УДК 519.21 **DOI** 10.15622/sp.2019.18.4.809-830

Н.А. Кузнецов, К.В. Семенихин

ОПТИМИЗАЦИЯ ПАРАМЕТРОВ ПЕРЕДАЧИ ДАННЫХ ПРИ НАЛИЧИИ МЕХАНИЗМА ПОВТОРНОЙ ОТПРАВКИ ПАКЕТОВ

Кузнецов Н.А., Семенихин К.В. **Оптимизация параметров передачи данных при наличии механизма повторной отправки пакетов.**

Аннотация. Процесс передачи данных описывается марковской моделью замкнутой сети массового обслуживания, которая состоит из двух систем (основной и вспомогательной). Основная система является конечной и одноканальной; она реализует процесс отправки пакетов по каналу связи с потерями. Вспомогательная система, будучи многоканальной, накапливает пакеты, потерянные основной системой, и пересылает их обратно в основную систему для повторной отправки. Скорость передачи пакетов основной системой и скорость их пересылки вспомогательной системой находятся в заданных диапазонах и подлежат оптимизации с целью минимизации времени успешной доставки и объема использованных ресурсов сети. Для указанных характеристик в стационарном режиме определены явные выражения, которые позволяют сформулировать задачу двукритериальной оптимизации. Определены оптимальные стратегии в двух постановках: в первой задаче минимизируется среднее время успешной передачи при ограничении на ресурсы; во второй задаче минимизируется расход ресурсов сети с учетом ограничения на время успешной передачи. Описано множество Парето-оптимальных стратегий в двукритериальной постановке за счет решения задачи минимизации расширенного функционала. Проанализировано качество приближенных решений, не учитывающих интенсивность обслуживания во вспомогательной системе.

Ключевые слова: замкнутая сеть массового обслуживания, процесс рождения и гибели, оптимизация, скорость передачи данных, интенсивность повторной отправки.

1. Введение. Модели систем с механизмом повторного обслуживания активно используются для описания современных инфокоммуникационных сетей, в которых потеря данных (вызовов, запросов) является критичной с точки зрения качества обслуживания [1-3]. Теория систем массового обслуживания (СМО) с повторными вызовами востребована при анализе характеристик контакт-центров [4, 5], поскольку для их адекватного описания необходимо учитывать наличие повторных обращений, возникающих при сильной загрузке операторов, окончании времени ожидания и тому подобное.

Для оптимизации качества передачи однотипных данных (пакетов) используют различные дисциплины активного управления очередями (Active Queue Management) [6]. Они основаны на отбрасывании части трафика для предотвращения перегрузки буферов и сигнализации передающей стороне о наличии коллизий. Математической моделью системы обычно служит одноканальная СМО. Сравнение классических механизмов типа RED (Random Early Detection) [7], их обобщений и дисциплин,

основанных на отбрасывании части очереди (Tail Drop), проведено в [8, 9] с использованием двух характеристик: средней длины очереди и вероятности потерь. В [10] рассмотрена одна характеристика, подлежащая максимизации на классе пороговых стратегий доступа, — это средний доход от работы СМО с учетом стоимости простоя/эксплуатации системы и задержек/потерь пакетов. Указанные работы основаны на исследовании системы в стационарном режиме.

На конечном промежутке времени динамические задачи управления доступом и загрузкой в одноканальных СМО с нестационарным входным потоком решены в [11, 12]. В [13] в задаче об оптимальном управлении элементами марковской сети массового обслуживания (СеМО) получено каноническое представление решения уравнения динамического программирования и разобраны примеры аналитического нахождения оптимальных стратегий. Для тандемных моделей СеМО, моделирующих процесс передачи данных по ненадежному каналу связи, оптимальное управление загрузкой канала с учетом нескольких конкурирующих требований к качеству обслуживания построено в [14, 15]. Необходимо отметить, что в этих работах целевым показателем выступало среднее число потерь, которое требовалось минимизировать с учетом ограничений на время полного обслуживания и объем энергозатрат, связанных с процессом передачи информации.

Если же потери не допускаются, то модель очереди необходимо дополнить вспомогательной системой, называемой иногда «орбитой», которая аккумулирует потерянные пакеты для повторной отправки [1]. Такая модель использовалась в [16] при решении задачи о жидкостной и диффузионной аппроксимации замкнутой марковской СеМО в загруженном состоянии. В [17] доказана пороговая структура оптимальной децентрализованной стратегии при управлении интенсивностью обслуживания на каждом узле замкнутой сети по критерию минимума стоимости удержания заявки и ресурсов системы. Задача управления загрузкой замкнутой сети в форме параллельного соединения тандемных СМО рассмотрена в [18], где максимизируется взвешенное число законченных фаз обслуживания. Одна из последних публикаций по этой тематике [19] посвящена диффузионной аппроксимации сети, в которой основная и вспомогательная системы имеют вид $G/G/1/N$ и $M/M/\infty$ соответственно, а уход на орбиту происходит с заданной вероятностью.

Анализ этих и других публикаций показывает, что модель замкнутой сети с повторными вызовами не рассматривалась ранее с точки зрения оптимизации двух характеристик качества передачи данных — времени успешной доставки пакета и объема использованных ресурсов. Поэтому

в настоящей работе рассматривается замкнутая СеМО, которая включает основную систему — одноканальную СМО, как модель процесса передачи с потерями, и вспомогательную систему — многоканальную СМО для повторной отправки потерянных пакетов. При этом обе интенсивности, как в основной системе, так и во вспомогательной, подлежат оптимизации на классе постоянных стратегий. В этом состоит отличие настоящего исследования от работы авторов [20], где для аналогичной модели сети определяется оптимальное управление скоростью передачи данных в виде функции времени или состояния при фиксированной интенсивности повторной отправки.

2. Описание модели и постановка задачи. На рисунке 1 изображена схема рассматриваемой сети передачи данных. Она состоит из двух систем: основной и вспомогательной. Основная система содержит конечную очередь и один элемент обслуживания (передатчик), который посылает данные по каналу связи с потерями при наличии постоянного входного трафика, составленного из однотипных пакетов. Вспомогательная система (орбита) является многоканальной и используется для описания повторной отправки пакетов в случае их потери.

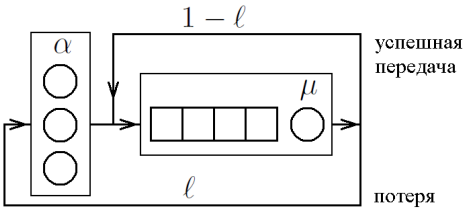

Рис. 1. Схема сети: слева — вспомогательная система, справа — основная система (α — интенсивность повторной отправки, ℓ — вероятность потери, μ — скорость передачи)

Если пакет успешно отправлен, то он покидает основную систему. Затем к ней мгновенно присоединяется новый пакет, подлежащий дальнейшей отправке. В этом случае число заявок в основной системе не меняется.

При потере, которая имеет место с вероятностью ℓ, число заявок в основной системе уменьшается на единицу, а число заявок во вспомогательной системе увеличивается на единицу. При этом внешний пакет блокируется и ожидает момента входа в основную систему, который наступает при первой успешной отправке. Блокировка внешнего трафика

происходит также в случае, если вспомогательная система заполнена полностью.

Тем самым рассматриваемая сеть является замкнутой, то есть в ней циркулирует постоянное число пакетов N.

