ИМИТАЦИОННОЕ И ЧИСЛЕННОЕ МОДЕЛИРОВАНИЕ СИСТЕМ С ОЧЕРЕДЯМИ[1]

Ю.И. Рыжиков (Санкт-Петербург)

Возникающие в самых различных сферах человеческой деятельности очереди являются неизбежной платой за совместное использоваие ограниченных ресурсов. Должен быть разумный баланс между затратами на организацию этого обслуживания и ущербом от вынужденного ожидания заявок. Соответствующие расчеты могут быть выполнены на основе численных методов теории очередей или средствами имитационного моделирования.

Основы численных методов

Появление в середине 1950-х гг. электронных вычислительных машин позволило преодолевать нарастающую сложность вычислений и дало мощный толчок развитию численных методов теории массового обслуживания (ТМО), которые, вопреки [1], отнюдь не сводятся к решению единственной системы уравнений. Итоги работы советских ученых подведены в юбилейном выпуске «Автоматики и телемеханики» [2]. Эти методы создали мост между ТМО и потребностями практики. Они способствовали критической оценке предлагаемых расчетных схем на основе анализа результатов их применения, выявления допущенных ошибок, осознания «узких мест». Получили широкое развитие и внедрение метод линейчатых марковских процессов [3–6], метод вложенных цепей Маркова, метод фиктивных фаз в двух вариантах расчетной схемы: итерационная [5,7–9] и матрично-геометрической прогрессии [8,9]. Переосмыслению и развитию этих методов способствовало осознание *законов сохранения теории очередей* и систематическое использование последних [10]. Решающим вкладом в расчет *сетей обслуживания* стал выход из условий известной теоремы ВСМР и переход к потокоэквивалентной декомпозиции сетей [8,11,12].

Подчеркнем, что теория очередей, как всякая теория, формулирует и выводит общие закономерности абстрактной предметной области, которые реализуются независимо от применяемых методов расчета – в частности, при использовании имитационных моделей. К их числу относятся:

- законы сохранения заявок, объема работы и стационарной очереди;
- постоянства взвешенной суммы ожиданий $\sum_i \rho_i w_i$ (инвариант Клейнрока) для консервативных приоритетных дисциплин;
- формула Полячека–Хинчина для среднего времени ожидания в системе M/G/1, позволяющая предвидеть последствия изменения коэффициента загрузки $\rho = \lambda b_1$ и коэффициента вариации обслуживания v_B.

Хорошо отработанные численные методы, как правило, позволяют получить достаточно высокую точность и менее трудоемки, чем имитационные. Перечислим основные результаты, полученные автором и его учениками.

Расчет приоритетных систем

Теория приоритетных систем обслуживания (см., например, [13,14]) – одна из наиболее сложных частей ТМО. К сожалению, авторы этих и многих других работ не были озабочены способами доведения результатов до числа, вследствие чего в упомянутых и иных книгах отсутствуют какие-либо таблицы и графики и расчет, к примеру, заканчивается получением

[1] Исследования проводились по гранту РФФИ 13-08-01250 и программе фундаментальных исследований ОНИТ РАН (проект 2.11).

системы функциональных уравнений для распределений периодов занятости в терминах преобразований Лапласа–Стилтьеса (ПЛС). В [9] значительная часть вычислений продвинута до непосредственной работы с моментами исходных и промежуточных распределений. Удалось значительно упростить расчет систем с повторной реализацией прежней длительности случайного обслуживания, исправить ошибки в теории систем с динамическим приоритетом, обобщить эту теорию на случай динамических приоритетов со стартовыми вкладами.

Попытки переноса методов расчета одноканальных систем на многоканальные к успеху не привели. Однако здесь благодаря обсуждаемым ниже новым методам расчета многоканальных систем удалось развить выдвинутую в [17] идею символических инвариантов отношения средних времен ожидания, применив ее в форме

$$\frac{M/G/1}{M/G/n} = \frac{\vec{M}/\vec{G}/1}{\vec{M}/\vec{G}/n}.$$

В [21] показано, что с ее помощью можно с допустимой погрешностью рассчитывать даже высшие моменты распределений ожидания и пребывания для систем как с абсолютным, так и с относительным приоритетом.

