

УДК 519.2:004.421.5:004.7

ПОЛУАНАЛИТИЧЕСКИЕ МЕТОДЫ ОПТИМИЗАЦИИ ТРАНСПОРТНЫХ СЕТЕЙ

В. Н. Задорожный

Омский государственный технический университет, zwn2015@yandex.ru

Для немарковских сетей массового обслуживания предлагается эффективный полуаналитический метод комплексной оптимизации маршрутной матрицы и быстродействий узлов. В качестве целевой функции используется среднее время прохождения заявки через сеть. Метод сочетает быстрый точный расчет градиента целевой функции в пространстве варьируемых переходных вероятностей и ускоренный аналитико-имитационный расчет градиента целевой функции в пространстве варьируемых быстродействий узлов.

Разрабатываются методы комплексной оптимизации транспортных сетей, основанные на их аппроксимации немарковскими сетями массового обслуживания с многолинейными узлами-дорогами, быстродействие которых падает с ростом нагрузки. Предлагаемые методы ориентированы на различные режимы транспортной нагрузки. Приводятся примеры применения предлагаемых методов. Демонстрируется, что наибольший эффект в борьбе с заторами дает комплексная оптимизация транспортных сетей, ориентированная на режим высокой транспортной нагрузки.

Ключевые слова: немарковские сети массового обслуживания, градиентные методы оптимизации, транспортные сети, переменная интенсивность обслуживания.

SEMI-ANALYTICAL METHODS OF OPTIMIZATION OF TRANSPORT NETWORKS

V. N. Zadorozhnyi

Omsk State Technical University, zwn2015@yandex.ru

An effective semi-analytical method for the integrated optimization of the routing matrix and node speeds for non-Markov queuing networks is proposed. As the objective function, the average time spent on the application through the network is used. The method combines a fast accurate calculation of the objective function gradient in the space of variable transition probabilities and an accelerated analytic-simulation calculation of the objective function gradient in the space of variable nodes speeds.

Methods of integrated optimization of transport networks are developed. They are based on their approximation by non-Markov networks of queuing with multi-linear nodes (roads), the speed of which decreases with increasing load. The proposed methods are oriented to different modes of transport load. Examples of the application of the proposed methods are given. It is demonstrated that the greatest effect in combating blockings is provided by the integrated optimization of transport networks oriented to the regime of high transport load.

Keywords: non-Markov queueing networks, gradient methods of optimization, transport networks, variable service rate.

В общем случае для оптимизации немарковских сетей массового обслуживания (далее – СеМО) приходится использовать имитационное моделирование (далее – ИМ). При этом если число варьируемых параметров выходит за пределы десятка, то без применения градиентных методов оптимизация становится практически нереализуемой, а расчет градиентов в ИМ чрезвычайно затруднен наличием стохастических погрешностей у вычисляемых оценок целевой функции [1]. Тем не менее проблему вычисления градиентов в ИМ СеМО удается решать путем применения полуаналитических методов [1–5].

При оптимизации СеМО в качестве варьируемых параметров могут выступать быстродействие узлов [6], число каналов в узлах [7], объемы буферов [8], переходные

вероятности или совокупности параметров нескольких таких видов [1, 6, 9]. В реальных сетях общий ресурс, используемый для улучшения показателей заданного вида, как правило, ограничен, поэтому задачи оптимизации ставятся обычно как задачи поиска оптимального распределения ресурса заданного вида [10] или оптимального сочетания распределений ресурсов нескольких видов. Целевые функции часто включают такие показатели качества, как среднее время прохождения заявки через сеть (среднее время ответа), вероятность превышения допустимого времени ответа, объемы буферов, вероятность потери заявки в узлах сети, общее число каналов в узлах и т. д.

Разработка новых эффективных методов оптимизации немарковских СеМО актуальна в области проектирования компьютерных [10–12] и транспортных [13–17] сетей. Значительные затруднения при оптимизации СеМО возникают в случаях, когда варьируемые параметры включают переходные вероятности [1, 6], т. е. когда оптимизируются маршрутные матрицы. В данной статье разрабатываются методы, позволяющие существенно, по сравнению с непосредственным ИМ, ускорить комплексную оптимизацию маршрутных матриц и быстродействия узлов немарковских сетей, включая транспортные сети.

Совместная оптимизация переходных вероятностей и быстродействия узлов.

Рассмотрим задачу оптимизации открытых немарковских СеМО путем одновременного перераспределения быстродействия узлов сети и ее переходных вероятностей.

В сеть поступает рекуррентный поток заявок с интенсивностью Λ . Длительности интервалов поступления заявок описываются функцией распределения (далее – ф.р.) $A(t)$. Заявка входного потока сети с вероятностью p_{0i} попадает в i -й узел, $i = \overline{1, n}$. Независимое время обслуживания заявки в любом из K_i каналов i -го узла имеет ф.р. $B_i(t)$. После обслуживания в i -м узле заявка случайно, в соответствии с заданными переходными вероятностями p_{ij} , выбирает один из узлов $j = \overline{1, n}$ для продолжения своего маршрута или (с вероятностью p_{i0}) уходит из сети (рис. 1). Вероятности p_{ij} ($i, j = \overline{0, n}$) задаются неразложимой маршрутной матрицей $\mathbf{P} = \| p_{ij} \|$. Если потоков, входящих в сеть из внешней среды, больше одного, они объединяются в один поток, разветвляющийся с подходящими вероятностями на входы разных узлов.

