

22. Винер Н. Кибернетика или управление и связь в животном и машине. М.: Сов. радио, 1958.
23. Охтилев М. Ю., Соколов Б. В., Юсупов Р. М. Возможный подход к созданию единой информационно-вычислительной среды для системы воздушно-космической обороны // Вопросы оборонной техники: науч.-техн. сб. 2010. Сер. 9, вып. 1(242)—2(243). С. 85—90.

**Сведения об авторах**

- Михаил Юрьевич Охтилев** — д-р техн. наук, профессор; СПИИРАН, лаборатория информационных технологий в системном анализе и моделировании; E-mail: oxt@mail.ru
- Николай Габдрахманович Мустафин** — канд. техн. наук, доцент; Санкт-Петербургский государственный электротехнический университет „ЛЭТИ“, кафедра автоматизированных систем обработки информации и управления; E-mail: nikolay.mustafin@gmail.com
- Владимир Евгеньевич Миллер** — канд. техн. наук; ОАО Радиотехнический институт им. акад. А. Л. Минца, Санкт-Петербург; директор филиала; E-mail: miller@progsystema.ru
- Борис Владимирович Соколов** — д-р техн. наук, профессор; СПИИРАН, лаборатория информационных технологий в системном анализе и моделировании; зам. директора по научной работе; E-mail: sokol@iias.spb.su

Рекомендована СПИИРАН

Поступила в редакцию  
10.06.14 г.

УДК 519.872

Ю. И. РЫЖИКОВ

**ОПТИМИЗАЦИЯ МАРШРУТНОЙ МАТРИЦЫ  
В СЕТЯХ ОБСЛУЖИВАНИЯ**

Описан алгоритм расчета временных характеристик разомкнутой сети обслуживания. Предложен метод оптимизации сети обслуживания по среднему времени пребывания заявки в сети путем выравнивания загрузки узлов. Приводятся и обсуждаются результаты численного эксперимента.

**Ключевые слова:** разомкнутая сеть, время пребывания, выравнивание загрузки узлов.

**Расчет сети.** Реальные процессы обслуживания связаны с прохождением нескольких его этапов, реализуемых в отдельных узлах сети. Сеть обслуживания состоит из рабочих узлов, пронумерованных от 1 до  $M$ , источника (узел „0“) и стока (узел „ $M+1$ “). Для каждого  $j$ -го узла задаются моменты распределения „чистой“ длительности обслуживания  $\{b_{j,l}\}$ ,  $l = \overline{1, L}$ , число каналов  $n_j$  и дисциплина обслуживания. Маршрут заявки в сети определяется неразложимой матрицей передач  $R = \{r_{i,j}\}$ ,  $i, j = \overline{0, M+1}$ , образованной вероятностями перехода из  $i$ -го узла в  $j$ -й. Важнейшей оперативной характеристикой работы сети является среднее время пребывания в ней заявки. Первым шагом процесса оптимизации сети должна быть минимизация этого времени.

Проблема расчета сетей обслуживания активно обсуждается в сотнях статей и монографий (см., например, список литературы в работе [1]). К концу 1980-х гг. выяснилось, что строгое решение этой задачи возможно лишь при весьма ограниченных условиях теоремы ВСМР (Baskett, Chandy, Muntz, Palacios [2]). Методы решения были непомерно трудоемкими, а получаемые характеристики — недостаточными. Как отмечал в ходе дискуссии на

конференции 1983 г. [3] П. Швейцер, „мы дошли до конца дороги с точными моделями...Мы затратили слишком много времени на такие модели, как мультипликативные сети и матрично-геометрические решения“. Альтернативой является только *потокэквивалентная декомпозиция* сетей обслуживания.

В настоящей статье ограничимся рассмотрением разомкнутых однородных сетей с простейшими потоками. Последнее предположение базируется на известных теоремах о суммировании и случайном прореживании потоков. Интенсивности потоков определяются из уравнений баланса заявок:

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

где  $\Lambda$  — суммарная интенсивность потока, поступающего из внешних источников.

Далее для всех узлов должно быть проверено условие отсутствия перегрузки  $\lambda_i b_{i,1} / n_i < 1$ , обеспечивающее существование в сети стационарного режима.

