

С. В. КОКОРИН, Ю. И. РЫЖИКОВ

ОПТИМИЗАЦИЯ ПАРАМЕТРОВ СЕТЕЙ МАССОВОГО ОБСЛУЖИВАНИЯ НА ОСНОВЕ КОМБИНИРОВАННОГО ИСПОЛЬЗОВАНИЯ АНАЛИТИЧЕСКИХ И ИМИТАЦИОННЫХ МОДЕЛЕЙ

Рассматриваются методы оптимизации параметров сетей массового обслуживания, задаваемых с использованием как аналитических, так и имитационных моделей.

Ключевые слова: имитационная модель, аналитическая модель, теория очередей, численная оптимизация с ограничениями, однородные системы и сети массового обслуживания.

Введение. Задача поиска оптимальных вероятностно-временных характеристик современных информационных систем (ИС) является весьма актуальной. ИС представляют собой сложные объекты, подверженные структурной динамике, поэтому для их моделирования необходимо создание совокупности моделей с различной степенью детализации.

Исследования показывают, что вопрос об оптимальном выборе характеристик ИС остается, как правило, за рамками постановки и решения классических задач анализа систем и сетей массового обслуживания (СМО) либо рассматривается применительно к простым моделям теории очередей, для которых существует аналитическое решение задач оптимизации, т.е. при экспоненциальном распределении времени обслуживания и интервалов между поступлениями заявок [1].

В настоящей статье предлагается использовать аналитические и имитационные модели (АМ, ИМ) теории очередей для ИС, описывающих рассматриваемую предметную область. По сравнению с существующими подходами исследуем более широкий класс моделей теории

очередей, допускающих произвольное распределение времени обслуживания в узлах сети в рамках задачи поиска ее оптимальных параметров.

Постановка задачи. Рассмотрим открытую однородную СМО, которая задается набором параметров $(\mathbf{a}, \mathbf{n}, \{\mathbf{b}_i\}_{i=1}^M, R)$, где $\mathbf{a} = (a_1, \dots, a_m)^T$ — первые m начальных моментов распределения интервалов между появлением заявок в источнике (выбираются в зависимости от требуемой точности решения задачи); $\mathbf{n} = (n_1, \dots, n_M)^T$ — количество каналов в каждом из M узлов сети; $\mathbf{b}_i = (b_{i1}, \dots, b_{im})^T$ — первые m начальных моментов распределения времени обслуживания заявок, задаваемых независимо для каждого узла сети; $R = \|r_{ij}\|_{i,j=0}^{i,j=M+1}$ — матрица стационарных вероятностей переходов между узлами сети. При рассмотрении СМО под нулевым узлом будем понимать источник заявок, а под узлом с индексом $M+1$ — сток сети. Два данных узла не являются частью сети и вводятся лишь для удобства обозначений. Представляющими интерес параметрами сети являются: $\mathbf{v} = (v_1, \dots, v_m)^T$ — моменты распределения времени пребывания заявок в сети и $\boldsymbol{\rho} = (\rho_1, \dots, \rho_M)^T$ — коэффициент загрузки для каждого узла сети.

Исследуем влияние структуры однородной сети на ее пропускную способность. В данном случае оптимизируемыми параметрами являются элементы матрицы R . В качестве целевых функций (ЦФ) выберем две, позволяющие количественно оценить такие аспекты, как эффективность и отказоустойчивость исследуемой СМО соответственно:

$$J_1 = v_1(R) \rightarrow \min_{R \in \Omega},$$

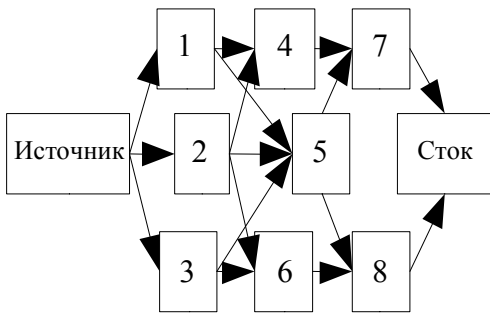
где Ω — множество допустимых значений элементов матрицы R ;

$$J_2 = \rho(R) = 1/M \sum_{i=1}^M \rho_i^2(R) - 1/M^2 \left(\sum_{i=1}^M \rho_i(R) \right)^2 \rightarrow \min_{R \in \Omega}.$$

ЦФ J_1 можно интерпретировать как среднее время прохождения заявки через СМО, а ЦФ J_2 — как показатель равномерности распределения нагрузки между узлами сети.

