

© 2021 г. А.С. МАНДЕЛЬ, д-р техн. наук (almandel@yandex.ru)
(Институт проблем управления им. В.А. Трапезникова РАН, Москва),
В.А. ЛАПТИН (straqker@bk.ru)
(Московский государственный университет им. М.В. Ломоносова)

ОПТИМАЛЬНОЕ УПРАВЛЕНИЕ СИСТЕМАМИ МАССОВОГО ОБСЛУЖИВАНИЯ С ПЕРЕКЛЮЧЕНИЕМ КАНАЛОВ ОБСЛУЖИВАНИЯ

Рассматривается задача оптимизации работы системы массового обслуживания, в которой число рабочих каналов обслуживания может управляемо изменяться в моменты контроля, отстоящие друг от друга на фиксированный временной шаг. Предполагается, что при переходе от шага к шагу интенсивность простейшего входящего потока изменяется в соответствии с некоторой однородной марковской цепью. Критерием выбора стратегии переключения каналов обслуживания является минимум суммарных средних затрат на многошаговом периоде планирования. Выявлена параметрическая структура оптимальной стратегии переключения каналов обслуживания.

Ключевые слова: управляемые системы массового обслуживания, марковский входящий поток, оптимизация, стратегии переключения каналов, параметризация стратегий.

DOI: 10.31857/S0005231021100093

1. Введение

Интерес к управляемым системам массового обслуживания (СМО) восходит к 70-м гг. XX в., когда была опубликована серия работ В.В. Рыкова. Отметим его фундаментальную публикацию [1], в которой приведена достаточно общая постановка задачи управления СМО и предложены подходы к ее решению. Оригинальным вкладом в теорию управляемых СМО стало исследование Ю.И. Неймарка [2], в котором изучались так называемые конфликтные системы, см. также развивающую эти результаты применительно к СМО публикацию [3].

К концу 80-х гг. интерес к управляемым системам массового обслуживания заметно вырос в связи с тем, что возник, начал и продолжает бурно развиваться такой класс прикладных систем как вычислительные и информационные сети и опирающиеся на них системы связи. Именно поэтому основные усилия специалистов в области теории массового обслуживания оказались направленными на решение задач управления, связанных именно с такими прикладными системами. Среди постановок задач и исследуемых моделей как у нас в стране, так и за ее рубежами рассматривались прежде всего

модели, в которых в качестве критериев выступали различные характеристики времен реакции на запрос [4], значения пропускных способностей [5, 6] и энергетической эффективности [7]. При этом решались вопросы оптимизации параметров [5, 7] и структуры этих систем [4, 6]. На эти темы имеется весьма значительное число публикаций, см., например, обзор [8].

В отличие от упомянутых работ в настоящей статье рассматривается задача оптимизации функционирования управляемой СМО по критерию минимума суммарных средних затрат за счет включения дополнительных или отключения лишних каналов обслуживания. Соответствующие модели весьма удобны для описания процессов, которые характерны для различных систем социально-экономической природы, включая производственно-торговые комплексы, системы продажи различного рода билетов, бронирования мест в гостиницах и многие другие. Настоящая статья развивает и обобщает постановки задач и методы их решения, которые были рассмотрены в предыдущих публикациях авторов настоящей статьи [9–14].

2. Постановка и алгоритм решения задачи

2.1. Содержательная постановка задачи

Рассматривается СМО с параллельно работающими каналами обслуживания, структурная схема которой представлена на рис. 1. При этом помимо работающих каналов обслуживания, которые на рис. 1 названы основными, имеются и резервные каналы обслуживания. В некоторые моменты времени (назовем их моментами контроля) любое доступное число резервных каналов может быть переведено в статус основных (включение) так же, как и любое число имеющихся основных каналов может быть переведено в статус резервных (выключение). Система рассматривается на конечном интервале времени — периоде планирования — продолжительности $T = N\tau$, где τ — длительность промежутка времени между соседними моментами контроля. Этот промежуток будет называться шагом. Таким образом, рассматриваемая СМО является системой с периодическим контролем, когда в равноудаленные моменты контроля можно включать дополнительные или отключать лишние