С другой стороны, традиционное допущение о показательных распределениях времени обслуживания, как правило, является необоснованным, и порождаемые им ошибки могут быть сколь угодно велики. Это определяет целесообразность моделирования узлов сети системами с простейшим входящим потоком и произвольным распределением времени обслуживания. Последнее приходится аппроксимировать параллельно-последовательным набором фаз с экспоненциально распределенной задержкой в каждой. Приемлемую точность (выравнивание трех заданных моментов) обеспечивает, например, гиперэкспоненциальная аппроксимация с двумя составляющими. Заметим, что возможные случаи комплексных параметров и „парадоксальных“ вероятностей (одна отрицательна, а вторая больше единицы) не влияют на осмысленность конечных результатов. После такой аппроксимации расчет распределения числа заявок в узле можно выполнить итерационным методом Такахаси — Таками или методом матрично-геометрической прогрессии [4].

Для *разомкнутой* сети в целом среднее время пребывания заявки можно, как и для отдельного узла, вычислить на основе формулы Литтла

$$v = \sum_{i=1}^M k_i / \Lambda$$

(среднее число заявок в сети  $k_i$  делится на суммарную интенсивность входящего потока), которая проверена многократно (в том числе, автором данной статьи).

**Оптимизация сети.** Эта задача в первом приближении решается как минимизация среднего времени пребывания заявки в сети и имеет множество аспектов: в частности, выбор производительности и количества обслуживающих устройств в узлах (см., например, [5]). Реально производительность таких устройств должна выбираться из *конечного ряда* значений, а количество устройств должно быть *целочисленным*. Поэтому постановка задачи о комплексной оптимизации сети *формальными* методами представляется непродуктивной. Иначе обстоит дело с маршрутной матрицей, оптимизация которой вообще не затрагивает аппаратную часть сети обслуживания и материализуется „бесплатно“. Именно с такой оптимизации и следует начинать. Кроме того, необходимо учитывать, что на маршрутную матрицу могут быть наложены ограничения, диктуемые технологическими и/или организационными соображениями.

Формула Литтла позволяет считать, что минимизация ожидаемого количества заявок в сети одновременно минимизирует среднее время пребывания в ней заявки. Естественно проводить оптимизацию маршрутной матрицы передач путем последовательной „расшивки“ узких мест сети, что достигается выравниванием ожидаемого числа заявок в узлах или

коэффициентов загрузки последних. Опишем алгоритм выравнивания (ради экономии места без разбивки на абзацы).

Рассчитать потоки на входе узлов; коэффициенты загрузки узлов; моменты распределения времени пребывания в узлах для однократного захода заявки; среднее количество  $q_i$  заявок в каждом узле сети и среднее время пребывания заявки в сети

$T = \sum_{i=1}^M q_i / \Lambda$ . Выбрать в качестве объекта разгрузки узел  $j^-$  с максимальным значением

$q = q_{j^-}$ . Выбрать непосредственного „предшественника“ этого узла  $j^*$  минимум с двумя приемниками. Среди его приемников выбрать узел  $j^+$  с наименьшим  $q = q_{j^+}$ . Долю  $x$  потока интенсивностью  $\lambda_{j^*r_{j^*,j^-}}$  переадресовать в узел  $j^+$  посредством коррекции маршрутной матрицы.

Указанные действия выполняются, пока остаются значимыми уменьшения среднего времени пребывания заявки в сети. Коррекция всегда производится для ненулевых элементов матрицы передач, что позволяет запрещать недопустимые ребра маршрутов. Максимально допустимое значение  $x$  определяется из условия

$$\rho'_{j^+} + x\lambda_{j^*r_{j^*,j^-}}b_{j^+} / n_{j^+} \leq 0,95,$$

где  $\rho'_{j^+}$  — исходный коэффициент загрузки.

Ниже рассматриваются два субоптимальных алгоритма оптимизации маршрутной матрицы.

