации переходных матриц при $j > n$.

Перепишем уравнения системы (4) для $j \geqslant 1$ в виде

$$t_j^{(m)} D_j = z_j^{(m)} t_{j-1}^{(m)} A_{j-1} + x_j^{(m)} t_{j+1}^{(m-1)} B_{j+1}, \qquad j = 1, 2, \ldots,$$

где верхний индекс указывает номер итерации. Теперь ясно, что

$$t_j^{(m)} = z_j^{(m)} \beta_j' + x_j^{(m)} \beta_j'', \tag{7}$$

где

$$\begin{aligned} \beta_j' &= t_{j-1}^{(m)} \quad A_{j-1} D_j^{-1}, \\ \beta_j'' &= t_{j+1}^{(m-1)} B_{j+1} D_j^{-1}. \end{aligned} \tag{8}$$

При $j = N$ (предельный индекс) считается, что

(9) $$\beta''_N = t^{(m)}_{N-1} B_N D_N^{-1}.$$

Осталось указать способ расчета $\{z^{(m)}_j\}$ и $\{x^{(m)}_j\}$. Перепишем (6) с учетом (7):

$$(z^{(m)}_j \beta'_j + x^{(m)}_j \beta''_j) B_j j - 1 = z^{(m)}_j t^{(m)}_{j-1} A_{j-1} j.$$

Отсюда следует пропорциональность

(10) $$z^{(m)}_j = c x^{(m)}_j$$

с коэффициентом

(11) $$c = \frac{\beta''_j B_j j - 1}{\beta'_j B_j j - 1 - t^{(m)}_{j-1} A_{j-1} j}.$$

В этой и последующих формулах произведения матриц перехода на вектор j равны суммам строк соответствующих матриц и могут быть вычислены до начала итераций.

Подстановка (9) в (7) и умножение обеих частей результата на j дают

$$1 = t^{(m)}_j j = x^{(m)}_j (c\beta'_j + \beta''_j) j.$$

Итак,

(12) $$x^{(m)}_j = 1/[(c\beta'_j + \beta''_j) j].$$

Удобным критерием прекращения итераций является условие

$$\max_j |x^{(m)}_j - x^{(m-1)}_j| \leqslant \varepsilon_1.$$

Предельное отношение смежных вероятностей. Для такого отношения в моделях $A/B/n$ обычно используется формула

(13) $$x_\infty = \rho^{2/(v_A^2 + v_B^2)}$$

(в нашем случае $v_A = 1$). В работе Такахаси [14] получен точный, но довольно сложный алгоритм для расчета x_∞. Сопоставим (таблица 2) предельные отношения вероятностей состояний 10-канальных систем, получаемые по аппроксимации (А) согласно (13) и по формулам Такахаси (Т). Как показали более детальные расчеты, эти отношения не зависят от числа каналов n.

Таблица 2. Оценки предельных отношений вероятностей

ρ	$\beta = 0.25$		$\beta = 1.0$		$\beta = 3.0$		$\beta = 10^9$	
	Т	А	Т	А	Т	А	Т	А
0.7	0.8623	0.8670	0.7000	0.7000	0.5945	0.5857	0.5089	0.4900
0.9	0.9582	0.9587	0.8999	0.9000	0.8549	0.8538	0.8128	0.8100

Значение параметра гамма-распределения 10^9 практически соответствует вырожденному распределению и использовалось для унификации расчетной схемы. Из этой таблицы можно сделать следующие выводы:

1. Представленные в таблице результаты соответствуют качественным ожиданиям: убывают по параметру β, что означает увеличение скорости убывания вероятностей при уменьшении вариации соответствующих распределений.

2. При $\alpha = \beta = 1$, т. е. для марковской системы, они равны коэффициенту загрузки ρ, что подтверждает корректность вычислений.

3. Расхождения в результатах счета по сопоставляемым формулам убывают по коэффициенту загрузки системы и возрастают — с увеличением отклонения β от единицы. По-видимому, для систем с загрузкой $\rho \geqslant 0.7$ может применяться существенно более простой подсчет согласно (13).

Отметим также, что обсуждаемые предельные значения достигаются при достаточно большой длине очереди и никак не могут служить для расчета вероятностей для всех $j > n$.