Расчет многоканальных систем методом фиктивных фаз

Классический метод Эрланга применим к системам с простейшим входящим потоком и показательно распределенным временем обслуживания. Однако если первое из названных допущений по опыту и на основании ряда теоретических соображений можно признать вполне реалистическим, то второе в большинстве приложений диктуется лишь соображениями «считаемости». Расчет многоканальных систем практически возможен только после «фазовой» аппроксимации распределения обслуживания с экспоненциально распределенной задержкой в каждой. Научная общественность фактически реализует «фольклорную» (по М. Ньютсу) концепцию: для распределений с коэффициентом вариации, большим единицы, используют гиперэкспоненциальную (H_2-) аппроксимацию, а в остальных случаях – эрлангову.

Эрланговы распределения позволяют строго выравнивать первый и лишь приближенно второй моменты распределения обслуживания. Коэффициент вариации эрлангова распределения порядка k составляет $v = 1/\sqrt{k} < 1$. Ширина диаграмм (максимальное количество микросостояний на ярусе) для моделей с эрланговым обслуживанием быстро растет по числу каналов n и порядку k распределения обслуживания – см. табл. 1.

Таблица 1

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

Число n каналов	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

Для наибольшего из включенных в табл. 1 значения $k = 6$ коэффициент вариации v равен 0.408. При обсчете моделей с меньшими коэффициентами вариации могут потребоваться значительно большие значения k и соответствующий рост ширины диаграммы. Ясно, что возможности этой схемы весьма ограниченны.

С другой стороны, распределение H_2 допускает *любые* коэффициенты вариации (при $v_B < 0.7$ – *с комплексными параметрами*) и имеет ширину диаграммы $n+1$. Обширная серия численных экспериментов показала, что отмеченная патология (для непривычных) проявляется только в *промежуточных* результатах – вероятностях «фиктивных» микросостояний, на которые расщепляются «физические». Поскольку микросостояния фиктивны – почему бы их вероятностям не быть комплексными?! Возврат к физическим состояниям всегда приводил к контрольным вещественным вероятностям из диапазона [0,1]. Из этих соображений развивающие идеи [7] работы [15,16] были сосредоточены на модели $M/H_2/n$. Последняя может быть интерпретирована как обслуживание потока заявок двух типов, причем выбор типа определяет параметр показательно распределенного обслуживания. Здесь «ключ» микросостояния указывает количество находящихся в каналах обслуживания заявок каждого типа. Диаграмма переходов по прибытию заявок для $M/H_2/3$ представлена на рис.1. Прибывающая с интенсивностью λ заявка с вероятностью y_i относится к i-му типу. Прибытие заявок в полностью занятую систему ключ микросостояния не меняет. На рис. 2 показаны переходы по завершению обслуживания.

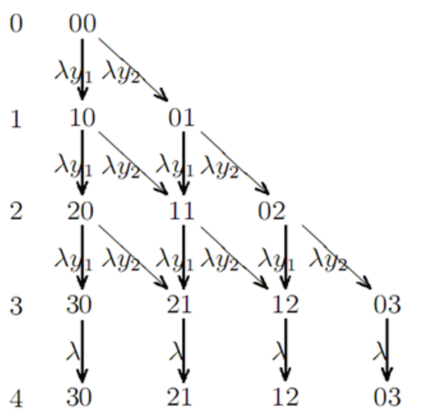

Рис. 1. Переходы по прибытию заявки

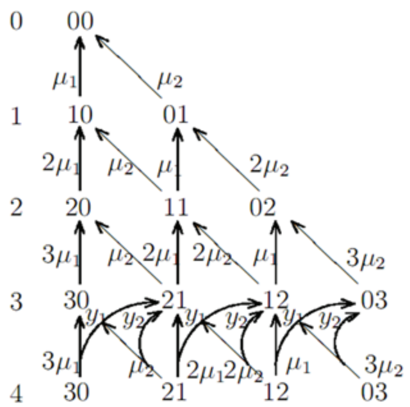

Рис. 2. Переходы по завершению обслуживания

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 с вероятностями $\{y_i\}$, равными долям заявок каждого типа во входящем потоке, в зависимости от типа выбранной из очереди заявки приводит в то или иное микросостояние вышележащего яруса.