**Выравнивание числа заявок** предполагает, что оптимальное значение  $x$  выбирается как абсцисса минимума параболы, аппроксимирующей зависимость от  $x$  суммарного числа заявок в узлах  $j^-$  и  $j^+$ . Парабола  $Ax^2 + Bx + C$  строится по точкам для  $x_0 = 0$ ,  $x_1 = x_{\max} / 2$  и  $x_2 = x_{\max}$ , причем слагаемые  $x_0 = 0$  определяются по результатам первого этапа текущей итерации. Координата минимума параболы

$$x^* = \frac{x_1^2(y_2 - y_0) - x_2^2(y_1 - y_0)}{2[x_1(y_2 - y_0) - x_2(y_1 - y_0)]}.$$

В примере сети с шестью рабочими узлами при исходной маршрутной матрице стартовое среднее время пребывания заявки в сети составляет 16,247, после первой итерации — 7,869, после второй — 6,821. Далее происходят осцилляции в диапазоне [6,80, 7,02].

**Выравнивание коэффициентов загрузки узлов.** Приравнивая правые части выражений для новых коэффициентов загрузки, приходим к условию

$$x = \frac{\rho_{j^-} - \rho_{j^+}}{\lambda_{j^*r_{j^*,j^-}}(b_{j^-} / n_{j^-} + b_{j^+} / n_{j^+})}.$$

Реализация этого подхода обеспечивает монотонное уменьшение целевого показателя, которое по шагам составляет 7,84; 1,27; 0,183;  $3,86 \cdot 10^{-2}$ ;  $1,02 \cdot 10^{-2}$ ;  $2,94 \cdot 10^{-3}$ ;  $2,58 \cdot 10^{-4}$ ;  $7,71 \cdot 10^{-5}$ ;  $6,87 \cdot 10^{-6}$ . Последний результат составил 6,896.

**Обсуждение результатов.** Из сопоставления результатов следует:

— оба рассмотренных метода работоспособны, несложны, быстро (в примере — за три шага) приводят к практически приемлемому результату и обеспечивают значительное уменьшение среднего времени пребывания заявки в сети;

— минимизация суммарного числа заявок в сети после некоторого числа монотонных улучшений порождает осциллирующий процесс;

— выравнивание коэффициентов загрузки монотонно приводит практически к тому же результату (разница в третьем знаке) и в каждой итерации исключает необходимое для параболической аппроксимации дополнительное двукратное обращение к процедуре расчета модели  $M/H_2/n$  для двух узлов сети.

Таким образом, для реальных расчетов предпочтительно применять выравнивание коэффициентов загрузки узлов.

Предложенный подход можно обобщить и на неоднородный поток заявок.

**Заключение.** Алгоритм коррекции маршрутных матриц прост, эффективен и позволяет легко учесть ограничения на допустимость коррекции их отдельных элементов. Если работа с маршрутной матрицей не дает приемлемых результатов, следует использовать эту технологию в комбинации с последовательным повышением производительности наиболее загруженных узлов — увеличением числа каналов или их быстродействия. Такую оптимизацию следует вести в диалоговом режиме.

#### СПИСОК ЛИТЕРАТУРЫ

1. *Ивницкий В. А.* Теория сетей массового обслуживания. М.: Физматлит, 2004.
2. *Baskett F., Chandy K. M., Muntz R. R., Palacios J. G.* Open, closed, and mixed networks of queuing with different classes of customers // *J. of the ACM*. 1975. Vol. 22, N 2. P. 248—260.
3. *Mathematic computer performance and reliability // Proc. of the Intern. Workshop, Piza, 1983.* Amsterdam: North-Holland Publ. Co, 1984. 429 p.
4. *Рыжиков Ю. И.* Алгоритмический подход к задачам массового обслуживания: Монография. СПб: ВКА им. А. Ф. Можайского, 2013. 496 с.
5. *Янбых Г. Ф., Столяров Б. А.* Оптимизация информационно-вычислительных сетей. М.: Радио и связь, 1987. 232 с.

#### Сведения об авторе

**Юрий Иванович Рыжиков**

— д-р техн. наук, профессор; СПИИРАН, лаборатория информационных технологий в системном анализе и моделировании; Военно-космическая академия им. А. Ф. Можайского, кафедра математического обеспечения ЭВМ, Санкт-Петербург; профессор; E-mail: ryzhbox@yandex.ru

Рекомендована СПИИРАН

Поступила в редакцию  
10.06.14 г.