Если обозначить через μ интенсивность обслуживания в основной системе, а через α — интенсивность обслуживания во вспомогательной системе, то можно утверждать следующее: μ — скорость передачи пакета без учета того, доставлен ли он успешно или нет; $(1-\ell)\mu$ — скорость успешной отправки; $\ell\mu$ — интенсивность потерь, то есть отправки пакета на орбиту; α — скорость отсылки наугад взятого пакета из вспомогательной системы в основную; $\alpha(N-n)$ — интенсивность схода с орбиты одного из пакетов, если в основной системе находится n пакетов.

Итак, рассматриваемая СеМО состоит из двух узлов:

$\cdot/M_\alpha/N/0$ — вспомогательная система (узел «0»);

$\cdot/M_\mu/1/N-1$ — основная система (узел «1»).

Маршрутизация между ними осуществляется с помощью марковской цепи, у которой переход $0 \to 1$ имеет место всегда, а переход $1 \to 0$ осуществляется с вероятностью ℓ. Поэтому работу сети будем описывать с помощью марковского процесса рождения и гибели $X(t) \in E = \{0, 1, \ldots, N\}$, где значения $X(t)$ и $N - X(t)$ равны числу пакетов в основной и вспомогательной системах соответственно. Матрица интенсивностей $\Lambda = \{\lambda_{n,p}\}_{n,p\in E}$ однозначно определяется интенсивностями двух переходов $n \to n \pm 1$:

$$\lambda_{n,n+1} = \alpha(N-n) \quad \text{при} \quad n < N \quad \text{и} \quad \lambda_{n,n-1} = \ell\mu \quad \text{при} \quad n > 0.$$

Качество работы рассматриваемой сети будет определяться двумя показателями: $\mathfrak{S}$ — среднее время полной передачи пакета; $\mathfrak{R}$ — объем ресурсов, использованных при передаче данных и обработке запросов на повторную отправку. Оба показателя рассматриваются как функции интенсивностей обслуживания μ и α.

Цель работы — оптимизация введенных показателей качества в рамках рассматриваемой модели сети за счет выбора двух управляемых характеристик: скорости передачи данных μ и интенсивности обработки повторных запросов α.

Поставленная цель обуславливает необходимость анализа проблемы двукритериальной оптимизации:

$$\big(\mathfrak{S}[\mu,\alpha], \mathfrak{R}[\mu,\alpha]\big) \to \min_{\mu,\alpha}.$$

Для решения этой проблемы будем рассматривать три варианта оптимизационных постановок:

1) наискорейшая передача данных с учетом ограничения на ресурсы сети:

$$\mathfrak{S}[\mu,\alpha] \to \min_{\mu,\alpha}: \quad \mathfrak{R}[\mu,\alpha] \leqslant R^{\max}; \tag{1}$$

2) минимизация потребления ресурсов при ограниченном времени успешной отправки данных:

$$\mathfrak{R}[\mu,\alpha] \to \min_{\mu,\alpha}: \quad \mathfrak{S}[\mu,\alpha] \leqslant S^{\max}; \tag{2}$$

3) синтез оптимальных стратегий относительно расширенного функционала:

$$\mathfrak{L}[\mu,\alpha,\lambda] = \mathfrak{S}[\mu,\alpha] + \lambda\,\mathfrak{R}[\mu,\alpha] \to \min_{\mu,\alpha} \tag{3}$$

при любом значении множителя $\lambda \geqslant 0$.

Перечисленные задачи будут изучаться применительно к модели однородного марковского процесса $X(t)$, рассматриваемого в стационарном режиме при использовании постоянных стратегий μ и α, которые подчиняются априорным ограничениям:

$$\mu \in [m_{\min}, m_{\max}], \quad \alpha \in [a_{\min}, a_{\max}], \tag{4}$$

где $0 < m_{\min} < m_{\max} < \infty$, $0 < a_{\min} < a_{\max} < \infty$ — заданные нижние и верхние границы интенсивностей.

3. Функционалы качества. Подробный вывод выражений для функционалов $\mathfrak{S}$ и $\mathfrak{R}$ представлен в [20]. Поэтому здесь опишем только конструкцию соответствующих случайных величин и окончательные выражения для оптимизируемых показателей.

Сначала рассмотрим величину S — время успешной доставки пакета:

$$S = W + B_1 + A_2 + B_2 + \ldots + A_\nu + B_\nu,$$

где W — время ожидания входа в систему; A_k и B_k — времена пребывания пакета во вспомогательной и основной системе соответственно при k-й попытке ($k = 1, 2, \ldots$ и $A_1 = 0$); ν — номер попытки, при которой пакет будет успешно передан.

Величины W и B_k имеют представление:

$$W = W_0\,\mathrm{I}\{X = 0\} + (\tau_1 + \ldots + \tau_{\varkappa})\,\mathrm{I}\{X > 0\};$$

$$B_k = (\tau_1 + \ldots + \tau_X)\,\mathrm{I}\{X > 0\},$$

где W_0 — время ожидания разблокировки при полной загрузке вспомогательной системы; $\tau_1, \tau_2, \ldots$ — времена обслуживания заявок, находящихся в основной системе; $\varkappa$ — номер первого успешно переданного пакета; X — состояние основной системы; $\mathrm{I}\{\ldots\}$ — индикатор случайного события.

В стационарном режиме условные законы распределения величин W_0, τ_k, A_k являются экспоненциальными:

$$\begin{cases} \mathsf{Law}\{W_0 \mid X = 0\} = E(\alpha N), \\ \mathsf{Law}\{\tau_k \mid X = n\} = E(\mu), \qquad n > 0, \\ \mathsf{Law}\{A_k \mid X = n\} = E(\alpha), \qquad n < N, \end{cases}$$

а величины ν и $\varkappa$ подчиняются геометрическому распределению $G(1-\ell)$.

Выражения для математических ожиданий величин S, W, B можно получить с помощью тождества Вальда [21]:

$$\mathsf{M}\,S = \mathsf{M}W + \mathsf{M}B + (\mathsf{M}A + \mathsf{M}B)\,\mathsf{M}(\nu - 1),$$

$$\mathsf{M}W = \frac{\mathsf{P}\{X = 0\}}{\alpha N} + \mathsf{M}\varkappa\,\frac{\mathsf{P}\{X > 0\}}{\mu},$$

$$\mathsf{M}A = \frac{\mathsf{P}\{X < N\}}{\alpha}, \quad \mathsf{M}B = \frac{\mathsf{M}(X\,\mathrm{I}\{X > 0\})}{\mu}.$$

В стационарном режиме функционал, описывающий среднее время успешной доставки пакета, принимает вид:

$$\mathfrak{S}[\mu, \alpha] = \frac{\mathsf{P}\{X = 0\}}{\alpha N} + \frac{\mathsf{P}\{X < N\}\ell}{(1-\ell)\alpha} + \frac{\mathsf{M}\big((X+1)\,\mathrm{I}\{X > 0\}\big)}{(1-\ell)\mu}.$$

Предположим, что скорость потребления ресурсов пропорциональна интенсивностям обслуживания μ, α, и введем коэффициент c, равный отношению цены загрузки одного сервера вспомогательной системы к стоимости загрузки передатчика. Теперь можно записать выражение для функционала, определяющего средний расход ресурсов сети в стационар-

ном режиме:

$$\mathfrak{R}[\mu,\alpha] = \mu\,\mathsf{P}\{X>0\} + c\alpha(N - \mathsf{M}X).$$