Обозначим через S_j множество всех возможных микросостояний системы, 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

$\gamma_j = \{\gamma_{j,1}, \gamma_{j2}, \ldots, \gamma_{j,\sigma_j}\}$ нахождения СМО в состоянии (j,i), $j = 0,1,\ldots$ Теперь можно записать векторно-матричные уравнения баланса переходов между состояниями

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

Решение этой системы может проводиться итерационным методом Такахаси–Таками или методом матрично-геометрической прогрессии. Первый из них основан на переходе к условным (нормированным к единице в пределах яруса) вероятностям микросостояний и итерационном пересчете этих вероятностей и отношений смежных суммарных вероятностей $x_j = p_{j+1}/p_j$ в серии прогонок по всем ярусам. Расчетная схема замыкается предположением, что условные вероятности микросостояний при $j \to \infty$ стабилизируются. После прекращения существенных уточнений $\{x_j\}$ определяется вероятность свободного состояния p_0 и через $\{x_j\}$ вычисляются суммарные вероятности ярусов $\{p_j\}$.

Для этой схемы был предложен ряд модификаций [9,18]. Разработаны улучшенные начальные приближения: биномиальные – пропорционально отношениям $\{y_i/\mu_i\}$, а также предельные для $j > n$. Последние получаются из уравнений баланса между смежными ярусами при вышеупомянутых допущениях относительно $\{x_j\}$ и стабилизации векторов условных вероятностей. Исследовались также влияние направления прогонки, эффект применения метода Люстерника для ускорения сходимости итераций, работа с чисто вещественными переменными при соответствующих исходных данных, использование «антисимметрии» в случае комплексных параметров. Было установлено, что итерации для H_2-аппроксимированной модели $M/E_3/n$ сходятся при $n < 40$, а для $M/H_2/n$ с коэффициентом вариации обслуживания $v_B = 2$ при $n < 250$.

Метод *матрично-геометрической прогрессии* (МГП) основан на предположении, что для ярусов $j > n$ векторы вероятностей микросостояний вычисляются согласно $\gamma_j = \gamma_{j-1} R$, где R – матричный знаменатель прогрессии, удовлетворяющий квадратному уравнению $R^2 B - RD + A = 0$. Последнее может быть решено методом итераций. Были исследованы [19] варианты метода простых итераций и задания начальных приближений, а также уточнение поправок методом Ньютона. В последнем случае вместо решения сильвестрова матричного уравнения на каждом ньютоновом шаге дополнительно применялись *внутренние итерации.*

После расчета знаменателя МГП появляется возможность записать систему линейных уравнений баланса для вероятностей микросостояний ярусов $j = \overline{0, n}$ размерности $(n+1)(n+2)/2$, трудоемкость решения которой также имеет порядок n^6. Трудоемкость данного шага при больших n является фактором, ограничивающим возможность решения задачи. Кроме того, в этом случае разброс модулей переменных измеряется десятками порядков, вследствие чего при стандартной разрядности вычислений (формат двойного слова для вещественных переменных) малые начальные вероятности оказываются отрицательными. С учетом этого обстоятельства в обсуждаемом диапазоне условий «надлежащие» результаты получались до $n = 50$. Была предпринята работа с *учетверенной* разрядностью – табл. 2. В ней данные по времени счета (с) приводятся для тех моделей, где сходился расчет знаменателя МГП.

