

## АЛГОРИТМ СТРУКТУРНОЙ ОПТИМИЗАЦИИ СЕТЕЙ С ОЧЕРЕДЯМИ

В. Н. Задорожный, А. А. Донец (Омск)

Имитационное моделирование (ИМ) предоставляет универсальные возможности для исследования систем с учетом разнообразных свойств и взаимосвязей их элементов. За это приходится расплачиваться большими расходами машинного времени на выполнение имитационных экспериментов (ИЭ), в том числе (и особенно) при решении оптимизационных задач. Хорошую помощь в решении проблем оптимизации в ИМ могут оказать простые эмпирические аппроксимации отклика [1], позволяющие ориентироваться в факторном пространстве (ФП) модели при минимальном числе ИЭ. В практических приложениях ИМ требования к точности используемых при оптимизации приближений весьма умерены. Часто бывает достаточно того, чтобы пошаговый процесс оптимизации приводил к решению, заметно лучшему, чем выбранное случайно.

Для задачи оптимального распределения каналов по узлам сети с очередями, решаемой в дискретном ФП, предлагается аналитико-имитационный метод, использующий эмпирическую аппроксимацию, и исследуется его эффективность.

**Постановка задачи**

Некоторая сеть с очередями, узлами которой являются системы массового обслуживания (СМО) типа  $\bullet | GI | n | \infty$ , усиливается путем ввода в ее состав  $M$  дополнительных каналов. Все каналы в сети технически идентичны, т. е. любой из вводимых каналов может быть добавлен в состав любой СМО. Различие вероятностных распределений времени обслуживания заявок в разных узлах связано с разной трудоемкостью соответствующих этапов обслуживания заявок. Каналы в составе одной СМО характеризуются одинаковым вероятностным распределением времени обслуживания. Целью усиления сети является сокращение среднего времени  $E$  пребывания в ней заявок (времени ответа). Задача состоит в нахождении оптимального распределения  $M$  дополнительных каналов по узлам, обеспечивающего минимум времени ответа (среднего)  $E$ .

Методы решения этой задачи рассмотрим на примере тестовой сети (рис. 1). Слева на рис. 1 изображена исходная сеть, для которой  $E = 15,8$  (единицы времени условные), справа – сеть, усиленная тремя дополнительными каналами, добавленными в СМО 4, 5 и 7. Интенсивность входного потока сети (пуассоновского)  $\Lambda = 1$ . Функции распределения вероятностей (ф.р.в.)  $B_i(t)$  для времени  $x_i$  обслуживания заявок в узлах  $i = 1, \dots, 9$  указаны в табл.1, где  $R$ ,  $M$  и  $E^2$  обозначают, соответственно, равномерное (на интервале от нуля до двух средних), экспоненциальное и второго порядка эрланговское распределения. Переходные вероятности сети следующие:  $p_{0,1} = 0,2$ ,  $p_{0,2} = 0,3$ ,  $p_{0,3} = 0,5$ ,  $p_{2,4} = 0,7$ ,  $p_{2,5} = 0,3$ ,  $p_{4,6} = 0,3$ ,  $p_{4,7} = 0,4$ ,  $p_{4,9} = 0,3$ ,  $p_{5,8} = 0,9$ ,  $p_{5,9} = 0,1$ .

Таблица 1

**Параметры узлов исходной сети**

Номер $i$ узла	1	2	3	4	5	6	7	8	9
Число каналов $K_i$	1	2	1	1	1	1	1	3	1
Тип ф.р.в. $B_i(t)$	$R$	$R$	$R$	$M$	$M$	$E^2$	$E^2$	$E^2$	$R$
Интенсивность обслуживания $b_i^{-1} = 1/M(x_i)$	0,3	0,3	0,8	1,8	7,25	0,8	0,8	5,5	2,25

В рассматриваемом тестовом примере требуется найти оптимальный вариант усиления сети при числе дополнительных каналов  $M = 3$ .