Множество Ω содержит естественные ограничения: $\sum_{j=0}^{M+1} r_{ij} = 1 \forall i \in \{0, \dots, M+1\}$ [2].

В общем случае предполагается, что далеко не все узлы СМО связаны друг с другом, поэтому для многих элементов матрицы R существует ограничение вида $r_{ij} = 0$. Предполагается так-



же, что имеется возможность выбирать не все элементы матрицы R . Модель сети с подобными ограничениями приведена на рисунке, где элементы матрицы R , которые могут принимать ненулевое значение, обозначены стрелками.

В дополнение к сказанному необходимо отметить, что можно задавать ограничения на элементы матрицы R следующим образом:

$$f'(R) = I(R \in \Omega') f(R),$$

где $f(R)$ — некоторая целевая функция, $I(R)$ — индикатор принадлежности матрицы R множеству Ω' , которое задается произвольным образом и является подмножеством Ω .

Кроме того, примем, что:

— $a \leq r_{ij} \leq A$, где (a, A) — произвольные константы; при этом проверка согласованности ограничений для различных r_{ij} не проводится; в случае когда данные ограничения не выполняются, то с использованием предлагаемого в статье метода решение не может быть найдено;

— $b \leq \rho_i(R) \leq B$, где (b, B) — произвольные константы.

Введение перечисленных ограничений позволяет учесть факторы износа оборудования, а также специфику различных транспортных путей между узлами, что в целом обеспечивает возможность моделирования реальных ИС с учетом особенностей их функционирования.

Расчет характеристик сети. Известно, что в общем случае [2] численный расчет параметров открытой СМО с учетом преобразования потоков заявок включает в себя следующие шаги.

Шаг 1. Аппроксимация входных данных по методу моментов с помощью экспоненциального либо гиперэкспоненциального распределения в зависимости от особенностей модели.

Шаг 2. Задание начальных значений λ_i для потоков поступающих заявок. Данные потоки принимаются простейшими с интенсивностью, определяемой при решении системы уравнений, описывающих условие баланса потоков:

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

где Λ — суммарная интенсивность потока заявок, поступающих от внешних источников.

Шаг 3. Проверка отсутствия перегрузки во всех узла сети: $\lambda_i b_{i1} / n(i) < 1$. Коэффициенты немарковости [2] входящих потоков для каждого узла на первой итерации полагаются равными $\xi_i^{(0)} = 0, i = 1, \dots, M$.

Шаг 4. Решение поочередно для всех узлов следующих подзадач:

— оценка распределений временных интервалов между поступлением заявок в i -й узел со всех узлов сети;

— суммирование полученных оценок распределений;

— расчет новых коэффициентов немарковости $\left\{ \xi_i^{(k)} \right\}_{i=1}^M$, определение невязки

$$\Delta_i^{(k)} = \left| \xi_i^{(k)} - \xi_i^{(k-1)} \right| \text{ для текущей итерации } k;$$

— расчет узла как изолированной системы массового обслуживания;

— расчет оценки распределений интервалов потока заявок, выходящего из i -го узла;

Шаг 5. Если $\max_{i \in \{1, \dots, M\}} \Delta_i > \varepsilon$, где ε — заданный порог невязки, возвращение к шагу 4.

Шаг 6. Расчет моментов распределения времени пребывания заявки в узлах при каждом посещении.

Шаг 7. Расчет моментов распределения времени пребывания заявки в сети в целом.

Оценки характеристик рассматриваемого класса СМО также можно получить с помощью имитационных моделей, однако в этом случае вычислительные затраты, связанные с получением окончательных решений, значительно возрастают по сравнению с подходом, использующим аналитические модели.

Методы оптимизации параметров СМО. Для вычисления и оптимизации целевых функций J_1, J_2 предлагается использовать метод глобального поиска — *метод пси-*

преобразования [3] — и метод численной оптимизации без вычисления производных — метод главных осей Брента [4].

Метод пси-преобразования является методом поиска глобального экстремума целевой функции и не критичен к выбору начальной оценки. Однако для его реализации требуются значительные вычислительные ресурсы в том случае, если увеличивается размерность пространства оптимизируемых параметров. Для ЦФ J_1 в качестве пси-функции была выбрана вероятностная мера множества точек, на котором пропускная способность сети выше заданного уровня, а для ЦФ J_2 — вероятностная мера множества точек, на котором величина $\rho(R)$ меньше заданного положительного значения. В этом случае задача оптимизации параметров СМО сводится к поиску решения уравнения с многими переменными (оптимизируемыми параметрами).