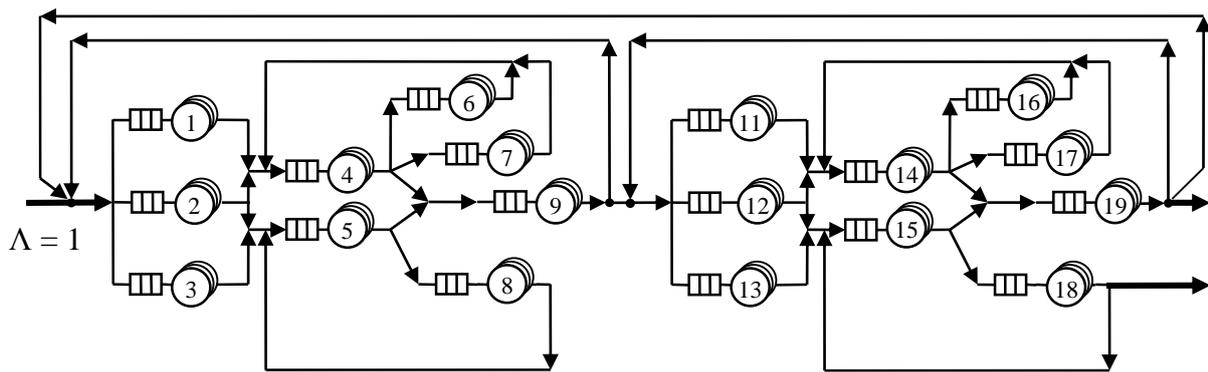


Рис. 1. Пример открытой сети с многолинейными узлами

Среднее время E прохождения открытой сети (среднее время ответа) можно выразить формулой:

$$E = \sum_{i=1}^n \alpha_i \left(w_i + \frac{1}{\mu_i} \right), \tag{1}$$

где α_i – среднее число (частота) посещения i -го узла заявкой за время прохождения сети;

w_i – среднее время ожидания в очереди i -го узла;

$\mu_i = b^{-1}$ – интенсивность обслуживания в канале i -го узла (соответственно b_i – среднее время обслуживания в i -м узле).

Частоты α_i однозначно определяются вероятностями p_{ij} посредством системы линейных алгебраических уравнений (далее – СЛАУ):

$$\alpha_i = \sum_{j=0}^n \alpha_j p_{ji}, \quad i = \overline{0, n}, \quad \alpha_0 \equiv 1. \quad (2)$$

Через решения α_i этой СЛАУ определяются интенсивности $\lambda_i = \Lambda \alpha_i$ потоков, входящих в узлы, и затем – коэффициенты загрузки узлов $\rho_i = \lambda_i / (\mu_i K_i)$, $i = \overline{1, n}$. Значения w_i для целевой функции (1) в общем случае определяются посредством моделирования и эффективно используются для оптимизации быстродействия узлов сети методом направляющих гипербол [1]. Метод направляющих гипербол является градиентным методом. Он использует адаптивную сепарабельную гиперболическую аппроксимацию зависимостей $w_i(\mu_i)$ и позволяет с высокой точностью решать задачу оптимального распределения быстродействия по узлам немарковской СеМО, состоящей из сотен одноканальных и многоканальных узлов.

Более общая задача совместной оптимизации переходных вероятностей и быстродействия узлов ставится в [6] следующим образом: ресурс M быстродействия (стоимость) сети как функция вектора $\vec{\mu} = (\mu_1, \dots, \mu_n)$ интенсивностей обслуживания в узлах $i = \overline{1, n}$ задается в виде $M(\vec{\mu}) = \sum_{i=1}^n c_i \mu_i^{\beta_i}$, где $c_i > 0$ – стоимостные коэффициенты, $\beta_i > 0$ – коэффициенты нелинейности. Варьируемые переходные вероятности, выписанные в фиксированном порядке, являются координатами вектора \vec{p}_v размерности m .

Требуется найти векторы $\vec{\mu} = \vec{\mu}_{\text{opt}}$ и $\vec{p}_v = \vec{p}_{v, \text{opt}}$, доставляющие минимум среднего времени ответа:

$$E(\vec{\mu}, \vec{p}_v) = \sum_{i=1}^n \alpha_i(\vec{p}_v) \left(w_i(\vec{\mu}, \vec{p}_v) + \frac{1}{\mu_i} \right) \rightarrow \min_{\vec{\mu}, \vec{p}_v} \quad (3)$$

и принадлежащие области допустимых решений:

$$M(\vec{\mu}) = \sum_{i=1}^n c_i \mu_i^{\beta_i} = M^* = \text{const}; \quad 0 \leq \rho_i \leq 1; \quad \sum_{j=0}^n p_{ij} = 1, \quad (i = \overline{0, n});$$

$$p_{vi}^{\min} \leq p_{vi} \leq p_{vi}^{\max}, \quad (i = \overline{1, m}). \quad (4)$$

Разработанный и испытанный в [6] метод решения задачи (3), (4) состоит в выполнении цикла итераций, приближающих варьируемые параметры к решению $(\vec{\mu}_{\text{opt}}, \vec{p}_{v, \text{opt}})$. Каждая итерация состоит из двух шагов. На первом шаге при фиксированном \vec{p}_v методом направляющих гипербол [1] оптимизируется распределение $\vec{\mu} = (\mu_1, \dots, \mu_n)$ ресурса M^* по узлам сети. На втором шаге при найденном $\vec{\mu}$ оптимизируется вектор \vec{p}_v переходных вероятностей. Далее итерации поочередной оптимизации векторов $\vec{\mu}$ и \vec{p}_v повторяются, пока не наступит заданное условие останова. На шаге оптимизации вектора \vec{p}_v для определения градиентов рассчитываются частные производные (далее – ЧП) $\partial E / \partial p_{jk}$ времени E по всем компонентам вектора \vec{p}_v , причем вычисление этих производных сводится к вычислению

производных $\partial\alpha_i(\bar{p}_v)/\partial p_{j,k}$. Действительно, дифференцируя (1) по любой переходной вероятности p_{jk} , получаем равенство:

$$\frac{\partial E}{\partial p_{jk}} = \sum_{i=1}^n \frac{\partial(\alpha_i/\mu_i)}{\partial p_{jk}} + \sum_{i=1}^n \frac{\partial(\alpha_i w_i)}{\partial p_{jk}} = \sum_{i=1}^n \frac{1}{\mu_i} \frac{\partial\alpha_i}{\partial p_{jk}} + \sum_{i=1}^n \frac{\partial\alpha_i}{\partial p_{jk}} w_i + \sum_{i=1}^n \alpha_i \frac{\partial w_i}{\partial p_{jk}},$$