Безусловные вероятности. После прекращения итераций можно переходить к нахождению абсолютных значений вероятностей. Прежде всего отметим, что из определения чисел $\{x_j\}$ следуют равенства

$$p_{j+1} = p_j x_j, \qquad j = \overline{0, N-1}. \tag{14}$$

Подставив их в условие баланса прибытия и ухода заявок

$$= n - \lambda b,$$

получаем формулу для вероятности свободного состояния системы

$$p_0 = \frac{n - \lambda b}{n + \sum\limits_{j=1}^{n-1} (n-j) \prod\limits_{i=0}^{j-1} x_i}. \tag{15}$$

Последующие вероятности для $j = \overline{1, N}$ определяются рекуррентно с помощью (14). При необходимости та же формула может быть применена для бо́льших значений j с использованием $x_j = x_\infty$.

Начальные приближения. Простейший способ задания начальных векторов вероятностей микросостояний базируется на равновероятном распределении в пределах яруса.

Более продвинутым является восходящее к уже упоминавшемуся аналогу модели $M/H_2/n$ — процессу обслуживания неоднородных заявок — предположение о том, что условное распределение числа обслуживаемых заявок каждого типа является биномиальным с вероятностями каждого типа, пропорциональными $\{y_i/\mu_i\}$.

Можно, наконец, для $j \geqslant n$ принять в качестве начальных приближений предельный при $j \to \infty$ вектор условных вероятностей микросостояний (подробности см. в [4, 5]).

Направление прогонки. В варианте с «предельным» вектором вероятностей микросостояний напрашивается смена направления прогонки – от бо́льших индексов к меньшим (от «хороших» приближений к менее удачным). В этом случае горизонтальный разрез проводится между j-м и $(j-1)$-м ярусами, соответствующие изменения в расчетной схеме алгоритма очевидны.

Численные результаты. Приведем сведения о количестве итераций и (через слэш) трудоемкости в секундах процессорного времени Pentium 4, 2.4 ГГц, для системы обслуживания с комплексными параметрами распределения времени обслуживания — таблица 3. В ней через слэш указаны количество итераций и время счета. Сопоставляются обсуждавшиеся выше варианты задания начальных условных векторов вероятностей микросостояний и различные направления прогонки. Точность оценки времени счета определялась разрешающей способностью системных часов (порядка 0.01 с). Обсчитывалось 70 ярусов диаграммы состояний. Точность стабилизации отношений смежных вероятностей назначалась $\varepsilon = 10^{-6}$.

Более подробное описание и обоснование описанных выше методов приведено в [6]. Они были запрограммированы и дали результаты, совпадающие с высокой степенью точности и подтвержденные перекрестным тестированием на ряде моделей систем обслуживания — см. [7].

Таблица 3. Итерационный обсчет модели $M/E_3/n$

Метод	Число каналов n				
	2	3	5	10	20
Равновер., $\downarrow$	8/0	13/0	23/0	50/0.031	84/0.110
Равновер., $\uparrow$	28/0	37/0	55/0	100/0.047	154/0.218
Бином., $\downarrow$	7/0	12/0	20/0	38/0.031	70/0.109
Бином., $\uparrow$	18/0	23/0	33/0	55/0.031	96/0.141
Lim, равновер., $\downarrow$	8/0	14/0	24/0	49/0.016	78/0.109
Lim, бином., $\downarrow$	8/0	14/0	24/0	52/0.032	78/0.109
Lim, равновер., $\uparrow$	21/0	28/0	42/0	79/0.047	146/0.187
Lim, бином., $\uparrow$	21/0	28/0	42/0	75/0.031	151/0.235

Из опыта расчетов можно сделать следующие выводы:

1. Все рассмотренные варианты позволяют обсчитывать модели с несколькими десятками каналов за десятые доли секунды.

2. Требуемое число итераций возрастает несколько медленнее, чем число каналов.

3. Наименее устойчив в работе метод, использующий предельный вектор микросостояний.

4. Наилучшим вариантом выбора начальных приближений является биномиальный, причем для модели $M/E_3/n$ была предпочтительна прогонка сверху вниз, а для $M/H_2/n$ — снизу вверх.

Анализ результатов расчетов показал, что в случае комплексных параметров H_2-обслуживания компоненты векторов $\{t_j\}$, симметричные относительно центрального элемента вектора, комплексно сопряжены. Это позволяет рассчитывать примерно половину компонент и тем уменьшить трудоемкость. С другой стороны, в случае действительных параметров можно объявить все объекты расчетной процедуры вещественными, что исключит работу с заведомо нулевыми мнимыми частями и снизит трудоемкость минимум вдвое.