Рассмотрим алгоритм, реализующий метод пси-преобразования [3].

Шаг 1. Оценка разброса значений целевой функции методом случайных испытаний.

Шаг 2. Выбор множества значений уровней $v_1(R) \geq \zeta_l$, $l \in \{1, \dots, L\}$, где L — заданное количество уровней, а ζ_l — величины, определяемые по оценкам разброса целевой функции, полученным на шаге 1.

Шаг 3. Вычисление средних значений целевой функции для каждого уровня методом случайных испытаний:

$$\Psi_l = 1/q \sum_{\{R^{(s)}: v_1(R^{(s)}) \geq \zeta_l\}} (v_1(R^{(s)}) - \zeta_l),$$

где q — число испытаний, $R^{(s)}$ — матрица переходов при s -м испытании и $l \in \{1, \dots, L\}$.

Шаг 4. Вычисление средних значений параметров элементов матрицы R для каждого уровня:

$$x_{ijl} = q / \Psi_l \sum_{\{R^{(s)}: v_1(R^{(s)}) \geq \zeta_l\}} r_{ij}^{(s)} (v_1(R^{(s)}) - \zeta_l)^\alpha,$$

где $\alpha = [1; \infty]$ — параметр метода оптимизации.

Шаг 5. Параболическая аппроксимация и экстраполяция величин Ψ_l на уровень 0.

Вследствие высокой вычислительной сложности метода пси-преобразования и его недостаточной точности при оптимизации ЦФ J_1 и J_2 при большой размерности СМО предлагается использовать **метод главных осей Брента**. Данный метод ориентирован на локальную оптимизацию функции многих переменных без вычисления производных. На практике метод эффективен при решении задач оптимизации структуры СМО, в том числе, и для случая, когда пространство параметров имеет большую размерность. Основным недостатком алгоритма, реализующего данный метод, является необходимость задания начальной оценки, которая должна быть вычислена отдельно для каждой конкретной задачи. Кроме того, необходимо задать два параметра — δ и Δz : параметр δ определяет момент останова итерационного процесса, Δz — максимальную длину шага алгоритма.

Приведем шаги алгоритма, реализующего метод главных осей для вычисления целевых функций J_1 и J_2 .

Шаг 1. Вычисление начальной оценки матрицы $R^{(0)}$.

Шаг 2. Задание начальных направлений одномерного поиска: $U^0 = \left\| (u_i)^{(0)} \right\|_{i=1}^{M(M+2)} = I$,

где I — единичная матрица.

Шаг 3. Поиск оптимального значения ЦФ поочередно вдоль каждого направления методом золотого сечения с начальным интервалом $2\Delta z$. Перенос текущих значений оптимальных параметров во вновь найденную точку, которую обозначим как $R^{(k)}$.

Шаг 4. Вычисление величины $\Delta z = 2 \cdot \|R^{(k)} - R^{(k-1)}\|$.

Шаг 5. Исключение вектора $u_{M(M+2)}^{(k-1)}$, переход к матрице направлений следующего вида: $U^k = (R^{(k)} - R^{(k-1)}, u_1^{(k-1)}, u_2^{(k-1)}, \dots, u_{M(M+2)-1}^{(k-1)})$.

Шаг 6. После полной смены направляющего набора векторов $U^{M(M+2)}$ обновление набора полностью с замещением его ортогональной матрицей, аппроксимирующей гессиан целевой функции в текущей точке.

Шаг 7. Повтор шагов 3—6 до тех пор, пока абсолютная разность значений ЦФ на последовательных итерациях не станет меньше δ .

Выбор начальной оценки при реализации метода главных осей предлагается осуществлять с использованием модели сети Джексона [4], для которой существует аналитическое решение поставленной задачи.

Многоканальные узлы замещаются одноканальными с пропорционально увеличенной интенсивностью времени обслуживания $\mu' = 1/b_i n(i)$. Экспоненциальное распределение времени обслуживания в каждом узле аппроксимируется с учетом первых моментов заданных распределений. Последовательный обход сети в ширину позволяет выбрать начальные оценки для каждого узла на основе решения следующей оптимизационной задачи:

$$\sum_{j=1}^{M+1} \frac{r_{ij}}{\mu'_j - r_{ij} \lambda_j} \rightarrow \min_{R \in \Omega}$$

Использование данного метода выбора начальных оценок эффективно (в вычислительном отношении) при небольшом числе каналов в узлах и малых значениях коэффициентов вариации времени обслуживания в узлах. В качестве альтернативы используется метод пси-преобразования. Его реализация предпочтительнее для сетей, в которых закон распределения времени обслуживания значительно отличается от экспоненциального (например, равномерный закон распределения на отрезке или гамма-распределение с высокой вариацией.).