из которого в [1] выводится приближение:

$$\frac{\partial E}{\partial p_{jk}} \approx \sum_{i=1}^n \left(\frac{1}{\mu_i} - \mu_i \frac{\partial w_i}{\partial \mu_i} \right) \frac{\partial\alpha_i}{\partial p_{jk}},$$

содержащее значения всех переменных и производных, известные (за исключением $\partial\alpha_i/\partial p_{jk}$) после шага оптимизации вектора $\bar{\mu}$ методом направляющих гипербол. В [6] все ЧП $\partial\alpha_i/\partial p_{jk}$ рассчитываются в ходе последовательных упрощений графа вложенного полумарковского процесса, сопровождаемых рекурсивным пересчетом ЧП от суперпозиций функций многих переменных. Ниже метод расчета ЧП $\partial\alpha_i/\partial p_{jk}$ существенно уточняется и ускоряется за счет сведения расчета ЧП $\partial\alpha_i/\partial p_{jk}$ к решению нескольких вспомогательных СЛАУ.

Расчет ЧП частот α_i по переходным вероятностям. Поскольку задача расчета ЧП $\partial\alpha_i/\partial p_{jk}$ представляет собой задачу расчета ЧП решений α_i СЛАУ (2) по коэффициентам p_{ij} этой СЛАУ, достаточно показать, как рассчитываются ЧП решений \bar{x} произвольной СЛАУ $C\bar{x} = \bar{e}$ по ее коэффициентам C, \bar{e} .

Пусть дана имеющая единственное решение СЛАУ:

$$\begin{cases} c_{11}x_1 + c_{12}x_2 + \dots + c_{1n}x_n = e_1, \\ c_{21}x_1 + c_{22}x_2 + \dots + c_{2n}x_n = e_2, \\ \vdots \\ c_{n1}x_1 + c_{n2}x_2 + \dots + c_{nn}x_n = e_n, \end{cases} \quad (5)$$

и требуется вычислить ЧП $\partial x_i/\partial c_{jk}$ всех решений x_i этой системы по выбранному ее коэффициенту c_{jk} . Для точного расчета искомых ЧП продифференцируем уравнения этой СЛАУ по выбранному коэффициенту c_{jk} . В результате получаем:

$$\begin{cases} c_{11} \frac{\partial x_1}{\partial c_{jk}} + \dots + c_{1k} \frac{\partial x_k}{\partial c_{jk}} + \dots + c_{1n} \frac{\partial x_n}{\partial c_{jk}} = 0, \\ \dots \\ c_{j1} \frac{\partial x_1}{\partial c_{jk}} + \dots + c_{jk} \frac{\partial x_k}{\partial c_{jk}} + \dots + c_{jn} \frac{\partial x_n}{\partial c_{jk}} = -x_k, \\ \dots \\ c_{n1} \frac{\partial x_1}{\partial c_{jk}} + \dots + c_{nk} \frac{\partial x_k}{\partial c_{jk}} + \dots + c_{nn} \frac{\partial x_n}{\partial c_{jk}} = 0. \end{cases} \quad (6)$$

Все n ЧП $\partial x_i/\partial c_{jk}$ теперь являются неизвестными в СЛАУ (6) и, следовательно, могут быть найдены как ее решения. Матрица C коэффициентов СЛАУ (6) совпадает с матрицей C исходной СЛАУ (5). Столбец правых частей содержит нули везде, кроме строки j , в которую записывается значение $(-x_k)$. Таким образом, нами получен простой двухшаговый метод для расчета ЧП всех решений x_i СЛАУ (5) по ее выбранному коэффициенту c_{jk} :

- 1) решаем исходную СЛАУ (5);

2) решаем вспомогательную СЛАУ (6), получаемую из исходной СЛАУ (5), заменой нулями всех элементов столбца \vec{e} , кроме элемента e_j – он заменяется значением $(-x_k)$, известным после выполнения шага (1).

Здесь j – номер строки с ненулевым свободным членом, задается первым индексом выбранного коэффициента c_{jk} . Номер k решения x_k , определяющего ненулевой свободный член, задается вторым индексом коэффициента c_{jk} .

Решение $(\dot{x}_1, \dots, \dot{x}_n)$ вспомогательной СЛАУ (6), получаемое на шаге (2), есть искомым вектор $\left(\frac{\partial x_1}{\partial c_{jk}}, \dots, \frac{\partial x_n}{\partial c_{jk}} \right)$.

При расчете ЧП решений исходной СЛАУ (5) по выбранному коэффициенту e_j вспомогательная СЛАУ отличается от СЛАУ (6) только тем, что $c_j = 1$.

Как показывают численные эксперименты, метод вспомогательных СЛАУ позволяет рассчитывать соответствующие ЧП с высокой точностью и большой скоростью. Например, в Excel ЧП рассчитываются этим методом с 15-ю точными значащими десятичными цифрами. Решение рассчитывается по сравнению с шагами ИМ практически мгновенно. Так, решение СЛАУ из 100 уравнений в Excel (методом обращения матрицы **C**) занимает вместе с отображением результата долю секунды.

Задача оптимизации ПБУ-сетей и ее прикладное значение. Рассмотрим представленную на рис. 2 часть городской улично-дорожной сети (далее – УДС). Дороги 9–5, 9–6 и 9–4 этой сети однонаправленные (от узла 9), остальные дороги двунаправленные. Дороги во всех направлениях двухполосные.