4. Матрично-геометрическая прогрессия

Применительно к расчету *разомкнутых* систем обслуживания весьма многообещающим выглядит предложенный Ивэнсом [8] и развиваемый М. Ньютсом и его последователями (см., например, [1,2,9]) метод *матрично-геометрической прогрессии* (МГП). Идея этого метода заключается в представлении векторов вероятностей микросостояний полностью занятой системы соотношением типа

$$\gamma_j = \gamma_n R^{j-n}, \qquad j = n, n+1, \ldots, \tag{16}$$

где R — матричный знаменатель прогрессии. Можно показать, что искомый знаменатель прогрессии должен удовлетворять матричному квадратному уравнению

$$R^2 B - RD + A = 0. \tag{17}$$

Найдя этот знаменатель и имея вектор p_n вероятностей микросостояний n-го яруса, можно согласно (16) вычислить вероятности микросостояний для $j > n$.

Расчет знаменателя прогрессии. Основная проблема реализации МГП-метода заключается в определении знаменателя прогрессии R. *Матричное* уравнение вида (17) может быть решено лишь численными методами. Возможные подходы к решению полиномиальных матричных уравнений обсуждаются в [3]. Все они имеют очень сложную теорию и соответственно — программную реализацию, что вынуждает искать решение в классе итерационных подходов.

Простые итерации могут быть применены в вариантах

$$R = R^2 B D^{-1} + A D^{-1} \tag{18}$$

или

$$R = A(D - RB)^{-1}. \tag{19}$$

Разумеется, постоянные части матричных произведений следует вычислять однократно.

Еще один вариант итерационного подхода — сверхрелаксация — восходит к работе [10], см. также [1,9]. После прибавления к обеим частям уравнения (17) слагаемого $R\omega$ можно получить формулу

$$R = \frac{1}{\omega}[R^2 B - R(D - I\omega) + A]. \tag{20}$$

Параметр ω в [9] предлагалось выбирать равным трем, а в [1] — как минимальный модуль диагонального элемента матрицы D_n. В наших экспериментах оба этих подхода приводили к расходящемуся вычислительному процессу. Обнаружилось, что приемлемым решением является выбор среднего геометрического из крайних элементов D_n. Возможна, наконец, адаптивная версия сверхрелаксации — с периодическим уточнением параметра ω (согласно [12] — после каждых 10 шагов). По нашим данным, с целью минимизации общего числа шагов коррекцию параметра ω следует проводить чаще (после 7 шагов) и менее энергично (не на 25, а на 15%).

Рассмотрим, наконец, «ньютоново» (в линейном приближении) вычисление матричных поправок к знаменателю прогрессии. Определим поправку Δ из условия

$$(R + \Delta)(R + \Delta)B - (R + \Delta)D + A = 0.$$

Раскрывая скобки и пренебрегая квадратом поправки, имеем

$$R^2B + \Delta \cdot RB + R\Delta \cdot B - RD - \Delta \cdot D + A \approx 0,$$

или

$$\Delta \cdot (RB - D) + R\Delta \cdot B \approx RD - R^2B - A.$$

Последнее уравнение приводится к виду

$$\Delta \cdot F + R\Delta \cdot B = G \tag{21}$$

и является *линейным* относительно поправочной матрицы Δ. Однако она участвует в произведении трех матриц как *внутренний* сомножитель, что исключает применение стандартных методов. Это затруднение снимается после выражения коэффициентов развернутой формы системы уравнений (21) через элементы известных матриц A, B и D. Число неизвестных результирующей системы для модели $M/H_2/n$ составляет $(n+1)^2$, что дает оценку трудоемкости каждого шага итераций $O(n^6)$.

По указанной причине имеет смысл определять поправки в методе Ньютона с помощью итерационной формы уравнения (21):

$$\Delta = (R \cdot \Delta \cdot B + A)(D - RB)^{-1} - R. \tag{22}$$

Увеличение числа внутренних итераций уменьшает количество внешних. Численные эксперименты показали, что целесообразно проведение одной внутренней итерации.

Для получения начального приближения к поправке в правой части (22) имеет смысл предположить матрицы R и Δ коммутирующими. Тогда уравнение (22) можно записать в виде

$$2\Delta \cdot RB - \Delta \cdot D = RD - R^2B - A,$$

откуда следует

$$\Delta_0 = (RD - R^2B - A)(2RB - D)^{-1}.$$

В качестве начального приближения во всех рассмотренных вариантах обычно предлагается применять

$$R_0 = xI, \tag{23}$$

где x — обсуждавшееся выше предельное при $j \to \infty$ отношение суммарных вероятностей смежных ярусов. Значительно лучшие результаты дает