4. Приближенный анализ оптимизационной модели. При любом выборе стратегий $\mu>0$ и $\alpha>0$ однородный марковский процесс $X(t)$ имеет стационарное распределение $\pi=\{\pi_n\colon n\in E\}$, то есть:

$$\pi_n = \lim_{t\to\infty} \mathsf{P}\{X(t)=n \mid X(0)=x\} \quad \forall\, x,n\in E.$$

Для этих вероятностей известно рекуррентное представление:

$$\pi_n = \frac{\alpha(N+1-n)}{\ell\mu}\pi_{n-1}, \quad n=1,\dots,N, \tag{5}$$

где π_0 находится из условия нормировки. Просуммировав уравнения (5) по n, получим:

$$\ell\mu(1-\pi_0) = \alpha(N-\mathsf{M}X).$$

Пользуясь этим равенством, функционалы можно представить следующим образом:

$$\mathfrak{S}[\mu,\alpha] = \frac{\pi_0}{\alpha N} + \frac{(\pi_0-\pi_N)\ell}{(1-\ell)\alpha} + \frac{N+1-\pi_0}{(1-\ell)\mu}, \tag{6}$$

$$\mathfrak{R}[\mu,\alpha] = \mu(1+c\ell)(1-\pi_0). \tag{7}$$

Вероятности π_0,π_N можно представить в явной форме:

$$\pi_0 = \Big(\sum_{n=0}^{N}\frac{N!\,\rho^n}{(N-n)!}\Big)^{-1}, \quad \pi_N = \Big(\sum_{n=0}^{N}\frac{1}{n!\,\rho^n}\Big)^{-1}, \quad \text{где} \quad \rho=\frac{\alpha}{\ell\mu}.$$

Численный анализ показывает, что если

$$N>30, \quad \max(2,1/(\ell N))/N < \rho < 1/7, \tag{8}$$

то $\max(\pi_0,\pi_N)<10^{-3}$. Поэтому можно пренебречь вероятностями π_0,π_N и воспользоваться аппроксимацией:

$$\mathfrak{S}[\mu,\alpha] \approx \overline{S}[\mu] = \frac{N+1}{(1-\ell)\mu}, \qquad \mathfrak{R}[\mu,\alpha] \approx \overline{R}[\mu] = \mu(1+c\ell), \tag{9}$$

причем приближенные значения $\overline{S}, \overline{R}$ отличаются от исходных менее чем на 0.03%.

Это дает право сделать следующий вывод. Оптимизируемые функционалы мало чувствительны к изменению параметра α, если отношение $\rho = \alpha/(\ell\mu)$ лежит в диапазоне (8). Это имеет место, когда исключены слишком маленькие и чересчур большие значения α, так как они приводят к переполнению одной из систем.

Приведем решения задач (1), (2) и (3), в которых исходные функционалы заменены их аппроксимациями $\overline{S}[\mu]$, $\overline{R}[\mu]$.

1. Из ограничения $\overline{R}[\mu] \leqslant R^{\max}$ получаем, что μ не превосходит величины $R^{\max}/(1 + c\ell)$, которая определяет оптимальную скорость передачи при минимизации среднего времени успешной отправки $\overline{S}[\mu]$ с учетом ограниченных ресурсов:

$$\mu^* = \frac{R^{\max}}{1 + c\ell}.$$

2. Для минимизации потребления ресурсов $\overline{R}[\mu]$ при наличии ограничения на время полной передачи $\overline{S}[\mu] \leqslant S^{\max}$ аналогично получается:

$$\mu^* = \frac{N + 1}{(1 - \ell)S^{\max}}.$$

3. Для нахождения оптимальной стратегий $\mu_\lambda \in (0, \infty)$ относительно расширенного функционала:

$$\overline{L}[\mu, \lambda] = \overline{S}[\mu] + \lambda\, \overline{R}[\mu], \quad \lambda \geqslant 0,$$

достаточно записать его в виде функции $a/\mu + b\mu$ с положительными коэффициентами $a = \lambda(1 + c\ell)$, $b = (N + 1)/(1 - \ell)$ и определить ее точку минимума:

$$\mu_\lambda = \sqrt{b/a} = \sqrt{\frac{N + 1}{\lambda(1 - \ell)(1 + c\ell)}}.$$

Перечисленные выше решения μ^*, μ_λ получены без учета априорных ограничений (4). Чтобы их учесть, достаточно спроектировать величины μ^*, μ_λ на отрезок $[m_{\min}, m_{\max}]$. Однако в разделах 1, 2 необходимо сначала проверить условие разрешимости, то есть: 1) неравенство $\overline{R}[m_{\min}] \leqslant R^{\max}$, гарантирующее соблюдение ограничения на ресурсы

при минимальной скорости передачи; 2) ограничение на время отправки при максимальной скорости передачи: $\overline{S}[m_{\max}] \leqslant S^{\max}$.

Взаимосвязь всех трех задач можно наглядно проследить по кривой:

$$R = \frac{(N+1)(1+c\ell)}{(1-\ell)S}, \tag{10}$$

уравнение которой записано в переменных $S = \overline{S}[\mu]$, $R = \overline{R}[\mu]$. Точки пересечения этой кривой с прямыми $R = R^{\max}$ и $S = S^{\max}$ определяют решения задач из разделов 1 и 2 соответственно. Кроме того, множество точек $(\overline{S}[\mu_\lambda], \overline{R}[\mu_\lambda])$, образующих решения задачи из п. 3 при всевозможных $\lambda > 0$, полностью совпадает с кривой (10). При этом увеличение параметра λ, характеризующего относительную стоимость потребления ресурсов, соответствует росту S и снижению R.

На рисунке 2 на плоскости переменных (S, R) изображены: сплошная кривая (10), описывающая зависимость $S = \overline{S}[\mu]$, $R = \overline{R}[\mu]$ между приближенными функционалами (9); штриховая кривая, определяющая истинные значения функционалов $S = \mathfrak{S}[\mu_\lambda, a_{\mathrm{med}}]$, $R = \mathfrak{R}[\mu_\lambda, a_{\mathrm{med}}]$ на приближенном решении расширенной задачи при различных $\lambda > 0$, где $a_{\mathrm{med}} = (a_{\min} + a_{\max})/2$; точки $(\mathfrak{S}[\mu, \alpha], \mathfrak{R}[\mu, \alpha])$, полученные на случайной выборке стратегий (μ, α).

Расчеты проведены для следующего набора параметров:

$$\begin{cases} m_{\min} = 0.4, \quad m_{\max} = 4, \quad \ell = 0.2, \\ a_{\min} = 0.01, \quad a_{\max} = 0.39, \quad c = 2.5. \end{cases} \tag{11}$$

На верхнем графике рисунка 2 представлены результаты для случая а) $N = 3$, а на нижнем — для случая б) $N = 30$. Таким образом, при маленьком числе пакетов N погрешность аппроксимации (9) является существенной, а разброс значений по критериям (особенно по времени отправки) — значительным. Однако при большом N приближенные значения $\overline{S}[\mu]$, $\overline{R}[\mu]$ визуально неотличимы от точных значений $\mathfrak{S}[\mu, \alpha]$, $\mathfrak{R}[\mu, \alpha]$ при всех $\alpha \in [a_{\min}, a_{\max}]$.