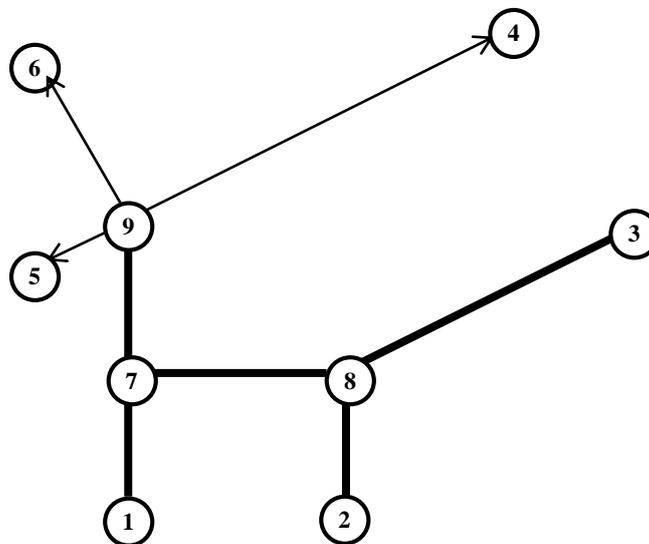


Рис. 2. Пример моделируемого участка УДС

Проиллюстрируем на примере этой УДС возможность комплексной оптимизации (по быстродействию дорог и по переходным вероятностям) транспортных сетей методами теории очередей. Для этого условимся представлять дороги УДС многолинейными системами массового обслуживания. Число K -каналов каждой многолинейной системы положим равным емкости соответствующей дороги – максимальному числу автомобилей среднего размера, которые могут двигаться по дороге одновременно, т. е. могут «обслуживаться» ею параллельно во времени. Маршрутная матрица **P** полученной таким образом сети с многолинейными узлами-дорогами представлена в табл. 1. На практике переходные вероятности подобных маршрутных матриц можно определять путем наблюдения исследуемых УДС.

Специфика дорог как обслуживающих узлов учитывается в теории транспортного моделирования путем использования для каждой дороги индивидуальной характеристики, называемой фундаментальной диаграммой (далее – ФД), которая отражает снижение средней скорости автотранспорта на дороге при увеличении плотности движения. Сети с такими обслуживающими узлами далее называются ПБУ-сетями (сетями с подавляемым быстроедействием узлов). Разрабатывается и исследуется ускоренный метод комплексной оптимизации транспортных сетей, использующий их представление ПБУ-сетями.

Имеется несколько общепринятых форм представления ФД. Определим ФД как функцию $\delta(\rho)$ коэффициента ρ загрузки дороги:

$$\delta(\rho) = \frac{v(\rho)}{v_0}, \quad (7)$$

где $\rho = \lambda b / K$, λ – интенсивность входа автомобилей на дорогу, $b = b(\rho) = l / v(\rho)$ – среднее время проезда по дороге, l – длина дороги ($0 \leq \rho \leq 1$);

$v(\rho)$ – средняя скорость на дороге при коэффициенте загрузки ρ ;

$v_0 = v(0)$ – базовая средняя скорость, т. е. скорость, с которой автомобили движутся по пустой дороге.

Таблица 1

Переходные вероятности участка УДС

Дорога i	Дорога j													
	0	7-1	2-8	9-7	7-8	3-8	1-7	7-9	8-7	8-2	8-3	9-5	9-6	9-4
0			0,3			0,3	0,4							
7-1							1							
2-8									0,5	0,2	0,3			
9-7		0,3			0,2			0,5						
7-8									0,5	0,2	0,3			
3-8									0,5	0,2	0,3			
1-7		0,3			0,2			0,5						
7-9				0,2								0,3	0,2	0,3
8-7		0,3			0,2			0,5						
8-2			1											
8-3						1								
9-5	1													
9-6	1													
9-4	1													

Емкости узлов-дорог в сети достаточно велики, поэтому в случае значительной загрузки дороги случайное число автомобилей в соответствующем многолинейном узле имеет распределение, близкое к распределению Гаусса [12]. Это позволяет приближенно рассчитывать стационарные режимы и оптимизировать их характеристики, ориентируясь на среднее число машин на дорогах и на их средние скорости. Ориентация на такое приближение оправдывает использование соотношения $b(\rho) = l / v(\rho)$ (из наблюдений за плотными транспортными потоками видно, что скорости автомобилей в плотном потоке практически одинаковы). В общем случае соотношение $b(\rho) = l / v(\rho)$ несостоятельно, так как в нем средние b и v рассматриваются как средние по множеству автомобилей, движущихся с разными скоростями. Впрочем, в общем случае этот вопрос можно снять, определяя и измеряя ФД сразу как зависимость $b(\rho)$ среднего времени прохождения дороги от нагрузки.

Как правило, ФД $\delta(\rho)$ (7) – монотонно невозрастающие функции, поэтому, согласно (7), $0 \leq \delta(\rho) \leq 1$ и при известной ФД фактическая средняя скорость $v(\rho)$ на дороге определится формулой $v(\rho) = v_0 \delta(\rho)$.

Для целей данной статьи можно использовать упрощенную ФД – одну и ту же для всех многолинейных узлов сети. Для сети, изображенной на рис. 2, выберем, например, ФД $\delta(\rho) = \frac{1}{1+\rho^2}$, изображенную верхней линией на рис. 3.

Основные параметры дорог приведены в табл. 2. В последней строке указаны базовые средние скорости v_{0i} , низкие значения которых объясняются плохим качеством дорог. На ремонт дорог выделен финансовый ресурс $\Phi^* = \$ 63\ 820$. Задача оптимизации сети заключается в нахождении такого распределения выделенного ресурса по дорогам, при котором среднее время E прохождения сети автомобилем будет минимально. При этом необходимо учесть, что после ремонта дорог водители личного транспорта могут менять свои прежние маршруты прохождения сети на новые, ставшие более выгодными.