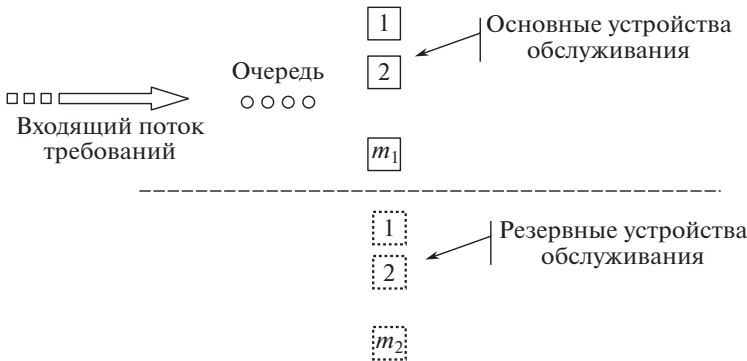


Рис. 1. Структурная схема исследуемой управляемой СМО.

каналы обслуживания. В предположении, что СМО должна выполнить все поступающие в нее требования, требуется минимизировать суммарные средние затраты на функционирование СМО в периоде планирования $[0, T]$. На каждом шаге в перечень затрат включаются затраты на содержание и эксплуатацию основных и резервных каналов обслуживания, затраты на решения о включении или отключении каналов (затраты на переключения) и затраты СМО на требования, пребывающие в очереди. При этом считается, что в СМО поступает простейший входящий поток, интенсивность которого $\lambda(t)$ на протяжении шага постоянна, а в моменты контроля претерпевает скачкообразные изменения, принимая конечное число k значений λ_i из дискретного множества $\Lambda = \{\lambda_i, i \in \overline{1, k}\}$.

2.2. Математическая постановка и решение задачи

Итак, рассматривается СМО, в которой число рабочих каналов обслуживания является управляемой величиной и может быть изменено в периодически (с шагом, равным единице) расположенные на оси времени моменты контроля за состоянием СМО. Задана матрица вероятностей перехода соответствующей однородной марковской цепи $\mathbf{P} = \|p_{ij}\|$, где p_{ij} — вероятность перехода (в момент контроля) от интенсивности λ_i , $i \in \overline{1, k}$, на предыдущем шаге к интенсивности λ_j , $j \in \overline{1, k}$, на следующем шаге.

Если на данном шаге интенсивность входящего потока равна λ_i , а интенсивность обслуживания на одном рабочем канале составляет μ , то, чтобы в рассматриваемой СМО на данном шаге мог установиться стационарный в вероятностном смысле режим функционирования, число основных рабочих каналов u в СМО должно выбираться, чтобы удовлетворять неравенству [4, 15]:

$$(1) \quad u \geq u_{\text{крит}}(\lambda_i) = \underline{u}_i = \left\lceil \frac{\lambda_i}{\mu} \right\rceil + 1,$$

где через $\lceil \bullet \rceil$ обозначена целая часть числа.

Рассмотрим состояние СМО в момент контроля за n шагов длительности τ до конца периода планирования. Будем считать, что в начальный момент времени случайная интенсивность входящего потока приняла на этом шаге значение λ_i , к этому моменту в системе имелось m основных каналов обслуживания и принимается решение о том, что в СМО должно работать u основных каналов обслуживания.

Запишем выражение для суммарных средних затрат $C^{(1)}(\lambda_i, m, u)$ на первом шаге процесса управления (стартующем за n шагов до конца периода планирования), считая, что затраты на эксплуатацию резервных устройств обслуживания равны нулю. В силу сделанных выше замечаний эти затраты могут быть представлены в виде

$$(2) \quad C^{(1)}(\lambda_i, m, u) = C_{\text{эспл}}(u) + C_{\text{очереди}}(\lambda_i, u) + C_{\text{переключ}}(\lambda_i, m, u).$$