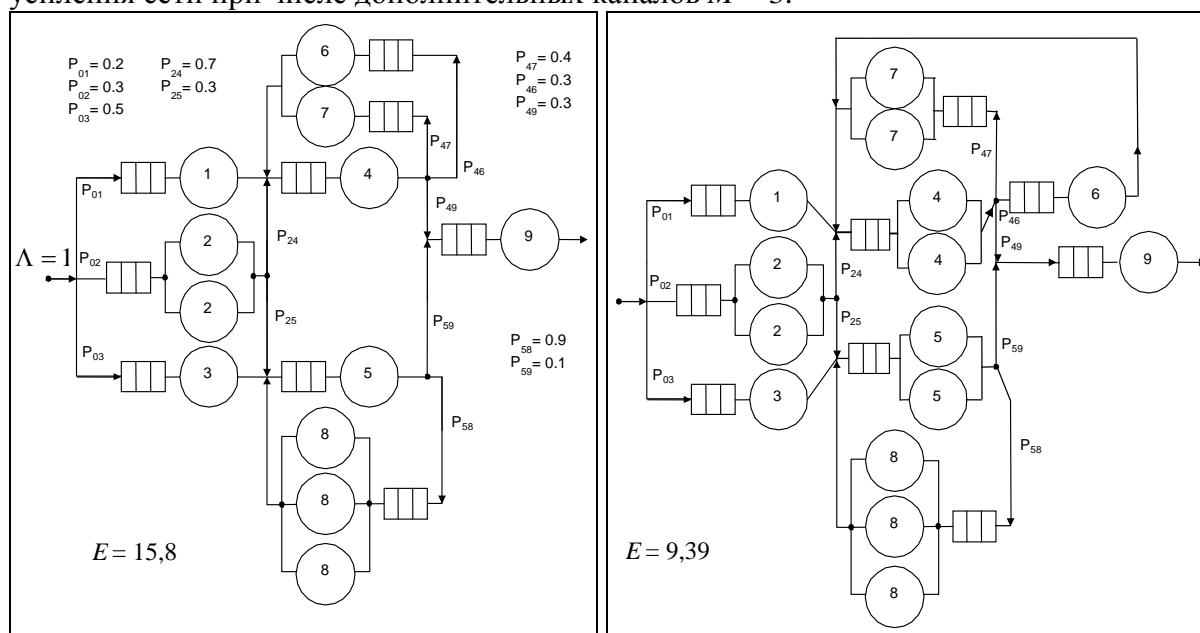


Рис. 1. Исходная тестовая сеть (слева) и сеть с усиленными узлами (справа)

### Методы решения

Путем ИМ приведенной исходной сети на GPSS World наряду со временем ответа  $E$  определены значения среднего времени ожидания  $w_i$  в ее узлах (табл. 2). Длина прогонов модели определялась прохождением через сеть 1 000 000 заявок, что обеспечило приемлемые относительные погрешности для всех используемых оценок, лежащие на уровне 1–3%. Наименьшее возможное время ответа, определяемое моделированием сети с бесконечным числом каналов в узлах, составляет  $E_- \approx 6,47$ .

Таблица 2

### Среднее время ожидания в узлах исходной сети и в узлах пробной сети

Номер $i$ узла	1	2	3	4	5	6	7	8	9
Среднее время ожидания $w_i$	4,54	0,774	1,415	1,57	0,579	1,02	2,19	0,006	0,353
$w'_i$ (для пробной сети)	0,295	0,113	0,097	0,086	0,027	0,071	0,136	0,001	0,023

В принципе, оптимальный вариант распределения  $M$  каналов по  $N$  узлам сети может быть найден полным перебором всех  $C_{N+M-1}^M$  возможных вариантов распределения. В случае сети, представленной на рис. 1, это при  $M = 3$  составляет 165 вариантов. В результате их ИМ (перебор конфигураций сети осуществлялся автоматически с помощью внешней для GPSS программы) был найден глобально оптимальный вариант, который и представлен на рис. 1 справа. При этом время ИЭ на персональном компьютере средней мощности составило более 13 ч. Локальные характеристики оптимально усиленной конфигурации сети приведены в табл. 3.