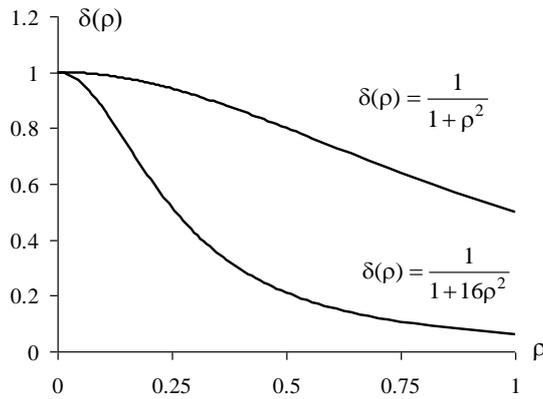


Рис. 3. Примеры фундаментальных диаграмм

Транспортные службы также могут влиять на маршруты, в том числе на маршруты движения общественного транспорта. Такая оптимизация маршрутов накладывается на оптимизацию распределения финансового ресурса Φ^* после его реализации. Поэтому распределение ресурса Φ^* разумно оптимизировать одновременно с маршрутной матрицей \mathbf{P} . Возникает следующая задача оптимизации, в общих чертах похожая на задачу (3), (4). Требуется найти вектор добавляемых за счет ремонта скоростей $\vec{\gamma} = \vec{\gamma}_{opt}$ и маршрутную матрицу $\mathbf{P} = \mathbf{P}_{opt}$, доставляющие при заданной Λ минимум среднего времени ответа:

$$E(\vec{\gamma}, \mathbf{P}) = \sum_{i=1}^n \alpha_i(\mathbf{P}) \frac{l_i}{(v_{0i} + \gamma_i)\delta(\rho_i)} \rightarrow \min_{\vec{\gamma}, \mathbf{P}} \quad (8)$$

и принадлежащие области допустимых решений:

$$\Phi(\vec{\gamma}) = \sum_{i=1}^n \varphi_i(\gamma_i) = \sum_{i=1}^n c_i l_i \gamma_i^{\beta_i} = \Phi^*; \quad 0 \leq \rho_i \leq 1; \quad \sum_{j=0}^n p_{ij} = 1, \quad (i = \overline{0, n});$$

$$p_{ij}^{\min} \leq p_{ij} \leq p_{ij}^{\max}, \quad (i, j = \overline{1, n}). \quad (9)$$

В (8) $\rho_i = \lambda_i b_i / K_i = \alpha_i \Lambda b_i / K_i = \alpha_i(\mathbf{P}) \Lambda l_i / (v_{0i} + \gamma_i) / K_i$. В (9) все коэффициенты нелинейности β_i примем равными единице. Стоимостные коэффициенты c_i указаны в табл. 2 в единицах измерения $[\text{доллар} / (\text{км}/\text{час})] / \text{км} = \text{доллар} \times \text{час} / \text{км}^2$.

Исходные параметры дорог

Параметр	Дорога <i>i</i>												
	7-1	2-8	9-7	7-8	3-8	1-7	7-9	8-7	8-2	8-3	9-5	9-6	9-4
Длина, м	234	242	257	234	305	176	242	257	305	484	207	072	176
Емкость	47	48	51	47	61	35	48	51	61	97	41	14	35
c_i	399,42	414,04	440,6	1398	523,3	305,7	1412	1437	524,3	1824	1352	1123	1300
α_i	0,75	0,74	0,25	0,5	0,96	1,15	1,25	1,1	0,44	0,66	0,375	0,25	0,375
v_{0i} , км/ч	10	10	10	10	10	10	10	10	10	10	10	10	10

В (9) в ограничениях $p_{ij}^{\min} \leq p_{ij} \leq p_{ij}^{\max}$ на изменения маршрутной матрицы положим, что все p_{ij}^{\min} равны $p_{ij}^0 - 0,1$ и все p_{ij}^{\max} равны $p_{ij}^0 + 0,1$, где p_{ij}^0 – исходные значения переходных вероятностей, приведенные в табл. 1. При этом вероятности, заданные нулями и единицами, не варьируются. В практических задачах ограничения на изменение переходных вероятностей более естественно задавать в виде ограничений на частоты α_i [9]. Как показывают численные эксперименты, такой способ задания ограничений не влияет на эффективность предлагаемого подхода.

Оптимизация транспортных ПБУ-сетей в режиме низкой нагрузки. На практике распределение ресурса Φ^* (8) по многолинейным узлам ПБУ-сетей (по дорогам транспортных сетей) можно оптимизировать для режима низкой нагрузки $\rho \rightarrow 0$, т. е. для режима, когда автомобили движутся по пустым дорогам, не учитывая при этом возможность модификации маршрутной матрицы, если поводы для такой модификации оцениваются как незначительные. Задача оптимизации при этом значительно упрощается и решается аналитическими методами. Основной вопрос к правомерности такого подхода заключается в том, как изменится целевая функция E оптимизированной таким образом сети, когда нагрузка дороги будет значительной, например, в часы пик перед выходными днями.

Решим эту упрощенную задачу оптимизации и оценим качество упрощенного решения на примере ПБУ-сети, построенной в предыдущем разделе статьи. Задача (8), (9) в результате упрощения ставится следующим образом. Требуется найти вектор скоростей $\vec{\gamma} = \vec{\gamma}_{opt}$, добавляемых за счет ремонта дорог, доставляющий минимум среднего времени ответа:

$$E(\vec{\gamma}) = \sum_{i=1}^n \alpha_i \frac{l_i}{v_{0i} + \gamma_i} \rightarrow \min_{\vec{\gamma}} \quad (10)$$

и принадлежащий области допустимых решений:

$$\Phi(\vec{\gamma}) = \sum_{i=1}^n \varphi_i(\gamma_i) = \sum_{i=1}^n c_i l_i \gamma_i^{\beta_i} = \Phi^* \quad (11)$$

Задача (10), (11) легко решается методом множителей Лагранжа, и при всех $\beta_i = 1$ оптимальное решение определяется выражением:

$$\gamma_i = \sqrt{\frac{\alpha_i}{\lambda c_i}} - v_{0i}, \quad i = 1, \dots, n, \quad (12)$$

где $\lambda = \left(\frac{\sum_{i=1}^n l_i \sqrt{c_i \alpha_i}}{\sum_{i=1}^n l_i c_i v_{0i} + \Phi^*} \right)^2$.