5. Оптимальные стратегии передачи данных и повторной отправки. Для анализа исходных задач (1), (2) и (3) удобно преобразовать функционал $\mathfrak{S}[\mu, \alpha]$ к виду:

$$\mathfrak{S}[\mu, \alpha] = \frac{\ell}{\alpha}\mathfrak{S}_0[\rho], \quad \rho = \frac{\alpha}{\ell\mu},$$

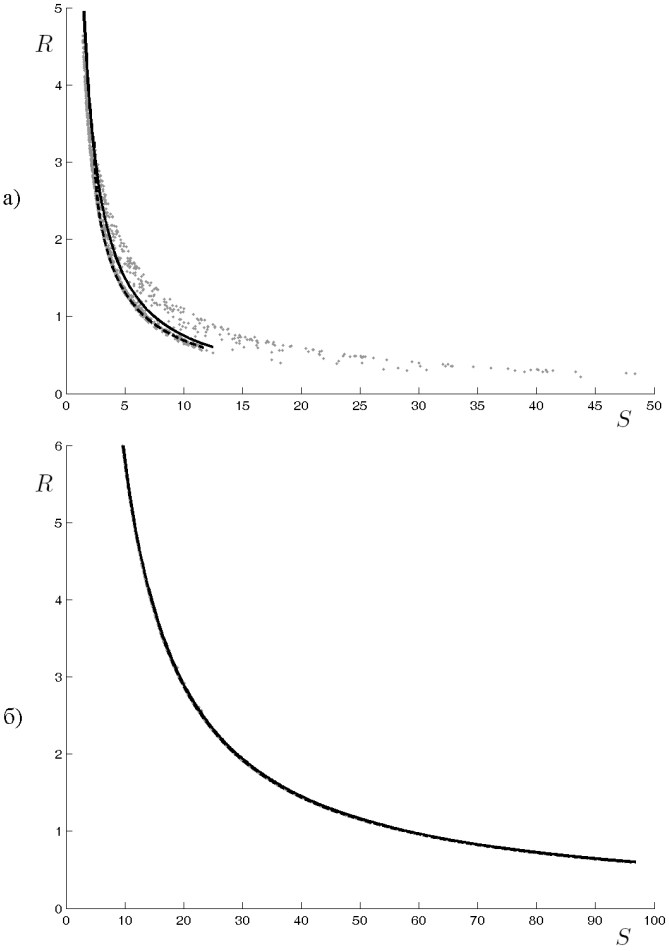

Рис. 2. Сравнение показателей в двух случаях: а) — сверху и б) — снизу

Труды СПИИРАН. 2019. Том 18 № 4. ISSN 2078-9181 (печ.),
ISSN 2078-9599 (онлайн), www.proceedings.spiiras.nw.ru

где в силу (6) функция $\mathfrak{S}_0[\rho]$ и вероятности $\pi_0 = \pi_0[\rho]$, $\pi_N = \pi_N[\rho]$ зависят только от переменной ρ:

$$\mathfrak{S}_0[\rho] = \frac{\pi_0[\rho]}{\ell N} + \frac{\pi_0[\rho] - \pi_N[\rho]}{1-\ell} + \frac{(N+1-\pi_0[\rho])\rho}{(1-\ell)}.$$

Относительно функции $\mathfrak{S}_0[\rho]$ можно утверждать, что она выпукла при $\rho > 0$, сходится к значению $1/(\ell N) + 1/(1-\ell)$ при $\rho \to 0$ и стремится к бесконечности при $\rho \to \infty$.

Итак, для минимизации функционала $\mathfrak{S}[\mu, \alpha]$ на множестве $\mathcal{U}$ пар (μ, α), допустимых в задаче (1), потребуется следующее: для коэффициента ρ определить отрезок значений $[\rho_1, \rho_2]$, при которых прямая $\alpha = \rho\ell\mu$ пересекает множество $\mathcal{U}$; для каждого $\rho \in [\rho_1, \rho_2]$ на пересечении прямой $\alpha = \rho\ell\mu$ с $\mathcal{U}$ взять максимальное значение α_ρ; для функции одной переменной $\mathfrak{S}_0[\rho]/\alpha_\rho$ найти точку минимума $\rho^* \in [\rho_1, \rho_2]$; положить $\alpha^* = \alpha_{\rho^*}$ и $\mu^* = \alpha^*/(\rho^*\ell)$. Теперь (μ^*, α^*) — искомое решение задачи (1).

Здесь стоит отметить вид множества $\mathcal{U}$. В силу (7) оно получается в результате пересечения прямоугольника (4) с областью $\mu(1 + c\ell) \leqslant R^{\max}/(1 - \pi_0[\rho])$. При фиксированном μ это неравенство задает ограничение снизу на убывающую функцию переменной ρ. Следовательно, значение ρ, то есть угол наклона прямых $\alpha = \rho\ell\mu$, ограничено сверху. Тем самым указанная область лежит ниже кривой $\mathfrak{R}[\mu, \alpha] = R^{\max}$.

Графическая иллюстрация этого факта представлена на рисунке 3, который описывает результаты вычислений для набора параметров (11) в двух случаях:

$$\text{а) } N = 3,\ R^{\max} = 2 \quad \text{и} \quad \text{б) } N = 30,\ R^{\max} = 5.$$

На рисунке 3 ограничения изображены сплошными линиями, линии уровня функционала $\mathfrak{S}$ — штриховыми кривыми, вектор антиградиента $(-\nabla\mathfrak{S})$ — стрелкой, а решение задачи (μ^*, α^*) — точкой.

Важно отметить, что ограничение на ресурсы порождает невыпуклое множество $\mathcal{U}$ (оно изображено на рисунке 3 в виде затемненной области). Искомая точка (μ^*, α^*), на которой достигается минимум среднего времени отправки, лежит на верхней правой части границы множества $\mathcal{U}$, что согласуется с направлением антиградиента $(-\nabla\mathfrak{S}[\mu, \alpha])$ и положением линий уровня $\mathfrak{S}[\mu, \alpha] = S_i$. Уровни заданы в виде $S_i = (5/i)S^*$, $i = 1, \ldots, 7$, где $S^* = \mathfrak{S}[\mu^*, \alpha^*]$. Линии уровня идут слева направо: $S_1 > \ldots > S_7$ на верхнем графике; $S_1 > \ldots > S_5$ на нижнем графике.

Существенное отличие случаев а) и б) состоит в том, чтобы при большой загрузке сети, то есть при большом N, функционал $\mathfrak{S}[\mu, \alpha]$ мало

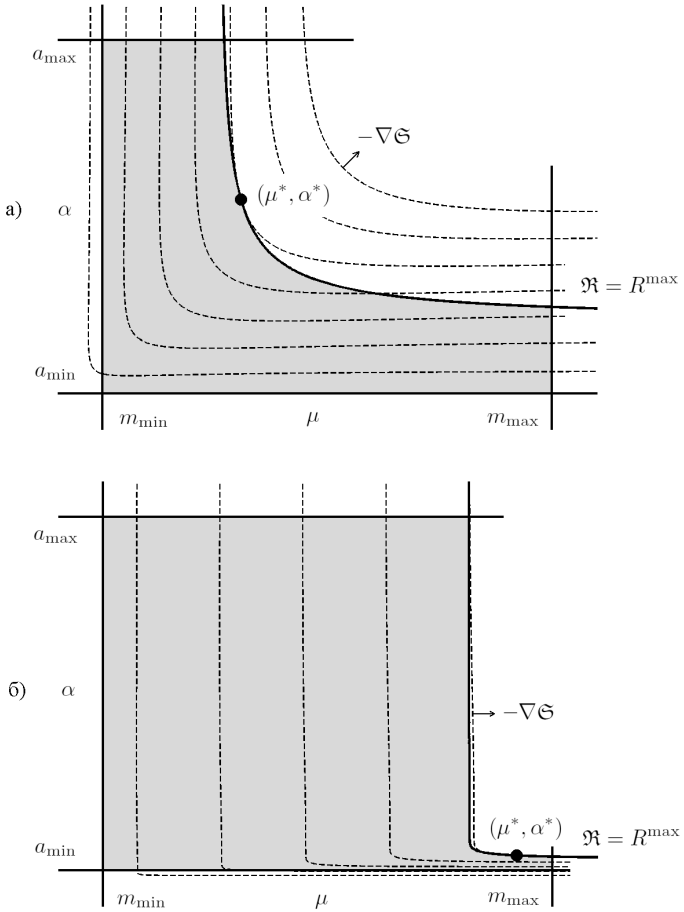

Рис. 3. Минимизация функционала $\mathfrak{S}$ в двух случаях: а) — сверху и б) — снизу