Таблица 2

Трудоемкость расчета начальных вероятностей (с)

Распределение обслуживания	Разрядность CPLX чисел	Число каналов				
		20	30	40	50	70
E_3	16	0.250	2.075	-	-	-
	32	0.686	7.332	35.30	132.37*	355.6*
H_2	16	0.140	1.810	11.50	42.95*	
	32	6.599	34.88	131.42	977.3	

Расчет 70-канальной системы был прерван на этапе расчета начальных вероятностей по истечении 14 ч процессорного времени. Варианты с отрицательными вероятностями отмечены звездочками. Безумный рост трудоемкости счета при учетверенной разрядности показал явную бесперспективность ее применения при большом числе каналов. По этим причинам представляется перспективной комбинация МГП с итерационным расчетом векторов стартовых вероятностей.

Сети обслуживания

Реально заявка, как правило, нуждается в *комплексном* обслуживании и вынуждена проходить сеть обслуживания со специализированными по видам обслуживания узлами. Численные методы расчета СеМО, разработанные в 50–80-х гг. прошлого века с достойным лучшего применения старанием, опирались на известную теорему ВСМР, условия которой с запросами практики не пересекаются. В СССР сравнительно рано [11,12] была осознана необходимость и возможность потокоэквивалентной декомпозиции сетей. При этом подходе на основе интенсивности внешнего потока Λ и матрицы передач $R = \{r_{i,j}\}$ составляются и решаются уравнения баланса потоков через узлы сети:

$$\lambda_i = \Lambda r_{0\,i} + \sum_{j=1}^{M} \lambda_j r_{j,i}, \qquad i = \overline{1, M}.$$

Далее узлы сети рассчитываются как изолированные системы, и на основе аппроксимаций распределения времени пребывания в узлах гамма-плотностью с поправочным многочленом получаются формулы расчета ПЛС этого времени. Затем происходит агрегация результатов: определяется ПЛС распределения времени пребывания в сети

$$\gamma(s) = P(I - \Gamma(s))^{-1} N(s) T.$$

Здесь P – вектор-строка вероятностей попадания заявки из источника в рабочие узлы, $N(s)$ – ПЛС длительности пребывания в рабочих узлах при однократном заходе, $\Gamma(s) = N(s)Q$, Q – матрица вероятностей переходов между рабочими узлами, T – вектор-столбец вероятностей попадания из рабочих узлов в сток. По таблице значений этой ПЛС в окрестности нуля строится интерполяционный многочлен Ньютона и его дифференцированием находятся (с точностью до знака) моменты искомого распределения. По найденным моментам может быть построена аппроксимирующая ДФР в виде распределения Вейбулла с поправочным многочленом.

В основном варианте алгоритма все потоки предполагались простейшими. В *уточненном* варианте узлы представлялись системами вида $H_2/H_2/n$ с выполнением для каждого узла:

• суммирования входящих потоков (построения распределения интервалов между заявками результирующего);

• расчета распределения интервалов между заявками выходящего потока;

• прореживания этого потока согласно маршрутной матрице при его распределении между преемниками данного узла.

Описанный подход был реализован и для сетей с *неоднородными заявками.* Здесь была выявлена и решена проблема расчета входящих в узлы потоков заявок каждого типа с учетом возможного пропуска некоторыми типами заявок различных в общем случае узлов. Распределение *времени ожидания* в узле заявки конкретного типа считалось по суммарной интенсивности потока и общим средневзвешенным моментам распределения обслуживания, распределение *времени пребывания* – сверткой в моментах распределений общего ожидания и специфического обслуживания.

Перечислим **недостатки** численных методов расчета СМО и практики их применения.

• Сложность используемого математического аппарата, что ограничивает число «бойцов передового фронта науки», подготовку их смены и использование полученных результатов (реально применяются почти исключительно марковские модели).

• «На поверхности» теории очередей легких задач не осталось, а новые (альтернирующие приоритеты, «отрицательные» заявки с негативным влиянием на процессы обслуживания, «нетерпеливые» заявки, системы с повторными вызовами, случаи зависимости длительности обслуживания от времени ожидания, нестационарные задачи и т. д.) с большим трудом поддаются теоретическому анализу.