Таблица 3

## Характеристики узлов оптимально усиленной сети

Номер $i$ узла	1	2	3	4	5	6	7	8	9
Число каналов $K_i$	1	2	1	2	2	1	2	3	1
Среднее время ожидания $w_i$	4,48	0,762	1,403	0,075	0,022	1,05	0,123	0,005	0,336

Очевидно, что в общем случае метод полного перебора неприемлем, поскольку число вариантов  $C_{N+M-1}^M$  резко возрастает при увеличении значений  $N$  и  $M$ .

Перебор можно значительно сократить, используя квазиградиентный метод поиска. Начиная с некоторого произвольного начального распределения  $M_1, \dots, M_N$ , где  $M_i \geq 0$  – число дополнительных каналов, вводимых в  $i$ -й узел,  $\sum M_i = M$ , можно сравнивать время ответа  $E(M_1, \dots, M_N)$  при таком распределении со временем ответа  $E(M'_1, \dots, M'_N)$  для всех «соседних» распределений  $M'_1, \dots, M'_N$ , отвечающих условию соседства  $\sum |M_i - M'_i| = 2$ ,  $\sum M'_i = M$ .

Далее наилучшее соседнее распределение, если оно лучше начального, используется в качестве нового начального распределения, и процесс продолжается, пока очередное начальное распределение не окажется лучше любого своего соседа. Этот метод требует на первой итерации выполнить  $C_N^2$  обращений к модели для оценки соседних вариантов и на последующих итерациях  $C_N^2 - 1$  обращений. Очевидно, что здесь переход к соседнему распределению отображает перемещение одного канала из узла, где его наличие наименее эффективно, в узел, где он наиболее нужен. Поскольку на последующих итерациях отмена подобных решений маловероятна, то общее число итераций квазиградиентного метода близко к числу  $M$  дополнительных каналов. Опыты с разными начальными условиями показывают, что число итераций вместе с моделированием неулучшаемой конечной конфигурации, как правило, составляет  $M + 1$ . Поэтому с учетом необходимости оценивать на каждой итерации соседние распределения общее количество  $K$  обращений к модели приблизительно равно  $K \approx (M + 1) \cdot (C_N^2 - 1)$ .

При оптимизации квазиградиентным методом тестовой сети в зависимости от начального распределения выполняется до  $K = 141$  обращений к модели, и процесс сходится к глобально оптимальному варианту, представленному на рис. 1 справа.

Еще большего сокращения перебора можно достичь последовательным методом размещения дополнительных каналов. Метод состоит в том, что вначале выбирается наилучшая позиция для одного канала, затем для следующего и т. д., пока не будут размещены все дополнительные каналы. В этом случае необходимо  $K = MN$  обращений к имитационной модели. Этот метод для вышеприведенной сети при числе обращений к модели  $K = 27$  также привел к глобально оптимальному решению (для дополнительных каналов последовательно были выбраны 5-й, 4-й и 7-й узлы) и потребовал около 2,5 ч машинного времени. Дополнительное преимущество последовательного метода состоит в том, что лицо, принимающее решение, может не фиксировать жестко число дополнительных каналов, а определять его в процессе сопоставления текущего числа размещенных каналов и достигаемого при их добавлении эффекта.

**Использование эмпирической аппроксимации**

Рассмотренные выше методы оптимизации не позволяют сократить время ИМ настолько, чтобы можно было рекомендовать их для применения к сетям достаточно большого размера, содержащим десятки или сотни узлов. Наиболее радикального сокращения числа необходимых обращений к ИМ здесь можно добиться с помощью эмпирической аппроксимации зависимости  $w(n)$ , предложенной в [2], и имеющей следующий вид:

$$w(n)/w(1) = (1/n)^{1/k}, \quad (1)$$

где  $w(n)$  – среднее время ожидания в СМО при наличии в ней  $n$  каналов;  $w(1)$  – время ожидания при наличии одного канала;  $k$  – параметр, который может быть подобран с помощью ИМ. Законы распределения интервалов поступления и обслуживания заявок при изменении числа каналов предполагаются в (1) неизменными. Проведенное нами тестирование этой формулы для разнообразных систем  $GI | GI | n | \infty$  показало, что при правильном подборе коэффициента  $k$  рассчитанные по формуле графики зависимости  $w(n)$  в широком диапазоне изменения числа каналов  $n$  не отклоняются существенно от графиков, полученных с помощью ИМ.