$$R_0 = A(D - A)^{-1} = \lambda(D - \lambda I)^{-1}.$$

Расчет начальных вероятностей. При известном знаменателе прогрессии R векторы $\{\gamma_j\}$, $j = \overline{0, n}$, находим решением системы уравнений глобального баланса для микросостояний соответствующих ярусов, дополненной условием баланса заявок:

$$\begin{array}{lcll} \sum_{j=0}^{n-1}(n-j)\gamma_j j & = & n - b/a, & \\ \gamma_0 D_0 & = & \gamma_1 B_1, & \\ \gamma_j D_j & = & \gamma_{j-1}A_{j-1} + \gamma_{j+1}B_{j+1}, & j = \overline{1, n-1}, \\ \gamma_n D_n & = & \gamma_{n-1}A_{n-1} + \gamma_n R B_{n+1}. & \end{array} \tag{24}$$

Практически эту систему перед решением необходимо расписать по компонентам упомянутых в ней векторов. Решение может быть проведено как прямым методом типа гауссова исключения, так и методом итераций. Размер системы для модели $M/H_2/n$ составит

$$N = \sum_{j=0}^{n}(j+1) = (n+1)(n+2)/2.$$

Последующие векторы вероятностей определяются на основе (16) рекуррентно — домножением предыдущего вектора на R.

Трудоемкость определения начальных векторов вероятностей может быть снижена, если посредством введения пересчетных матриц для $\gamma_j, \quad j = \overline{1,n}$ свести эту задачу к определению вектора γ_n [11].

Численные результаты для МГП. Рассмотрим результаты применения МГП к тем же объектам, что при тестировании итерационных методов — на этот раз для модели $M/H_2/n$ с коэффициентом вариации 2. В таблице 4 S означает решение объединенной системы уравнений для компонент начальных векторов вероятностей, а N — только для компонент вектора γ_n.

Таблица 4. «Прогрессивный» обсчет модели $M/H_2/n$

Метод		Число каналов n				
		2	3	5	10	20
$R = R^2BD^{-1} + AD^{-1}$	S	49/0	52/0	54/0	59/0.015	62/0.313
	N	49/0	52/0	54/0	59/0.015	62/0.047
$R = A(D - RB)^{-1}$	S	35/0	36/0	37/0	39/0.016	41/0.328
	N	35/0	36/0	37/0	39/0.016	41/0.047
Ω	S	51/0	56/0	61/0	69/0.016	78/0.359
	N	51/0	56/0	61/0	69/0.016	78/0.063
Ньютон	S	5/0	6/0	5/0	4/0.156	3/7.422
	N	5/0	6/0	5/0	4/0.156	3/7.094

Из таблицы следует:

1. Все итерационные версии метода МГП существенно различаются по числу шагов, но довольно близки по трудоемкости.
2. Их сходимость заметно ухудшается с ростом числа каналов.
3. N-вариант расчета начальных векторов вероятностей сокращает время счета в десятки раз.
4. Ньютонов вариант, как и ожидалось, характеризуется рекордно малым числом итераций и быстрым ростом ростом их трудоемкости.
5. Применение N-варианта расчета начальных векторов незначительно уменьшает время счета по ньютонову варианту, поскольку трудоемкость данного этапа составляет небольшую долю от трудоемкости алгоритма расчета знаменателя прогрессии.

Заключение

В рамках исследований, проведенных по гранту РФФИ 10-08-00906-а и программе фундаментальных исследований ОНИТ РАН (проект 2.3), были выполнены доработки и сопоставление двух методов расчет многоканальных систем «фазового» обслуживания — итерационного и матрично-геометрической прогрессии.

Обоснована возможность и целесообразность использования H_2-аппроксимации обслуживания.

Для итерационного метода рассмотрены новые варианты расчета начальных приближений к векторам условных вероятностей и возможная смена направлений прогонки.

Применительно к методу МГП исследованы варианты расчета знаменателя прогрессии: два вида простой итерации, ω-схема сверхрелаксации, а также уточнение матричных поправок в схеме Ньютона. Разработаны два варианта расчета векторов вероятностей состояний для ярусов $j = \overline{0, n}$: из объединенной системы линейных уравнений и методом прогонки.

Предложенные рекомендации запрограммированы и сопоставлены по числу итераций и времени счета в зависимости от числа каналов.