• Математическая культура массового потребителя результатов ТМО заметно упала. Концепция «непрограммирующих пользователей» свела практически к нулю творческий потенциал большинства студентов и научных сотрудников.

• Рост размерности решаемых задач резко увеличил время счета и ухудшил (в ряде случаев исключил) сходимость итерационных процессов метода Такахаси–Таками и расчета знаменателя МГП.

• Обнаружилась недостаточная точность вычислений при расчете систем обслуживания с большим числом каналов (задача, весьма актуальная для call-центров, применения облачных вычислений и т. п.). Это относится и к методу Кроммелина для M/D/n (расчет корней знаменателя производящей функции вероятностей состояний). В аналогичной ситуации при расчете вероятностей начальных микросостояний из-за громадного разброса модулей их значений (десятки порядков) суммарные вероятности первых состояний системы оказывались отрицательными. Здесь надежды связываются с переходом к итерационному расчету начальных векторов микровероятностей.

Имитационное моделирование

Появление ЭВМ позволило поставить на службу теории очередей применявшийся до того эпизодически метод статистических испытаний. Так появилось имитационное моделирование (ИМ). Его сутью является воспроизведение на ЭВМ детального логического аналога работы моделируемой системы в сочетании со сбором и заключительной (или оперативной – *байесовские* сети) статистической обработкой результатов. Отметим, что первая отечественная книга по ИМ (Н.П. Бусленко) иллюстрирует идею ИМ именно на задаче массового обслуживания. Приходится отметить, однако, что проводимое в [1] его сопоставление с численными методами нуждается в существенной коррекции. В частности, никак нельзя согласиться с тем, что «численные методы дают точные значения неизвестных, а ИМ – их вероятностные характеристики». В обоих случаях речь идет именно о вероятностных характеристиках (например, моментах распределения времени ожидания), но во втором получаются *статистические оценки* последних.

Наиболее обстоятельным руководством по ИМ является книга [22], представленная издателями как «классика Computer Science». К сожалению, эта «классика» изобилует погрешностями, виновника которых (авторы, переводчики или научный редактор) установить трудно, но которые читатель должен иметь в виду. К примеру, на с. 49 утверждается, что к экспоненте при моделировании прибегают потому, что ее легко сгенерировать (генерировать равномерно распределенные числа много проще; на самом деле причиной является ее *марковское свойство*). На с. 53 заявлено, что программа «может иметь синтаксис, отличный от приведенного в листингах» (синтаксис – это *правила языка*, а не текст программы!). На с. 73 смешиваются понятия статики и стационарности, на с. 260 и 264 – статики и статистики. На с. 295 для иллюстрации различия в дисперсиях приведены два графика плотностей распределения с существенно различными площадями под кривыми.

Утверждается, что «медиана может быть лучшей мерой стремления к центру, чем среднее» (с. 294), а коэффициенты симметрии являются показателями…асимметрии (с. 309). На с. 317 написано, что критерии оценки показателей должны включать ЛПР (оно должно *включаться в выработку критериев*!). Заявлено (с. 318), что «имитация и ее результаты достоверны, если менеджер и другие руководители проекта признают их правильными. Из текста на с. 685 следует зависимость настоящего от будущего и т. д.

Ниже изложены предложения автора, приведенные в [23,24] и многократно применявшиеся в решении исследовательских задач и в учебном процессе.

Опыт **прямого** моделирования систем и в особенности сетей обслуживания на универсальных языках численного программирования (современный Фортан) подсказал ряд полезных технологических приемов:

• Обязательное использование *паспорта заявки*, представляющего ее в очереди и в каналах обслуживания.

• Иерархическое построение цепей будущих событий (ЦБС), к примеру по уровням: ближайшие события каждого вида; ближайшие моменты завершений обслуживания по узлам; моменты завершения обслуживания во всех каналах данного узла, моменты прибытия очередных заявок каждого типа). Разумеется, при выборе соответствующих минимумов фиксируются и их координаты.