Здесь λ – множитель Лагранжа.

Используя параметры сети, приведенные в табл. 2, по формуле (12) получаем на дорогах 7–1, 2–8, ..., 9–4 оптимальные добавляемые скорости (в км/час) 42,59; 41,31; 18,91; 12,95; 41,98; 64,45; 26,11; 23,57; 25,16; 13,09; 10,21; 8,11; 10,62. С учетом базовых скоростей средние скорости на дорогах составят в результате ремонта 52,59; 51,31; 28,91; 22,95; 51,98; 74,45; 36,11; 33,57; 35,16; 23,09; 20,21; 18,11; 20,62 (км/час).

Теперь рассмотрим, насколько это решение отличается от еще более простого решения, согласно которому ресурс Φ^* распределяется по дорогам равномерно, т. е. пропорционально длине дорог. На каждую дорогу при этом выделяется $\varphi_i(\gamma_i) = c_i l_i \gamma_i = \frac{l_i}{\sum l_i} \Phi^* = \frac{l_i}{3,191} \Phi^*$

долларов, и, следовательно, мы получаем добавленные скорости $\gamma_i = \Phi^* / (3,191 c_i)$. В сумме с базовыми скоростями в результате такого ремонта дорог 7–1, 2–8, ..., 9–4 на них достигаются средние скорости 60,07; 58,30; 55,39; 24,30; 48,22; 75,43; 24,16; 23,91; 48,15; 20,96; 24,79; 27,82; 25,39 (км/час).

Для сравнения качества двух рассмотренных решений в широком диапазоне нагрузок выполнены соответствующие имитационные эксперименты с транспортной сетью.

Имитационная модель воспроизводит следующие особенности движения транспортных потоков, не учитываемые ПБУ-сетью:

- скорость автомобиля, входящего на дорогу, определяется фактической (т. е. случайной) загрузкой дороги на текущий момент времени;
- при попытке перехода на следующую дорогу могут возникать блокировки (если на выбранную дорогу автомобиль перейти не может из-за ее полной занятости, то он не покидает текущую дорогу).

Модель имеет достаточно много и более тонких соответствий реальным транспортным потокам, в том числе возможность учета длительности фаз светофоров. Перечисление и обоснование этих соответствий не входит в задачи нашей статьи.

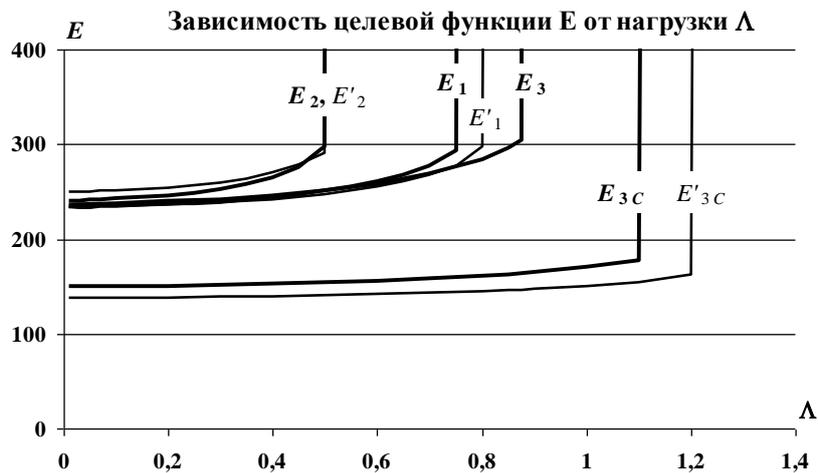


Рис. 4. Изменение качества оптимизированных сетей с ростом нагрузки

На рис. 4 сравниваются $E(\Lambda)$ -характеристики ПБУ-сети, оптимизированной в режиме малой нагрузки (тонкая линия E'_1), и соответствующей транспортной сети (полужирная кривая E_1). При изменяемых Λ характеристика E'_1 рассчитана по формуле (8) при добавочных скоростях (12) и заданной в табл. 1 матрице P . Характеристика E_1 получена путем прогонов ИМ. Кривые E_2 и E'_2 являются аналогичным образом полученными характеристиками транспортной сети и ПБУ-сети, в которых ресурс Φ^* распределялся по дорогам пропорционально их длине,

т. е. в равном количестве на каждый километр дорог. Интенсивность Λ указана в числе автомобилей в секунду, время E – в секундах. Сравнение полученных характеристик показывает, что оптимизированные по нулевой загрузке сети обладают заметно лучшими пропускными способностями, чем сети, ресурс ремонта которых распределялся равномерно. В стационарном режиме среднее время проезда через оптимизированные сети меньше, и они переходят в нестационарный режим (вертикальные участки графиков) при большей интенсивности Λ входящего в сеть потока. Соответственно, если в час пик заторы возникнут, то они наступят позже, будут мягче, короче и расформируются быстрее.

Характеристики E_3 , E_{3c} и E'_{3c} , показанные на рис. 4, достигаются за счет комплексной оптимизации, учитывающей нагрузку и возможность изменения переходных вероятностей.

Оптимизация транспортных ПБУ-сетей в режиме высокой нагрузки.