Чтобы формулу можно было применять и в случаях, когда при  $n=1$  стационарный режим отсутствует, представим ее в следующем виде, вытекающем из (1):

$$w(n) = w(m)(m/n)^{1/k},$$

и, достаточно произвольно распространяя область ее применения на узлы сети (формула предложена для изолированных СМО), в виде

$$w_i(n_i) = w_i(m_i)(m_i/n_i)^{1/k_i}, \quad i = \overline{1, N}. \quad (2)$$

Выполнив ИМ сети при некоторых фиксированных  $m_i = m_i^*$  и  $n_i = n_i^*$ ,  $i = \overline{1, N}$ , из (2) можно определить параметры  $k_i$ :

$$k_i = \frac{\ln m_i^* - \ln n_i^*}{\ln w_i(n_i^*) - \ln w_i(m_i^*)}, \quad i = \overline{1, N}, \quad (3)$$

и использовать их вместе с известными из ИМ  $w_i(m_i^*)$  для расчета  $w_i(n_i)$  при различных других  $n_i$ . Используя (2), можно сократить число обращений к ИМ до двух (независимо от  $N$  и  $M$ ). Алгоритм оптимизации включает в этом случае два основных шага.

*Первый*, подготовительный, шаг сводится к ИМ исходной сети и ИМ пробной сети, получаемой из исходной сети увеличением на единицу числа каналов во всех узлах одновременно. В табл. 2 (см. выше) приведено время ожидания  $w'_i$  в узлах пробной сети. Данные табл. 2 позволяют определить параметры  $k_i$  с помощью соотношения (3) и, таким образом, далее просто вычислять время ответа  $E$  (для различных распределений дополнительных каналов по узлам) по следующей формуле:

$$E = \sum_{i=1}^N \alpha_i \cdot [w_i(n_i) + b_i], \quad (4)$$

где  $w_i(n_i)$  вычисляется с помощью аппроксимации (2) при общем числе  $n_i$  назначаемых в узел каналов (вместе с дополнительными каналами);  $b_i$  – среднее время обслуживания заявок каналом  $i$ -го узла;  $\alpha_i$  – среднее число посещений  $i$ -го узла одной заявкой, которое легко определяется путем решения системы уравнений баланса. Например, для тестовой сети имеем  $(\alpha_1, \dots, \alpha_9) = (0,2, 0,3, 0,5, 1,3667, 5,9, 0,41, 0,54667, 5,31, 1,00)$ .

Второй, расчетный, шаг состоит собственно в выборе оптимального распределения каналов одним из рассмотренных выше методов. Но при этом, используя (4) и (2), уже можно обойтись без ИМ.

Применение описанного двухшагового метода к тестовой сети (при последовательном алгоритме размещения каналов на расчетной стадии) привело к тому же известному решению – для трех дополнительных каналов также последовательно выбирались 5-й, 4-й и 7-й узлы. При этом аппроксимация (4), (2) показала для оптимального варианта значение  $E = 9,47$ , которое отличается от полученного путем проверочного ИЭ значения  $E = 9,39$  менее чем на 1%.

В общем случае, возможно, могут понадобиться дополнительные обращения к ИМ, если на расчетной стадии число каналов в каких-либо узлах значительно превысит значения, использованные для начальной настройки аппроксимации. Тогда применение дополнительного прогона позволит скорректировать параметры  $k_i$ .