В (2) $C_{\text{экспл}}(u)$ – это одношаговые затраты на эксплуатацию u основных устройств ($C_{\text{экспл}}(u) = c_1 u$, где c_1 – цена эксплуатации основного устройства обслуживания на протяжении одного шага); $C_{\text{очереди}}(\lambda_i, u)$ – одношаговые затраты на очередь в стационарном режиме, которые будем считать пропорциональными средней длине очереди $\bar{L}_{\text{очереди}}(\lambda_i, u)$ как функции от интенсивности входящего потока λ_i и управляющего решения о числе u основных каналов; $C_{\text{переключ}}(\lambda_i, m, u)$ – затраты на переключения, которые зависят от интенсивности входящего потока λ_i , от начального (до выполнения операции переключения) числа основных каналов m и от управляющего решения о числе основных каналов u . Будем считать, что эта компонента затрат может быть представлена в виде

$$(3) \quad C_{\text{переключ}} = \begin{cases} A_1, & \text{если } u > m; \\ 0, & \text{если } u = m; \\ A_2 + c_2(m - u), & \text{если } u < m, \end{cases}$$

где A_1 – фиксированная цена принятия решения о подключении новых рабочих устройств (“включение”), A_2 – это фиксированная цена принятия решения об отключении новых рабочих устройств (“выключение”), а c_2 – стоимость отключения одного рабочего устройства обслуживания ($c_1 > c_2$). В формуле (2) осталось конкретизировать только значение $C_{\text{очереди}}(\lambda_i, u)$. Будем считать, что пребывание каждого требования в очереди на протяжении одного шага обходится в сумму d , т.е. $C_{\text{очереди}}(\lambda_i, u) = d\bar{L}_{\text{очереди}}(\lambda_i, u)$. Пользуясь классическими результатами [4, 15], можно записать:

$$(4) \quad \bar{L}_{\text{очереди}} = \left[\sum_{k=1}^{u-1} \frac{(u\rho_i)^k}{k!} + \frac{(u\rho_i)^u}{u!(1-\rho_i)} \right]^{-1} \frac{(u\rho_i)^u \rho_i}{u!(1-\rho_i)^2},$$

где $\rho_i = \frac{\lambda_i}{u\mu}$.

Собираясь воспользоваться для решения проблемы методом дискретного динамического программирования, введем следующий промежуточный функционал, а именно функционал $C_n^*(\lambda_i, m)$, описывающий минимально возможное (т.е. при использовании еще не найденной оптимальной стратегии переключения каналов) значение средних суммарных затрат на последних n шагах процесса управления, когда в начале первого шага текущее число основных рабочих каналов равно m , а интенсивность простейшего входящего потока принимает значение λ_i . В этих обозначениях окончательное решение проблемы переключений на протяжении всего периода планирования будет описываться по уровню минимально возможных средних затрат величиной $C_N^*(\lambda, m)$, где λ и m – текущие значения интенсивности входящего потока и числа основных каналов обслуживания, с которых стартует СМО в периоде планирования $[0, T]$.

Теперь нетрудно записать систему уравнений дискретного динамического программирования, которая и является решением проблемы управления

процессом переключения каналов:

$$(5) \quad C_1^*(\lambda_i, m) = \min_{u \geq \underline{u}_i} C^{(1)}(\lambda_i, m, u),$$

$$(6) \quad C_n^*(\lambda_i, m) = \min_{u \geq \underline{u}_i} \left\{ C^{(1)}(\lambda_i, m, u) + \alpha \sum_{j=1}^l p_{ij} C_{n-1}^*(\lambda_j, u) \right\}, \quad n \in \overline{2, N},$$

где значение $\underline{u}_i$ задается формулой (1), значение $C^{(1)}(\lambda_i, m, u)$ – формулой (2), $i \in \overline{1, k}$, а α – коэффициент дисконтирования, $0 \leq \alpha \leq 1$.