При ненулевой интенсивности λ_i входа автомобилей на дорогу i возникает положительный коэффициент ее загрузки $\rho_i > 0$. Это приводит к уменьшению значения ФД $\delta_i(\rho_i) = \delta(\rho_i)$ (рис. 3) и уменьшению фактической средней скорости $v_i = v_{0i}\delta(\rho_i)$ на дороге. Учитывая ФД, запишем для коэффициента загрузки ρ_i произвольной дороги уравнение:

$$\rho_i = \frac{\lambda_i b_i}{K_i} = \frac{\lambda_i}{K_i} \frac{l_i}{v_{0i} \cdot \delta(\rho_i)} = \frac{\rho_{0i}}{\delta(\rho_i)} = \rho_{0i}(1 + \rho_i^2), \quad (13)$$

где l_i – длина дороги;

$\rho_{0i} = \lambda_i l_i / (v_{0i} K_i)$ – виртуальный коэффициент загрузки, который имел бы место, если бы скорость автомобилей с ростом нагрузки не снижалась.

Уравнение (13) приближенно выполняется в стационарном режиме. Решение квадратного уравнения (13) имеет вид:

$$\rho_i = \frac{1}{2\rho_{0i}} - \sqrt{\frac{1}{(2\rho_{0i})^2} - 1} = \frac{v_{0i} K_i}{2\lambda_i l_i} - \sqrt{\left(\frac{v_{0i} K_i}{2\lambda_i l_i}\right)^2 - 1}, \quad (14)$$

где $\lambda_i = \alpha_i \Lambda$. Это решение корректно (соответствует стационарному режиму) при $\rho_{0i} \leq 1/2$ (в этом случае $\rho_i \leq 1$).

Используя решение (14), выразим целевую функцию (8) непосредственно через параметры ПБУ-сети и сформулируем задачу ее комплексной оптимизации следующим образом.

Требуется найти вектор скоростей $\vec{\gamma} = \vec{\gamma}_{opt}$ и маршрутную матрицу $\mathbf{P} = \mathbf{P}_{opt}$, доставляющие минимум среднего времени ответа $E(\vec{\gamma}, \mathbf{P})$ при заданном $\Lambda > 0$:

$$E(\vec{\gamma}, \mathbf{P}) = \sum_{i=1}^n \frac{\alpha_i(\mathbf{P}) l_i}{v_{0i} + \gamma_i} (1 + \rho_i^2) = \sum_{i=1}^n \frac{\alpha_i(\mathbf{P}) l_i}{v_{0i} + \gamma_i} \left(1 + \left(\frac{v_{0i} K_i}{2\alpha_i(\mathbf{P}) l_i} - \sqrt{\left(\frac{v_{0i} K_i}{2\alpha_i(\mathbf{P}) l_i}\right)^2 - 1} \right)^2 \right) \rightarrow \min_{\vec{\gamma}, \mathbf{P}} \quad (15)$$

и принадлежащие области допустимых решений:

$$\sum_{i=1}^n c_i l_i \gamma_i^{\beta_i} = \Phi^*; \quad 0 \leq \rho_i \leq 1; \quad \sum_{j=0}^n p_{ij} = 1, \quad (i = \overline{0, n}); \quad p_{ij}^{\min} \leq p_{ij} \leq p_{ij}^{\max}, \quad (i, j = \overline{1, n}). \quad (16)$$

Эту задачу комплексной оптимизации удобно решать путем многократной поочередной оптимизации текущего решения только по добавочным скоростям γ_i и только

по матрице \mathbf{P} – такие итерации очень быстро сходятся к решению, оптимальному одновременно и по добавочным скоростям, и по маршрутной матрице \mathbf{P} .

При небольших размерах сетей, когда число дорог находится в пределах сотни, сети можно оптимизировать с помощью стандартных программ оптимизации, использующих для вычисления градиентов численное дифференцирование. При этом затраты времени персонального компьютера средней мощности составляют от доли секунды до нескольких секунд. В случае, когда ПБУ-сети содержат сотни узлов, нужен точный расчет градиентов целевой функции по переходным вероятностям. Такой расчет быстро выполняется методом вспомогательных СЛАУ, предложенным в статье.

Решая задачу (15), (16) при некоторой $\Lambda > 0$, мы находим такое распределение ресурса Φ^* по ремонтируемым дорогам, которое наиболее эффективно проявится при интенсивности входящего в сеть потока, равной Λ .

На рис. 4 характеристика E_3 транспортной сети рассчитана путем ее имитационного моделирования при параметрах γ_i , полученных решением задачи (15), (16) при $\Lambda = 0,75$ автомобиля в секунду и при фиксированной матрице \mathbf{P} , заданной в табл. 1.

Характеристика E_{3C} рассчитана путем имитационного моделирования транспортной сети при параметрах γ_i и \mathbf{P} , полученных решением задачи (15), (16) при $\Lambda = 0,75$ автомобиля в секунду, при варьировании скоростей γ_i и при варьировании вероятностей p_{ij} в диапазоне $\pm 0,1$ (т. е. при параметрах, полученных путем комплексной оптимизации сети).

Кривая E'_{3C} характеризует соответствующую ПБУ-сеть и рассчитана без имитационного моделирования, непосредственно по выражению целевой функции (15).

Приведем решение $\vec{\gamma} = \vec{\gamma}_{opt}$, $\mathbf{P} = \mathbf{P}_{opt}$, задачи (15), (16), полученное для $\Lambda = 0,75$ и обеспечивающее характеристику E'_{3C} для ПБУ-сети и характеристику E_{3C} для соответствующей транспортной сети (см. рис. 4):

$\vec{\gamma}_{opt} = 46,63; 44,15; 13,60; 7,54; 33,43; 80,24; 39,16; 26,09; 21,28; 7,37; 9,63; 16,63; 19,25;$
 \mathbf{P}_{opt} представлено в табл. 3.