меняется вдоль ограничения $\mathfrak{R}[\mu,\alpha] = R^{\max}$. Напротив, при маленькой загрузке, что имеет место в случае а), функционал $\mathfrak{S}[\mu,\alpha]$ оказывается весьма чувствителен к вариации интенсивности обслуживания во вспомогательной системе α. Для подтверждения этого факта найдена стратегия $(\check{\mu},\check{\alpha})$, максимизирующая $\mathfrak{S}[\mu,\alpha]$ вдоль кривой $\mathfrak{R}[\mu,\alpha] = R^{\max}$. Результаты этих вычислений сведены в таблицы 1 и 2. Отметим, что некоторые из перечисленных там стратегий (μ,α) не являются допустимыми, поэтому на них значение $\mathfrak{S}[\mu,\alpha]$ оказывается меньше оптимума S^*, но за счет превышения лимита использования ресурсов $\mathfrak{R}[\mu,\alpha] \leqslant R^{\max}$.

Таблица 1. Характеристики стратегий в задаче (1) для случая а)

Скорость передачи μ	Интенсивность α	$\mathfrak{S}[\mu,\alpha]$	$\mathfrak{R}[\mu,\alpha]$
$\mu^* = 1.508$	$\alpha^* = 0.219$	$3.225 = S^*$	$2 = R^{\max}$
$\check{\mu} = 4$	$\check{\alpha} = 0.103$	4.787	$2 = R^{\max}$
$m_{\min} = 0.4$	$a_{\min} = 0.01$	49.593	0.195
$m_{\mathrm{med}} = 2.2$	$a_{\mathrm{med}} = 0.2$	2.667	2.508
$m_{\max} = 4$	$a_{\max} = 0.39$	1.411	4.691

Таблица 2. Характеристики стратегий в задаче (1) для случая б)

Скорость передачи μ	Интенсивность α	$\mathfrak{S}[\mu,\alpha]$	$\mathfrak{R}[\mu,\alpha]$
$\mu^* = 3.712$	$\alpha^* = 0.026$	$11.502 = S^*$	$5 = R^{\max}$
$\check{\mu} = 3.333$	$\check{\alpha} = 0.057$	11.625	$5 = R^{\max}$
$m_{\min} = 0.4$	$a_{\min} = 0.01$	96.867	0.6
$m_{\mathrm{med}} = 2.2$	$a_{\mathrm{med}} = 0.2$	17.475	3.3
$m_{\max} = 4$	$a_{\max} = 0.39$	9.605	6

Теперь перейдем к решению задачи (2). Пусть $\mathcal{V}$ — множество точек (μ,α), допустимых в этой задаче. Функционал $\mathfrak{R}[\mu,\alpha]$ равен произведению двух множителей $\mu(1+c\ell)$ и $1-\pi_0[\rho]$, первый из которых зависит только от μ, а второй полностью определяется значением ρ. Поэтому процедура минимизации $\mathfrak{R}[\mu,\alpha]$ аналогична: сначала определить границы ρ_1,ρ_2 значений ρ, при которых прямая $\alpha = \rho\ell\mu$ пересекает $\mathcal{V}$, затем на каждом таком пересечении взять минимальное значение μ_ρ, далее найти точку максимума $\rho^* \in [\rho_1,\rho_2]$ функции $\mu_\rho\pi_0[\rho]$, и в конце определить решение задачи (2) в виде $\mu^* = \mu_{\rho^*}$, $\alpha^* = \rho^*\ell\mu^*$.

Отметим, что множество $\mathcal{V}$ обязано быть выпуклым в переменных (ρ,α). Действительно, $\mathcal{V}$ — есть пересечение выпуклого четырехугольника

$$\{(\rho,\alpha)\colon \alpha/(\ell m_{\max}) \leqslant \rho \leqslant \alpha/(\ell m_{\min}),\ a_{\min} \leqslant \alpha \leqslant a_{\max}\}$$

и надграфика выпуклой функции $\mathfrak{S}_0[\rho]\ell/S^{\max} \leqslant \alpha$.

Графическая иллюстрация этого утверждения представлена на рисунке 4, на котором множество $\mathcal{V}$ представлено в виде затененной области, ограничения изображены сплошными линиями, линии уровня функционала $\mathfrak{R}$ — штриховыми кривыми, вектор антиградиента $(-\nabla\mathfrak{R})$ — стрелкой, а решение задачи (μ^*, α^*) — точкой.

Результаты вычислений получены для набора параметров (11) в двух случаях:

$$\text{а) } N = 3,\ S^{\max} = 3.225 \quad \text{и} \quad \text{б) } N = 30,\ S^{\max} = 11.502, \tag{12}$$

где значения $S^{\max}$ выбраны равными оптимуму S^* в предыдущей задаче (см. таблицы 1, 2).

Благодаря такому выбору параметров ограничения, решение задачи минимизации $\mathfrak{R}[\mu, \alpha]$ совпадает с точкой (μ^*, α^*), доставляющей минимум функционалу $\mathfrak{S}[\mu, \alpha]$ с учетом неравенства $\mathfrak{R}[\mu, \alpha] \leqslant R^{\max}$. Поэтому на полученном решении оптимум $R^* = \mathfrak{R}[\mu^*, \alpha^*]$ при ограничениях (12) оказывается равен $R^{\max} = 2$ в случае а) и $R^{\max} = 5$ в случае б).

При сравнении двух сценариев, когда нагрузка на сеть а) мала или, наоборот, б) велика, вновь наблюдается разная чувствительность критерия оптимизации к вариации интенсивности α. В случае а) имеет место четко выраженный минимум, а в случае б) относительное изменение функционала $\mathfrak{R}[\mu, \alpha]$ при движении вдоль границы $\mathfrak{S}[\mu, \alpha] = S^{\max}$ составляет около 1 %, в то время как α варьируется в очень широких пределах: от $\check{\alpha} = 0.058$ до $a_{\max} = 3.9$. О последнем факте свидетельствует также вид линий уровня $\mathfrak{R}[\mu, \alpha] = R_i$, которые на большей части диапазона $[a_{\min}, a_{\max}]$ являются почти вертикальными. Линии уровня идут слева направо по направлению роста значений $\{R_i\}$ с равным шагом: от $R_1 = (3/5)R^*$ до $R_9 = (11/5)R^*$ — на верхнем графике; от $R_1 = (12/15)R^*$ до $R_8 = (19/15)R^*$— на нижнем графике.