Рассмотренная задача решалась также в модифицированном варианте, когда интервал  $\tau$  поступления заявок в сеть разыгрывался по формуле  $\tau = a[(\eta + \varepsilon)^{-1} - (1 + \varepsilon)^{-1}]$ , где  $\eta$  – стандартная случайная величина, равномерно распределенная на интервале  $(0,1)$ ,  $a = 0,1692$ ,  $\varepsilon = 0,001$ , т.е.  $\tau$  имеет ограниченное гиперболическое распределение вероятностей с математическим ожиданием 1 и коэффициентом вариации 5,25. Число  $M$  заранее не фиксировалось. Первые три канала попали в узлы 5, 4 и 3. В исходной сети  $E \approx 54,5$  (вычислено путем ИМ), в усиленной тремя каналами сети  $E \approx 19,1$  (если каналы добавить в узлы 5, 4, 7, то  $E$  составит 20,56). Четвертый канал попал в узел 1. Погрешность аппроксимации (4), (2) здесь увеличилась. Однако сравнение вариантов по расчетным оценкам по-прежнему хорошо коррелирует со сравнением вариантов по оценкам ИМ (рис. 2 справа), а снижение времени ответа после каждого добавления канала (проверяемое посредством ИМ) получается весьма существенным (рис. 2 слева).

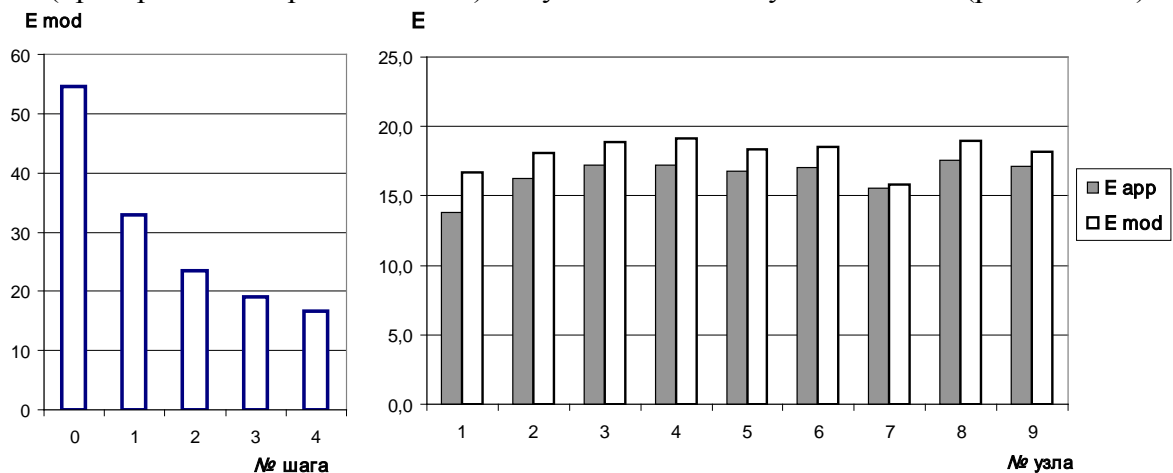


Рис. 2. Изменения  $E$  при последовательном добавлении каналов (слева) и сопоставление расчетных оценок для вариантов размещения пятого канала с оценками ИМ (справа)

### Выводы

Предлагаемый в статье метод можно применять для структурной оптимизации сети, содержащей сотни узлов и характеризуемой произвольными распределениями интервалов поступления и обслуживания заявок. Он не гарантирует нахождения глобального оптимума, но, как показывают тестовые испытания, позволяет при усилении сети дополнительными каналами существенно снижать время ответа и достигать его значений, близких к тем, которые достигаются в точке глобального оптимума. Возможны уточнения метода в плане определения параметров аппроксимации  $k_i$ . Основным достоинством метода является его скорость: он позволяет аналитически оптимизировать распределение каналов, основываясь на ИМ лишь двух конфигураций сети.

### Литература

1. **Рыжиков Ю. И.** Имитационное моделирование. Теория и технологии. СПб.: КОРОНА принт. М.: Альтекс-А, 2004. 384 с.
2. **Рыжиков Ю. И.** Имитационное моделирование в обосновании методик расчета многоканальных приоритетных систем//Материалы I Всероссийской конференции (ИММОД-2003). С.-Петербург. 23–24 окт. 2003. ФГУП ЦНИИТС. Режим доступа <http://www.gpss.ru/immod%2703/index.html>