Таблица 3

Переходные вероятности матрицы \mathbf{P}_{opt}

Дорога i	Дорога j													
	0	7-1	2-8	9-7	7-8	3-8	1-7	7-9	8-7	8-2	8-3	9-5	9-6	9-4
0			0,3			0,2	0,5							
7-1							1							
2-8									0,6	0,2	0,2			
9-7		0,3			0,1			0,6						
7-8									0,6	0,2	0,2			
3-8									0,6	0,2	0,2			
1-7		0,3			0,1			0,6						
7-9				0,1								0,2	0,3	0,4
8-7		0,3			0,1			0,6						
8-2														
8-3			1											
9-5	1					1								
9-6	1													
9-4	1													

Сравнивая характеристики E_{3C} и E'_{3C} с другими характеристиками, приведенными на рис. 4, можно увидеть, что комплексная оптимизация, ориентированная на большую нагрузку и учитывающая возможность частичного изменения (адаптации) маршрутов, позволяет существенно снизить среднее время проезда через сеть и повысить ее пропускную способность.

Экономический эффект, достигаемый оптимизацией транспортных сетей путем быстрой численной оптимизации соответствующих ПБУ-сетей, вполне очевиден.

Столь же эффективные решения получаются и в случае более «жестких» ФД, в том числе в случае, когда ФД описывается функцией, изображенной на рис. 3 нижней линией.

Заключение. Для немарковских сетей массового обслуживания разработан эффективный полуаналитический метод комплексной оптимизации маршрутной матрицы и быстродействий узлов. В качестве целевой функции используется среднее время прохождения заявки через сеть. Особенностью предлагаемого метода является эффективная реализация двух его основных частей:

- быстрого точного расчета градиента целевой функции в пространстве варьируемых переходных вероятностей методом решения дополнительных СЛАУ;
- ускоренного аналитико-имитационного расчета с повышенной точностью градиента целевой функции в пространстве варьируемых быстродействий узлов методом направляющих гипербол.

В методе направляющих гипербол время ожидания в очередях рассчитывается путем имитационного моделирования. Частные производные целевой функции по быстродействиям узлов рассчитываются в ходе моделирования путем построения адаптивных сепарабельных гиперболических аппроксимаций целевой функции. В целом разработанный метод позволяет за один сеанс работы на персональном компьютере осуществлять комплексную оптимизацию немарковских сетей, содержащих сотни одноканальных и многоканальных узлов – систем массового обслуживания.

Разработан метод комплексной оптимизации транспортных сетей для стационарных режимов в широком диапазоне изменения транспортной нагрузки. Метод основан на аппроксимации транспортных сетей немарковскими ПБУ-сетями с многолинейными узлами-дорогами, быстродействие которых падает с ростом нагрузки, а буферы для очередей отсутствуют, вследствие чего заявки ожидают освобождения в следующем узле, не покидая предыдущего.

Предложены эффективные методы комплексной оптимизации транспортных ПБУ-сетей, ориентированной на различные режимы транспортной нагрузки. Показано, что наибольший эффект в борьбе с транспортными заторами дает комплексная оптимизация, ориентированная на режим высокой транспортной нагрузки.

Литература

1. Задорожный В. Н. Оптимизация однородных немарковских сетей массового обслуживания // Проблемы упр. 2009. № 6. С. 68–75.
2. Клейнен Дж. Статистические методы в имитационном моделировании / пер с англ.; под ред. Ю. П. Адлера и В. Н. Варыгина. М. : Статистика, 1978. Вып. 1. 221 с.
3. Johnson M. E. Jackson J. Infinitesimal Perturbation Analysis: a Tool for Simulation // J. of the Operational Res. Soc., 1989. Vol. 40, № 3. P. 134–160.
4. Rubinstein R. Y. Sensitivity analysis of computer simulation models via the efficient score // Oper. Res., 1989. Vol. 37. P. 72–81.
5. Suri R, Zazanis M. Perturbation Analysis Gives Strongly Consistent Sensitivity Estimates for the M|G|1 Queue // Mgmt Science, 1988. Vol. 34. P. 39–64.
6. Zadorozhnyi V. N. Optimization of Uniform Non-Markov Queueing Networks using resources and transition probabilities redistribution // Communications in Computer and Information Science, 2016. Vol. 638. P. 366–381.
7. Zadorozhnyi V. N., Zakharenkova T. R. Optimization of Channel Distribution over Nodes in Networks with Fractal Traffic // Dynamics of Systems, Mechanisms and Machines: conference proceeding, 15–17 Nov. 2016. P. 1–5.
8. Zadorozhnyi V. N. Simulation modeling of fractal queues // Dynamics of Systems, Mechanisms and Machines (Dynamics), December, 2014. P. 1–7.
9. Tsitsiashvili G. Sh. Parametric and Structural Optimization of the Queueing Network Throughput // Automation and Remote Control, 2007. Vol. 68, № 7. P. 1177–1185.

10. Клейнрок Л. Вычислительные системы с очередями / пер. с англ.; под ред. Б. С. Цыбакова. М. : Мир, 1979. 600 с.
11. Вишневский В. М. Теоретические основы проектирования компьютерных сетей. М. : Техносфера, 2003. 512 с.
12. Моисеев А. Н, Назаров А. А. Бесконечнолинейные системы и сети массового обслуживания. Томск : Изд-во НТЛ, 2015. 240 с.
13. Хейт Ф. Математическая теория транспортных потоков. М. : Мир, 1966. 288 с.
14. Дрю Д. Теория транспортных потоков и управление ими / пер. с англ. Е. Г. Коваленко и Г. Д. Шермана ; под ред. Н. П. Бусленко. М. : Транспорт, 1972. 424 с.
15. Daganzo C. F. Urban Gridlock: Macroscopic Modeling and Mitigation Approaches // Transportation Part B. 2007. Vol. 41. P. 49–62.
16. Транспортная стратегия Российской Федерации на период до 2030 года. Проект, Москва, 5 Августа 2013 г. 183 с. URL: <https://mintrans.ru/documents/> (дата обращения: 27. 12. 2018).
17. Zadorozhnyi V. N., Kornach M. A. Optimization of transport queuing networks on the basis of the method of directing hyperbole // 2016 International Siberian conference on control and communications SIBCON, sec. Control of the Large-Scale Systems, Russia, Moscow, May 12–14, 2016. P. 1–4.