Теперь рассмотрим расширенный функционал (3). Через переменные ρ, α его можно записать как:

$$\mathfrak{L}[\mu, \alpha, \lambda] = \frac{\ell}{\alpha}\mathfrak{S}_0[\rho] + \lambda(1 + c\ell)\frac{\alpha(1 - \pi_0[\rho])}{\rho\ell}.$$

Если не учитывать априорные ограничения на α, то можно явно записать выражение для минимума по этой переменной:

$$\min_{\alpha>0} \mathfrak{L}[\alpha/(\rho\ell), \alpha, \lambda] = 2\sqrt{\lambda(1 + c\ell)\mathfrak{S}_0[\rho](1 - \pi_0[\rho])/\rho}.$$

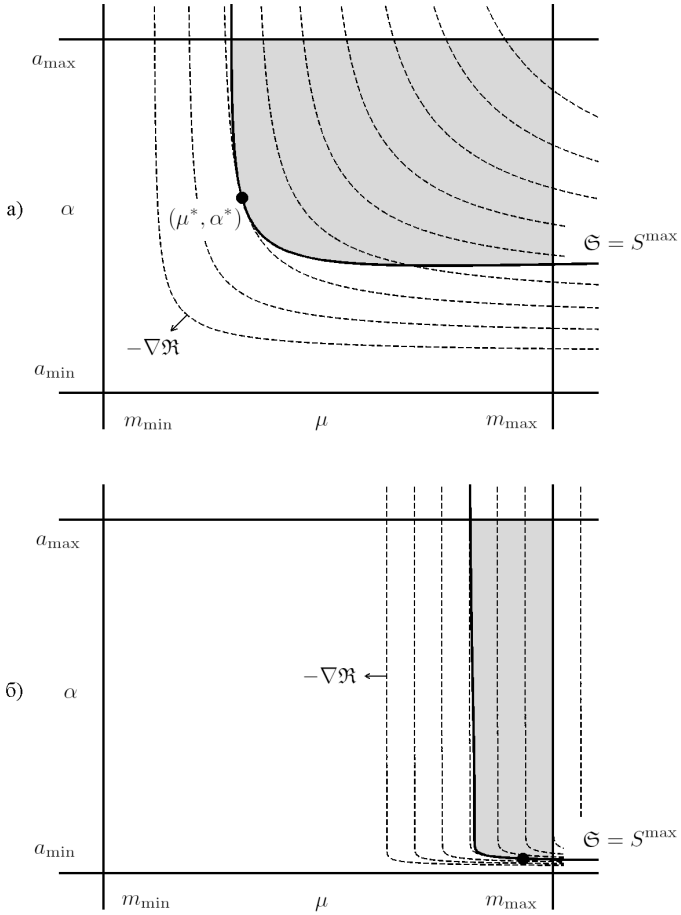

Рис. 4. Минимизация функционала $\mathfrak{R}$ в двух случаях: а) — сверху и б) — снизу

Полученная функция достигает минимума в единственной точке $\rho^* > 0$, не зависящей от выбора множителя λ. Если отбросить априорные ограничения на μ, то решение задачи (3) имеет вид:

$$\alpha_\lambda = \ell\left(\frac{\mathfrak{S}_0[\rho^*]\rho^*}{\lambda(1+c\ell)(1-\pi_0[\rho^*])}\right)^{1/2}, \quad \mu_\lambda = \frac{\alpha_\lambda}{\rho^*\ell}.$$

Таким образом, оптимальные стратегии $\mu_\lambda, \alpha_\lambda$ пропорциональны $1/\sqrt{\lambda}$. Это означает, что при увеличении параметра λ, характеризующего цену расходования ресурсов сети, интенсивности обслуживания в основной и вспомогательной системах должны снижаться. При этом соотношение между ними $\mu_\lambda/\alpha_\lambda$ остается неизменным.

Отметим, что расширенный функционал $\mathfrak{L}[\mu, \alpha, \lambda]$ не является выпуклым по (μ, α), но семейство решений $\{(\mu_\lambda, \alpha_\lambda)\}$ совпадает с множеством Парето-оптимальных стратегий.

Оба эти факта можно проиллюстрировать с помощью рисунка 5, на котором представлены результаты расчетов для случая $N = 3$ и того же набора параметров (11).

На верхнем графике рисунка 5 для расширенного функционала изображены его линии уровня $\mathfrak{L}[\mu, \alpha, \lambda^*] = L_i$ в виде штриховых кривых и точка минимума. Сплошные линии описывают априорные органичения. Значение множителя λ^* здесь выбрано таким образом, чтобы точка минимума расширенного функционала совпала с решением (μ^*, α^*) двух предыдущих задач. Согласно найденной выше зависимости $\lambda \mapsto \alpha_\lambda$, искомый множитель λ^* можно определить из равенства $\lambda^* = (\alpha_\lambda/\alpha^*)^2\lambda$. Величины $\{L_i\}$ взяты с равным шагом от $L_1 = 1.01\,L^*$ до $L_{20} = 1.2\,L^*$, где $L^* = \mathfrak{L}[\mu^*, \alpha^*, \lambda^*]$.

Нижний график рисунка 5 содержит набор точек, полученных на случайной выборке стратегий, и сплошную кривую $S = \mathfrak{S}[\mu_\lambda, \alpha_\lambda]$, $R = \mathfrak{R}[\mu_\lambda, \alpha_\lambda]$, которая была построена из решения расширенной задачи (3) при $\lambda \in (0.2, 3.3 \cdot 10^6)$. В отличие от приближенного решения, изображенного на рисунке 2 штриховой линией, сплошная кривая на рисунке 5 ограничивает снизу и слева область, которая включает все полученные точки. Это подтверждает оптимальность по Парето стратегий $(\mu_\lambda, \alpha_\lambda)$, найденных при решении расширенной задачи.

6. Заключение. В работе рассмотрена параметрическая модель оптимизации процесса передачи данных по ненадежному каналу связи при наличии механизма повторной отправки. Сеть передачи данных моделировалась в стационарном режиме на фиксированном уровне загрузки с помощью замкнутой СеМО, содержащей основной узел-передатчик

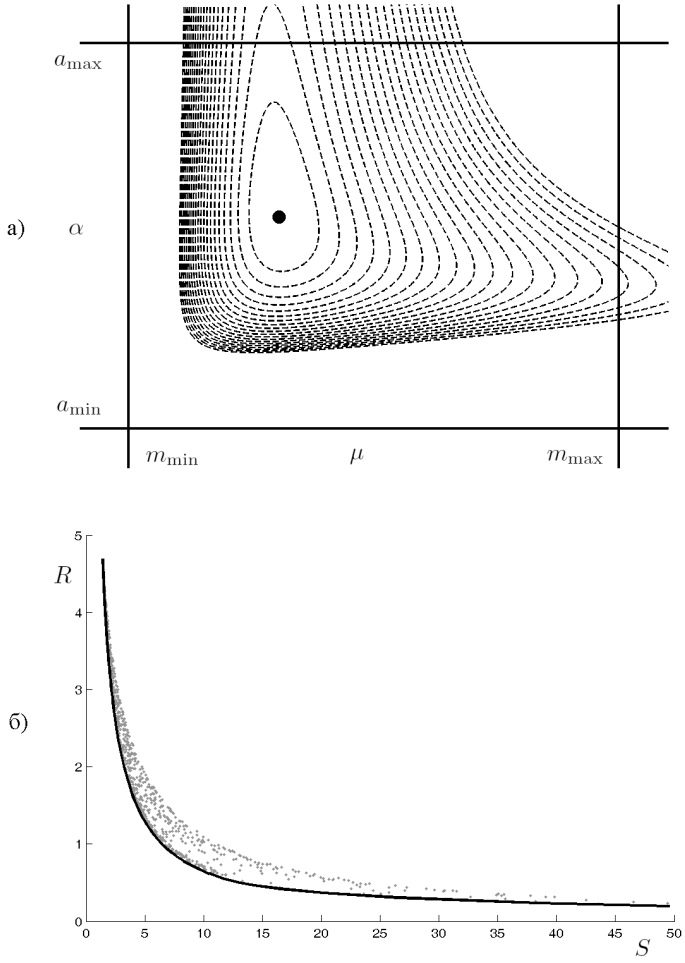

Рис. 5. Расширенная задача — а) и задача двукритериальной оптимизации — б)