• Кольцевое построение очередей (в этом случае нет необходимости сдвигать очередь: достаточно изменить значения указателей на голову и хвост очереди).

• Метод исследования сети посредством «меченых заявок» реализуется с помощью ключа KEY – аналога семафора Дийкстры. Первая заявка, пришедшая в сеть с KEY=0, получает в паспорте метку Z=1 и устанавливает KEY=1. Все последующие получают Z=0. Заявки, попавшие в сток, обрабатываются стандартным образом. Если такая заявка имеет признак Z=1, для нее собирается отдельная статистика, после чего вновь устанавливается KEY=0. Этот способ обеспечивает одновременное нахождение в сети *не более одной меченой заявки*, что исключает любую корреляцию в поведении таковых и облегчает обработку статистики.

• Подпрограмма вычисления равномерно распределенного псевдослучайного числа на основе двоичного разложения его номера позволяет гарантировать непересекаемость серий случайных чисел, генерируемых в различных целях, что считается обязательным условием корректности процесса моделирования.

Достоинствами ИМ являются:

• легкость учета особенностей ситуаций, практически не поддающихся численному анализу (многоканальные приоритетные системы; системы с немарковским нетерпением; нестационарные задачи – через получение временных сечений процесса по множеству реализаций; процессы расщепления и слияния заявок);

• наглядность реализующей модель программы (при необходимости – и процесса реализации), легкость интерпретации результатов, что облегчает отладку;

• простота освоения основных идей и соответствующих технологий.

Имитационное моделирование позволяет оценивать работу практически любых систем и сетей обслуживания и делает его (как поставщика первичных данных и эталонных результатов) незаменимым помощником при разработке численных методик.

Недостатки имитационного моделирования:

• каждый прогон ИМ – это частный результат, по ним труднее делать обобщения;

• ввиду неидеальности используемых датчиков псевдослучайных равномерно распределенных чисел надежды на принципиальную возможность добиться сколь угодно высокой точности увеличением числа испытаний безосновательны;

• получаемые оценки искомых величин имеют статистическую погрешность;

• имитационные модели неудобны для оптимизации, поскольку выигрыш в значении целевой функции может быть перекрыт статистическими и инструментальными погрешностями;

• прогон модели, как правило, имеет много большую трудоемкость, чем ее численный обсчет (если такая возможность имеется);

• относительная погрешность определения вероятностей редких событий велика и часто не может быть снижена до приемлемого значения при разумном числе испытаний;

• ветви алгоритма моделирования находятся в сильной взаимной зависимости, что затрудняет обнаружение ошибок и затягивает процесс отладки.

Отметим также существенную для соискателей ученых степеней «малую наукоемкость» ИМ. Идея метода столь проста, что внести в него полезную новацию, важную и *неочевидную для специалиста* (чего справедливо требуют диссертационные советы), весьма нелегко. Об этом можно судить по приведенному выше перечню новых технологических приемов.

Отдельного обсуждения заслуживает сопоставление «авторского» моделирования (например, на Фортране) и работы в моделирующих средах (GPSS World, AnyLogic и т. п.). Такие среды в современном исполнении предоставляют пользователю удобный интерфейс вплоть до графической сборки программы, а также средства многоаспектной визуализации процесса и его результатов. Это ускоряет подготовку и проведение имитационных экспериментов и облегчает анализ результатов. Следует, однако, иметь в виду, что все такие системы имеют встроенный и защищенный от изменений интерпретатор программы, который *принципиально не является универсальным* и для которого нельзя дать исчерпывающий перечень того, чего он *не может* делать. Кроме того, стремление разработчиков к расширению применимости приводит к непомерной избыточности порождаемых программ и как следствие – к серьезному замедлению их работы. Для иллюстрации его степени приведем данные по длительности прогона на GPSS World одной и той же модели СМО:

• все окна закрыты – 4 с;

• выводится таймер – 97 с;

• открыты окно блоков и таймер – 5271 с (почти полтора часа).