Из анализа вычислительных схем и сопоставления результатов расчета (в том числе не представленных в данной статье из-за ограниченности места) вытекают следующие **выводы**:

1) Итерационный метод легко модифицируется применительно к системам с интенсивностью входящего потока, зависящей от состояния системы (в частности, замкнутым), и к системам с ограниченной очередью. Он сравнительно легко обобщается на системы, где допустимы переходы между несмежными ярусами диаграммы (например, при потоке групповых заявок, в том числе случайного объема). Благодаря работе с векторами относительных вероятностей и наличию этапа агрегации на каждом ярусе его точность практически не зависит от числа обсчитываемых ярусов. При вещественных параметрах H_2-аппроксимации метод сходился при 100 каналах. Сходимость ухудшается при увеличении числа обсчитываемых ярусов, а также при малых загрузке и коэффициентах вариации (впрочем, при таких данных проблема ожидания снимается сама собой).

Для его реализации целесообразно иметь *три* процедуры: строго комплексную (учитывающую вышеупомянутую симметрию векторов условных вероятностей микросостояний), вещественную и общую — для работы в составе процедур обсчета сетей обслуживания методом потокоэквивалентной декомпозиции.

2) Метод матрично-геометрической прогрессии (МГП) уступает ему по универсальности — он принципиально применим только для QBD-процессов с переходами между *смежными* ярусами, зато его сходимость от количества ярусов не зависит. При большом числе каналов вследствие накопления погрешностей вероятности начальных состояний могут оказаться отрицательными. Как наиболее быстродействующий, он предпочтителен при высоких требованиях к точности; в массовых расчетах; при работе с бесконечными суммами вероятностей; при анализе сетей обслуживания (прежде всего при немарковских входящих потоках с итерационным уточнением последних). Практически он применим при числе каналов $n < 30$. Для МГП-метода аналогично итерационному при достоверно известном типе данных также возможна экономия вычислений.

3) Для расчета систем обслуживания с умеренным (менее 10) числом каналов предпочтителен метод матрично-геометрической прогрессии, а при большем числе каналов — итерационный. В последнем случае имеет смысл ограничить количество обсчитываемых ярусов диаграммы. Вероятности состояний со старшими индексами можно получить последовательным домножением их на предельное значение x, найденное по формуле (13).

4) Приведенные соображения могут быть отнесены и к модели с рекуррентным входящим потоком при H_2-аппроксимации интервалов между смежными заявками — после доработки расчетных схем.

Список литературы

1. Бочаров П.П., Печинкин А.В. Теория массового обслуживания: Учебник. М.: Изд-во РУДН, 1995. 529 с.
2. Вишневский В.М. Теоретические основы проектирования компьютерных сетей. М.: Техносфера, 2003. 512 с.
3. Икрамов Х.Д. Численное решение матричных уравнений. М.: Наука, 1984. 192 с.
4. Рыжиков Ю.И. Алгоритм расчета многоканальной системы с эрланговским обслуживанием // Автоматика и телемеханика. 1980. № 5. С. 30-37.
5. 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.
6. Рыжиков Ю.И. Компьютерное моделирование систем с очередями: Курс лекций. СПб.: ВКА им. А.Ф. Можайского, 2007. 164 с.
7. Рыжиков Ю.И. Пакет программ для расчета систем с очередями и его тестирование // Труды СПИИРАН. Вып. 7. СПб.: Наука, 2008. С. 265-284.
8. Evans R.D. Geometric Distribution in Some Two Dimensional Queuing Systems // Operat. Res. 1967. Vol. 15, No. 5. P. 830-846.
9. Daigle J.N. Queuing Theory with Applications to Packet Telecommunication. Boston: Springer, 2005. 326 p.
10. Latouche G., Ramaswami V. A Logarithmic Reduction Algorithm for Quasi-Birth-and-Death process // J. Appl. Prob. 1993. Vol. 30. P. 650-674.
11. Ryzhikov Yu.I. Realization of a Method of Matrix-Geometric Progression // Proceedings of the Third Conference on Smart Spaces, ruSMART 2010, and 10th International Conference, NEW2AN 2010. Smart Spaces and Next Generation Wired/Wireless Networking. Berlin etc, 2010. Springer. Lecture Notes on Computer Science. 2010. Vol. 6294. P. 266-274.
12. Seelen L.P. An Algorithm for Ph/Ph/c Queues. // Eur. J. of Operational Research. 1986. Vol. 23. P. 118-127.
13. Takahashi Y., Takami Y. A Numerical Method for the Steady-State Probabilities of a GI/G/c Queuing System in a General Class // J. of the Operat. Res. Soc. of Japan. 1976. Vol. 19, No. 2. P. 147-157.
14. Takahashi Y. Asymptotic Exponentiality of the Tail of the Waiting Time Distribution in a Ph/Ph/c Queue // Adv. in Applied Probability. 1981. Vol. 13. P. 619-630.