Преобразуем уравнения (5) и (6), введя вспомогательные функции $G_{\text{вкл}}^{(1)}(\lambda_i, u)$, $G_{\text{выкл}}^{(1)}(\lambda_i, u)$, $G_{\text{вкл}}^{(n)}(\lambda_i, u)$, $G_{\text{выкл}}^{(n)}(\lambda_i, u)$:

$$(7) \quad G_{\text{вкл}}^{(1)}(\lambda_i, u) = uc_1 + d\bar{L}_{\text{очереди}}(\lambda_i, u),$$

$$(8) \quad G_{\text{выкл}}^{(1)}(\lambda_i, u) = u(c_1 - c_2) + d\bar{L}_{\text{очереди}}(\lambda_i, u),$$

$$(9) \quad G_{\text{вкл}}^{(n)}(\lambda_i, u) = uc_1 + d\bar{L}_{\text{очереди}}(\lambda_i, u) + \alpha \sum_{j=1}^l p_{ij} C_{n-1}^*(\lambda_j, u), \quad n = 2, 3, \dots,$$

$$(10) \quad G_{\text{выкл}}^{(n)}(\lambda_i, u) = u(c_1 - c_2) + d\bar{L}_{\text{очереди}}(\lambda_i, u) + \alpha \sum_{j=1}^l p_{ij} C_{n-1}^*(\lambda_j, u),$$

$$n = 2, 3, \dots$$

Теперь уравнения (5) и (6) можно переписать в единообразном виде:

$$(11) \quad C_n^*(\lambda_i, m) =$$

$$= \min \begin{cases} -c_1 m + \min \begin{cases} A_1 + \min_{u \geq \underline{u}_i} G_{\text{вкл}}^{(n)}(\lambda_i, u), \\ G_{\text{вкл}}^{(n)}(\lambda_i, m), \end{cases} \\ -(c_1 - c_2)m + \min \begin{cases} A_2 + \min_{u \geq \underline{u}_i} G_{\text{выкл}}^{(n)}(\lambda_i, u), \\ G_{\text{выкл}}^{(n)}(\lambda_i, m), \end{cases} \end{cases} \quad n \in \overline{1, N}.$$

Поясним запись уравнений дискретного динамического программирования (5)–(6) в форме (11). Очевидно, что появление нижних индексов “включ” и “выключ” в функциях с общим обозначением G соответствует принятию решений о включении дополнительных резервных каналов или, наоборот, отключении части рабочих каналов. Решениям о включении и выключении соответствуют первая (включение) и третья (выключение) строки в крайней правой части формулы (11). В то же время вторая и четвертая строки в крайней правой части формулы (11) отвечают за принятие решений об отказе от каких-либо переключений. При этом сама оптимальная стратегия в новой переформулированной записи уравнений (11) пока еще не определена.

Прежде чем перейти к обсуждению решений уравнений (5)–(6) или (11), отметим, что поставленная задача в “близорукой” постановке ($n = 1$), которая сводится к решению задачи минимизации (2), подробно исследовалась в [13], где обсуждение опиралось на качественное использование аналогий с математическими задачами теории управления запасами и производством [16]. Опора на эту аналогию, в конечном счете, привела к тому, что для управляемых СМО удалось в несколько видоизмененной форме воспроизвести некоторые классические свойства стратегий оптимального управления запасами, например их параметрический характер (наличие двух критических уровней R и r). При этом упомянутые аналогии выстраивались между задачами управления запасами с непрерывными переменными состояниями, тогда как рассматриваемые задачи теории массового обслуживания по одной из переменных (значение интенсивности входящего потока) являются дискретными. Чтобы восполнить этот пробел, в разделе 3 настоящей статьи приводятся некоторые развивающие известные и новые результаты по математическим инструментам изучения “похожих” на выпуклые дискретных задач оптимизации.

3. А-выпуклость в дискретных задачах оптимизации

Рассматривается функция $g(i)$ от переменной i , принимающей значения из ряда натуральных чисел ($i = 1, 2, 3, \dots$). Напомним несколько известных понятий, которые тщательно изучены для случая вещественных (непрерывных) переменных и функций, а для случая дискретных переменных используются несколько реже.