Та же модель на Фортране была реализована за 0.05 с – в 80 раз быстрее «безоконной» версии!

Отметим, наконец, что обучение работе исключительно с имитирующими средами требует освоения множества технических деталей, но не способствует уяснению философии и сути моделирования и развитию творческого потенциала обучаемых. Оно, безусловно, противоречит требованию *фундаментальности* образования.

Заключение

Перечислим выводы из вышеизложенного.

1. Задачи расчета систем и сетей с очередями были, есть и останутся актуальными как в теоретическом, так и в практическом смысле.

2. В области своей применимости (там, где они существуют и реализованы программно) и (обязательно!) *в умелых руках* благодаря своему быстодействию и точности предпочтительны *численно-аналитические* методы. Как показал опыт автора, еще в 1987 г. передавшего в Государственный фонд алгоритмов и программ пакет МОСТ по расчету систем с очередями, для использования подобных пакетов часто требуется тонкая настройка, реализуемая через коррекцию текстов программ. Таким образом, работа с ними требует по крайней мере *авторского сопровождения* – если не исполнения. Классик теории очередей М. Ньютс именно из этих соображений отвергал предложения реализовать свои разработки как общедоступный пакет подпрограмм.

3. Развитие численных методов приблизилось к пределу своих возможностей, и дальнейший их прогресс дается все с большим трудом. Новые работы в этом направлении должны всячески приветствоваться. Однако научная общественность вправе требовать от их авторов *явных доказательств работоспособности и завершенности* представленных «изделий»: таблиц и графиков результатов счета, а также подтверждений их корректности сопоставлением с полученными другими методами, в том числе посредством имитационного моделирования.

4. Обучение теории очередей в подавляющем большинстве случаев ограничивается изложением метода Эрланга. Серьезное освоение численных методов теории очередей должно быть ориентировано на достаточно общие случаи (немарковские задачи) и строиться на базе обстоятельных курсов вычислительной математики и прикладного программирования – лучше всего на современном Фортране. Пятидесятилетний педагогический опыт автора в Военно-космической академии им. А.Ф. Можайского показал, что работа с пакетами *готовых* подпрограмм (в том числе с разработанным автором МОСТом) не способствует пониманию *сути* изучаемых алгоритмов: гораздо полезнее осознание уравнений баланса, правил построения диаграмм перехода и соответствующих матриц, программирование ключевых фрагментов задач. Кстати, то же самое можно сказать о попытках свести преподавание *вычислительной математики* к работе с MathCAD’ом и его аналогами.

Для специальностей с умеренной математической подготовкой (инженеры-экономисты), возможно, достаточно ограничиться теорией марковских систем и сетей, но не по Эрлангу, а на базе законов сохранения и обязательно – с упоминанием о существовании более общей теории, ее особенностях и возможном отличии результатов. С другой стороны, студенты математических факультетов должны уметь получать *численные* результаты расчетов.

5. Имитационные методы полезны при решении нестандартных задач с усложняющими факторами (нетерпеливые и «отрицательные» заявки, поломка каналов, их сменная работа, нестационарность, неполнодоступность, наличие операций типа «fork – join» и т. п.) для верификации новых численных методов и поиска полезных аппроксимаций. Они незаменимы при разработке *тренажеров*.

6. Техника ИМ всегда будет нуждаться в дальнейшем развитии в направлениях учета специфики моделей, их агрегирования и рационализации (пример – избавляющая от сдвигов кольцевая организация очередей). Для этого развития необходима «авторская» имитация.

7. Обучение имитационному моделированию обязательно должно начинаться с основ «авторской» имитации – для понимания идеи и возможностей ИМ. Имитационные среды могут использоваться при выполнении курсовых и дипломных работ. Разумеется, они остаются полезным инструментом решения особо сложных проектно-конструкторских и производственных задач.