Сравнительный анализ начальных оценок матрицы R , полученных методом построения сети Джексона и методом пси-преобразования для расчета оптимального значения ЦФ J_1 , представлен в таблице, где K — количество итераций.

M	Метод Джексона		Метод пси-преобразования		J_1
	$v_1(R)$	K	$v_1(R)$	K	
3	2,91	319	2,91	651	2,18
5	3,35	895	3,57	1293	3,19
8	4,9	6067	4,83	5419	4,53
13	7,02	21939	7,92	41345	6,57
21	9,05	212939	8,84	201323	8,64

Численные эксперименты. Для сравнения результатов, полученных с использованием аналитических и имитационных моделей СМО, рассмотрим простую сеть, состоящую из 8 узлов, расположенных в трех ярусах (см. рисунок). При варьировании элементов, входящих в матрицу R вероятностей переходов между узлами, можно получить одинаковые результаты как на АМ, так и на ИМ. По сравнению с численной оптимизацией на АМ процедура поиска параметров на ИМ отличается плохой сходимостью соответствующего локального метода, которая определяется тем, что при имитации целевые функции принимают случайные значения.

Для уменьшения влияния данного фактора требуется увеличить число имитационных экспериментов, при этом высокая точность требуется только на последних шагах метода при малых значениях Δz . Поэтому для повышения эффективности вычислений ЦФ на ИМ был реализован расчет с использованием метода Brenta, при котором точность расчета целевой функции зависит от длины шага как $q_0 / \Delta z$.

Выводы. В задачах оптимизации характеристик СМО вопросы повышения эффективности соответствующих вычислительных процедур, связанных с расчетом ЦФ, приобретают особую актуальность. При этом использование ИМ в данных задачах требует значительного количества информационных и временных ресурсов. Кроме того, такое моделирование характеризуется невысокой точностью расчета экстремальных значений ЦФ вследствие статистических погрешностей имитации. Поэтому в данной ситуации целесообразен поиск аналогичных параметров СМО с использованием АМ, что обусловлено значительно меньшими затратами информационных ресурсов, сокращением общего времени на проведение соответствующих расчетов ЦФ, а сами значения рассматриваемых функций вычисляются с меньшими погрешностями.

Предложенный в настоящей статье оригинальный подход к разработке и реализации методов оптимизации параметров различных типов сетей массового обслуживания основан на комбинированном использовании методов глобальной и локальной оптимизации, при этом недостатки метода, базирующегося на имитационной модели, компенсируются достоинствами аналитического метода.

Исследования по данной тематике проводились при поддержке Российского фонда фундаментальных исследований (гранты 10-07-00311, 09-07-00066а, 09-07-11004, 08-08-00346-а, 10-08-0906-а, 10-08-90027-Бел-а) и Отделения нанотехнологий и информационных технологий РАН (проект № О-2.3/03).

СПИСОК ЛИТЕРАТУРЫ

1. *Bramson M.* Stability of Queuing Networks [Электронный ресурс]: <<http://projecteuclid.org/DPubS?service=UI&version=1.0&verb=Display&handle=euclid.ps/1220879338>>.
2. *Рыжиков Ю. И.* Машинные методы расчета систем массового обслуживания. Л.: ВИКИ им А. Ф. Можайского, 1979.
3. *Чичинадзе В. К.* Решение невыпуклых нелинейных задач оптимизации. Метод пси-преобразования. М.: Наука, 1983.
4. *Brent R. P.* Algorithms for Minimization without Derivatives. NJ: Prentice-Hall, Inc., 1973.
5. *Охтилев М. Ю., Соколов Б. В., Юсупов Р. М.* Интеллектуальные информационные технологии мониторинга и управления структурной динамикой сложных технических объектов. М.: Наука, 2006.

Сведения об авторах

- Сергей Владимирович Кокорин** — аспирант; СПИИРАН, лаборатория информационных технологий в системном анализе и моделировании; E-mail: kokorins@rambler.ru
- Юрий Иванович Рыжиков** — д-р техн. наук, профессор; СПИИРАН, лаборатория информационных технологий в системном анализе и моделировании; E-mail: ryzhbox@yandex.ru

Рекомендована СПИИРАН

Поступила в редакцию
09.07.10 г.