в виде конечной очереди и вспомогательный узел для повторной отправки в виде многоканальной системы. В задаче на минимум среднего времени успешной передачи при наличии ограниченного потребления ресурсов определены оптимальные значения двух параметров сети: скорости передачи и интенсивности повторной отправки. Решена также задача о минимальном расходе ресурсов сети с учетом ограничения на время успешной передачи. Несмотря на то, что указанные задачи не являются задачами выпуклого программирования, множество Парето-оптимальных стратегий найдено методом множителей Лагранжа в результате минимизации расширенного функционала. Установлено, что при высокой загрузке сети достаточно использовать приближенные решения, основанные на оптимизации скорости передачи без учета интенсивности повторной отправки. Этот факт может быть использован при анализе и разработке реальных инфокоммуникационных сетей, предназначенных для обработки интенсивного пакетного трафика.

Литература

1. *Artalejo J.R.*, *Gomez-Corral A.* Retrial Queueing Systems: A Computational Approach // Springer. 2008. 318 p.
2. *Artalejo J.R.* Accessible bibliography on retrial queues: Progress in 2000-2009 // Mathematical and Computer Modelling. 2010. vol. 51. pp. 1071–1081.
3. *Phung-Duc T.* Retrial queueing models: A survey on theory and applications // Applied Stochastic Models in Business and Industry. 2019. vol. 35. 31 p.
4. *Степанов С.Н.*, *Степанов М.С.* Построение и анализ обобщенной модели контакт-центра // Автоматика и телемеханика. 2014. № 11. С. 55–69.
5. *Степанов С.Н.*, *Степанов М.С.* Алгоритмы оценки показателей пропускной способности обобщенной модели контакт-центра // Автоматика и телемеханика. 2016. № 7. С. 86–102.
6. *Adams R.* Active queue management: A survey // IEEE Communications Surveys & Tutorials. 2013. vol. 15. no. 3. pp. 1425–1476.
7. *Floyd S.*, *Jacobson V.* Random early detection gateways for congestion avoidance // IEEE/ACM Transactions on Networking. 1993. vol. 1. no. 4. pp. 397–413.
8. *Коновалов М.Г.*, *Разумчик Р.В.* Об управлении размером очереди в системе с одним сервером // Системы и средства информатики. 2017. Т. 27. № 4. С. 4–15.
9. *Konovalov M.G.*, *Razumchik R.V.* Comparison of two active queue management schemes through the M/D/1/N queue // Информатика и ее применения. 2018. Т. 12. № 4. pp. 9–15.
10. *Агаларов Я.М.*, *Агаларов М.Я.*, *Шоргин В.С.* Об оптимальном пороговом значении длины очереди в одной задаче максимизации дохода системы массового обслуживания типа M/G/1 // Информатика и ее применения. 2016. Т. 10. № 2. С. 70–79.
11. *Miller B.M.* Optimization of queuing system via stochastic control // Automatica. 2009. vol. 45. no. 6. pp. 1423–1430.
12. *Миллер А.Б.* Динамическое управление доступом и скоростью обслуживания при активных пользователях // Автоматика и телемеханика. 2010. № 9. С. 70–82.

13. *Солодянников Ю.В.* Управление и наблюдение для динамических сетей массового обслуживания. I // Автоматика и телемеханика. 2014. № 3. С. 14–45.
14. *Kuznetsov N.A.*, *Myasnikov D.V.*, *Semenikhin K.V.* Optimal control of data transmission in a mobile two-agent robotic system // Journal of Communications Technology and Electronics. 2016. vol. 61. no. 12. pp. 1456–1465.
15. *Kuznetsov N.A.*, *Myasnikov D.V.*, *Semenikhin K.V.* Optimization of two-phase queuing system and its application to the control of data transmission between two robotic agents // Journal of Communications Technology and Electronics. 2017. vol. 62, no. 12. pp. 1484–1498.
16. *Коган Я.А.*, *Липцер Р.Ш.*, *Смородинский А.В.* Гауссовская диффузионная аппроксимация марковских замкнутых моделей сетей связи ЭВМ // Проблемы передачи информации. 1986. Т. 22. № 1. С. 49–65.
17. *Schechner Z.*, *Yao D.* Decentralized control of service rates in a closed Jackson network // IEEE Transactions on Automatic Control. 1989. vol. 34. no. 2. pp. 236–240.
18. *Argon N.T.*, *Deng C.*, *Kulkarni V.G.* Optimal control of a single server in a finite-population queueing network // Queueing Systems. 2017. vol. 85. no. 1-2. pp. 149–172.
19. *Atar R.*, *Lev-Ari A.* Optimizing buffer size for the retrial queue: two state space collapse results in heavy traffic // Queueing Systems. 2018. vol. 90. no. 3-4. pp. 225–255.
20. *Кузнецов Н.А.*, *Семенихин К.В.* Анализ и оптимизация управляемой модели замкнутой сети массового обслуживания // Автоматика и телемеханика. URL: www.researchgate.net/publication/334126143 (дата обращения: 30.06.2019).
21. *Baccelli F.*, *Bremaud P.* Elements of Queueing Theory: Palm-Martingale Calculus and Stochastic Recurrences // Springer. 2003. 334 p.
22. *Иоффе А.Д.*, *Тихомиров В.М.* Теория экстремальных задач // М.: Наука. 1974. 480 с.

Кузнецов Николай Александрович — д-р техн. наук, академик РАН, профессор, советник РАН; заведующий лабораторией № 351 Института радиотехники и электроники им. В.А. Котельникова РАН (ИРЭ), заведующий кафедрой «Инфокоммуникационные системы и сети» Московского физико-технического института. Область научных интересов: управление инфокоммуникационными системами и сетями, теория рассинхронизованных систем, методы управления наблюдениями в стохастических системах, разработка алгоритмов идентификации для систем управления движущимися объектами, оптимизация транспортных потоков, обработка биомедицинской информации. Число научных публикаций — 224. kuznetsov@cplire.ru; ИРЭ, Моховая ул., д. 11, корп. 7, г. Москва, 125009, РФ; р.т. +7(495)6293005, факс +7(495)6293005.

Семенихин Константин Владимирович — д-р физ.-мат. наук, доцент; профессор кафедры «Теория вероятностей и компьютерное моделирование» Московского авиационного института (МАИ), старший научный сотрудник Института радиотехники и электроники им. В.А. Котельникова РАН. Область научных интересов: методы робастной статистики, минимаксная оптимизация, управление и оценивание в стохастических системах, оптимизация систем массового обслуживания. Число научных публикаций — 108. siemenkv@mail.ru; МАИ, Волоколамское шоссе, д. 4, г. Москва, 125993, РФ; р.т. +7(499)1584874, факс +7(499)1584560.