Литература

1. **Бабина О.И.** Сравнительный анализ имитационных и аналитических моделей // Материалы IV Всероссийской научно-практической конференции по имитационному моделированию и его применению в науке и промышленности. Сб. докладов. Т. 1. СПб. 2009. С. 73–77.
2. **Яшков С.Ф.** Столетие теории очередей: Предисловие к тематическому выпуску. // Автоматика и телемеханика. 2009. № 12. С. 38.
3. **Севастьянов Б.А.** Эргодическая теорема для марковских процессов и ее приложение к телефонным системам с отказами // Теория вероятностей и ее применения. 1957. Т. 2. Вып.1. С. 106–116.
4. **Рыжиков Ю.И.** Управление запасами. М.: Наука, Физмат, 1969. 344 с.
5. **Рыжиков Ю.И.** Машинные методы расчета систем массового обслуживания. Л.: ВИКИ им. А. Ф. Можайского, 1979. 177 с.
6. **Рыжиков Ю.И.** Расчет систем массового обслуживания с порогом включения и разогревом // Известия АН ССР. Техническая кибернетика. 1974, № 6. С. 125–131.
7. **Takahashi Y., Takami Y.** A numerical method for the steady-state probabilities of a GI/G/c queueing system in a general class // J. of the Operat. Res. Soc. of Japan, 1976. Vol. 19. № 2. P. 147–157.
8. **Рыжиков Ю.И.** Компьютерное моделирование систем с очередями: курс лекций. СПб.: ВКА им. А. Ф. Можайского, 2007.164 с.
9. **Рыжиков Ю.И.** Алгоритмический подход к задачам массового обслуживания: монография. СПб.: ВКА им. А. Ф. Можайского, 2013. 496 с.
10. **Krakowski M.** Conservation methods in queuing theory // Revue Française d'Automatique, Informatique et Récherhe Opérationnelle. 1973. Vol . P. 63–84.
11. **Башарин Г.П., Бочаров П.П., Коган Я.А.** Анализ очередей в вычислительных сетях. Теория и методы расчета. М.: Наука, Физмат, 1989. 336 с.
12. **Вишневский В.М.** Теоретические основы проектирования компьютерных сетей. М.: Техносфера, 2003. 512 с.
13. **Климов Г.П.** Стохастические системы обслуживания. М.: Наука, 1966. 243 с.
14. **Приоритетные системы обслуживания**. М.: МГУ, 1973. 447 с.
15. **Рыжиков Ю.И.** Алгоритм расчета многоканальной системы с эрланговским обслуживанием // Автоматика и телемеханика. 1980. № 5. С. 30–37.
16. **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.
17. **Бронштейн О.И., Духовный И.М.** Модели приоритетного обслуживания в информационно-вычислительных системах. М.: Наука, 1976. 220 с.

18. **Рыжиков Ю.И.** Итерационный метод расчета многоканальных систем обслуживания – основы, модификации и предельные возможности // Труды 9-й Российской мультиконференции по проблемам управления. Информационные технологии в управлении. ГНЦ РФ ОАО "Концерн "Электроприбор", СПб., 2016. С. 224–233.

19. **Рыжиков Ю.И.** Модификации и перспективы метода матрично-геометрической прогрессии // Там же. С. 234–243.

20. **Бочаров П.П., Печинкин А.В.** Теория массового обслуживания: учебник. М.: Изд-во РУДН, 1995. 529 с.

21. **Рыжиков Ю.И., Хомоненко А.Д.** Расчет многоканальных приоритетных систем с применением инвариантов отношения // Интеллектуальные технологии на транспорте, 2015. № 4. С. 34–39.

22. **Кельтон В.Д., Лоу А.М.** Имитационное моделирование /пер. с англ. СПб.: Питер; Киев, BHV. 2004. 847 с.

23. **Рыжиков Ю.И.** Имитационное моделирование. Теория и технологии. СПб.: КОРОНА Принт, 2004. 380 с.

24. **Рыжиков Ю.И.** Имитационное моделирование: курс лекций. СПб.: ВКА им. А. Ф. Можайского, 2007. 125 с.