DOI 10.15622/sp.2019.18.4.809-830

N.A. KUZNETSOV, K.V. SEMENIKHIN

PARAMETRIC OPTIMIZATION OF PACKET TRANSMISSION WITH RESENDING PACKETS MECHANISM

Kuznetsov N.A., Semenikhin K.V. **Parametric Optimization of Packet Transmission with Resending Packets Mechanism.**

Abstract. The data transmission process is modelled by a Markov closed queuing network, which consists of two stations. The primary station describes the process of sending packets over a lossy channel by means of a finite and single-channel queue. The auxiliary station, being a multichannel queuing system, accumulates packets lost by the primary station and forwards them back for retrial. The transmission rate at the primary station and the retrial rate at the auxiliary station are in the specified ranges and subject to optimization in order to minimize time of successful delivery and amount of network resources used. The explicit expressions for these characteristics are derived in the steady-state mode in order to formulate the problem of bi-criterion optimization. The optimal policies are established in two scenarios: the first problem is to minimize the average time of successful transmission with limited resources; the second problem is to minimize the consumption of network resources within time constraint for successful transmission. The set of Pareto-optimal policies is obtained by solving the problem of minimization of the augmented functional. The quality characteristics of approximate solutions that do not take into account the service rate in the auxiliary system are analyzed.

Keywords: closed queuing network, birth-and-death process, optimization, transmission rate, retrial rate.

Kuznetsov Nikolay Alexandrovich — Dr. Sc. in Technology and Engineering, Professor, Academician (Russian Academy of Sciences); Head of Laboratory no. 351 in Kotel'nikov Institute of Radioengineering and Electronics (IRE), Head of Department «Infocommunication systems and networks» in Moscow Institute of Physics and Technology. Research interests: control of infocommunication systems and networks, theory of non-synchronous systems, methods of observation control in stochastic systems, development of identification algorithms for control systems of moving objects, optimization of traffic flows, biomedical data processing. The number of publications — 224. kuznetsov@cplire.ru; IRE, Mokhovaya, 11-7, Moscow, 125009, Russian Federation; office phone +7(495)6293005, fax +7(495)6293005.

Semenikhin Konstantin Vladimirovich — Dr. Sc. in Physics and Mathematics, Associate Professor; Professor of Department «Probability theory and computer modeling» in Moscow Aviation Institute (MAI), Senior Research Fellow in Kotel'nikov Institute of Radio Engineering and Electronics. Research interests: methods of robust statistics, minimax optimization, control and estimation in stochastic systems, optimization of queuing systems. The number of publications — 108. siemenkv@mail.ru; MAI, Volokolamskoye shosse, 4, Moscow, 125993, Russian Federation; office phone +7(499)1584874, fax +7(499)1584560.

References

1. Artalejo J.R., Gomez-Corral A. Retrial Queueing Systems: A Computational Approach. Springer. 2008. 318 p.
2. Artalejo J.R. Accessible bibliography on retrial queues: Progress in 2000-2009. *Mathematical and Computer Modelling*. 2010. vol. 51. pp. 1071–1081.

3. Phung-Duc T. Retrial queueing models: A survey on theory and applications. *Applied Stochastic Models in Business and Industry*. 2019. vol. 35. 31 p.
4. Stepanov S.N., Stepanov M.S. [Construction and analysis of a generalized contact center model]. *Avtomatika i Telemekhanika — Automation and Remote Control*. 2014. no. 11. pp. 55-69. (In Russ.).
5. Stepanov S.N., Stepanov M.S. [Algorithms for estimating throughput characteristics in a generalized call center model]. *Avtomatika i Telemekhanika — Automation and Remote Control*. 2016. no. 7. pp. 86-102. (In Russ.).
6. Adams R. Active queue management: A survey. *IEEE Communications Surveys & Tutorials*. 2013. vol. 15. no. 3. pp. 1425–1476.
7. Floyd S., Jacobson V. Random early detection gateways for congestion avoidance. *IEEE/ACM Transactions on Networking*. 1993. vol. 1. no. 4. pp. 397–413.
8. Konovalov M.G., Razumchik R.V. [Controlling queue size in a single server system]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics*. 2017. vol. 27. no. 4. pp. 4–15. (In Russ.).
9. Konovalov M.G., Razumchik R.V. Comparison of two active queue management schemes through the M/D/1/N queue. *Informatics and Applications*. 2018. vol. 12. no. 4. pp. 9–15.
10. Agalarov Ya.M., Agalarov M.Ya., Shorgin V.S. [About the optimal threshold of queue length in a particular problem of profit maximization in the M/G/1 queuing system]. *Informatika i ee primeneniya — Informatics and Applications*. 2016. vol. 10. no. 2. pp. 70–79. (In Russ.).
11. Miller B.M. Optimization of queuing system via stochastic control. *Automatica*. 2009. vol. 45. no. 6. pp. 1423–1430.
12. Miller A.B. [Using methods of stochastic control to prevent overloads in data transmission networks]. *Avtomatika i Telemekhanika — Automation and Remote Control*. 2010. no. 9. pp. 70–82. (In Russ.).
13. Solodyannikov Yu.V. Control and observation for dynamical queueing networks. I. *Avtomatika i Telemekhanika — Automation and Remote Control*. 2014. no. 3. pp. 14-45. (In Russ.).
14. Kuznetsov N.A., Myasnikov D.V., Semenikhin K.V. Optimal control of data transmission in a mobile two-agent robotic system. *Journal of Communications Technology and Electronics*. 2016. vol. 61. no. 12. pp. 1456–1465.
15. Kuznetsov N.A., Myasnikov D.V., Semenikhin K.V. Optimization of two-phase queuing system and its application to the control of data transmission between two robotic agents. *Journal of Communications Technology and Electronics*. 2017. vol. 62, no. 12. pp. 1484–1498.
16. Kogan A.Ya., Lipcer R.Sh., Smorodinskij A.V. Gaussian Diffusion Approximation of Closed Markov Models of Computer Networks. *Problemy peredachi informacii — Problems of Information Transmission*. 1986. vol. 22. no. 1. pp. 49–65. (In Russ.).
17. Schechner Z., Yao D. Decentralized control of service rates in a closed Jackson network. *IEEE Transactions on Automatic Control*. 1989. vol. 34. no. 2. pp. 236–240.
18. Argon N.T., Deng C., Kulkarni V.G. Optimal control of a single server in a finite-population queueing network. *Queueing Systems*. 2017. vol. 85. no. 1-2. pp. 149–172.
19. Atar R., Lev-Ari A. Optimizing buffer size for the retrial queue: two state space collapse results in heavy traffic. *Queueing Systems*. 2018. vol. 90. no. 3-4. pp. 225–255.
20. Kuznetsov N.A., Semenikhin K.V. [Analysis and optimization of the closed queuing network as a controllable system]. *Avtomatika i Telemekhanika — Automation and Remote Control*. Available at: www.researchgate.net/publication/334126143 (accessed 30.06.2019). (In Russ.).

21. Baccelli F., Bremaud P. Elements of Queueing Theory: Palm-Martingale Calculus and Stochastic Recurrences. Springer. 2003. 334 p.

22. Ioffe A.D., Tihomirov V.M. *Teoriya ekstremal'nyh zadach* [Theory of extremal problems]. M: Nauka, 1974. 480 p. (In Russ.).