Определение 1. Функция $g(i)$ называется выпуклой (вниз), если для всех натуральных чисел i и j

$$(12) \quad g(i+j) - g(i) - \Delta^{(1)}g(i) \times j \geq 0,$$

где $\Delta^{(1)}g(i)$ – первая разность функции $g(i)$ в точке i : $\Delta^{(1)}g(i) = g(i) - g(i-1)$.

Необходимым и достаточным условием выпуклости (вниз) функции является выполнение неравенства $\Delta^{(2)}g(i) \geq 0$, где $\Delta^{(2)}g(i)$ – вторая разность функции $g(i)$ в точке i . Действительно, пусть j из определения 1 равно единице, тогда формула (11) записывается как $\Delta^{(1)}g(i+1) \geq \Delta^{(1)}g(i)$, т.е. первая разность $g(i)$ оказывается монотонно возрастающей функцией, т.е. $\Delta^{(2)}g(i) \geq 0$. Аналогично доказывается и достаточность.

Определение 2. Функция $g(i)$ называется А-выпуклой¹ ($A \geq 0$), если для всех натуральных чисел i и j :

$$(13) \quad A + g(i+j) - g(i) - \Delta^{(1)}g(i) \times j \geq 0.$$

¹ Понятие А-выпуклости было предложено Г. Скарфом [17] для анализа свойств оптимальных стратегий принятия решений (управления) в задачах логистики запасов при наличии фиксированных цен поставки, не зависящих от размера поставки, с добавлением к ним сумм, обусловленных размером партии поставки.

Для функций от дискретных переменных сохраняются все свойства A -выпуклых функций, отмеченные Гербертом Скарфом [8], а именно:

Свойство 1. Если функция $g(i)$ является A -выпуклой, то при любом натуральном j функция $g(i + j)$ также A -выпукла.

Свойство 2. Если функция $g_1(i)$ является A_1 -выпуклой, а функция $g_2(i)$ — A_2 -выпуклой, то при любых $\theta_1, \theta_2 > 0$ функция $g(i) = \theta_1 g_1(i) + \theta_2 g_2(i)$ также является $(\theta_1 A_1 + \theta_2 A_2)$ -выпуклой.

Свойство 3. Если функция $g(i)$ является A_1 -выпуклой, то она и A_2 -выпукла для любого $A_2 > A_1$.

Свойство 4. Пусть i — случайная величина с распределением $\{p_i\}_1^k$, $p_i > 0$, $\sum_{i=1}^k p_i = 1$, и пусть функция $g(i)$ является A -выпуклой. Тогда функция $\alpha \sum_{i=1}^k p_i g(i)$, где $\alpha \geq 0$, также A -выпукла.

Отметим, что доказательство устанавливаемых в разделе 4 настоящей статьи фактов опирается на аналогию между задачами теории управления запасами и задачами теории массового обслуживания. Первичный анализ этих аналогий был выполнен в публикации [16], в которой была рассмотрена задача управления запасами, названная авторами “фантазийной”². Тем не менее нельзя не признать, что для рассматриваемого класса задач (как “фантазийной” из теории управления запасами³, так и задачи о переключении каналов), была продемонстрирована некоторая неготовность, неприспособленность существующих теории оптимизации и теории выпуклости к решению этих задач. Дело в том, что в при математическом описании рассматриваемого класса задач используется не одно, а два разных видения оптимальности (критерия оптимальности). Говоря на языке, адекватном рассматриваемому в данной статье классу задач, это *оптимальность при включении* каналов (при подаче заказов в теории управления запасами), если перемещаться по оси целочисленной переменной *и слева-направо*, и *оптимальность при отключении* каналов (возвращении товара в теории управления запасами), если перемещаться по оси переменной *и справа-налево*. Используем это замечание, чтобы сформулировать новые определения.

Определение 3. Функция $g(i)$ называется A -выпуклой слева ($A \geq 0$) в интервале $(-\infty, s]$, если для некоторого натурального числа s , которое обладает тем свойством, что оно находится правее точки абсолютного минимума функции $g(i)$, и для всех натуральных чисел i и j , удовлетворяющих неравенствам $i, i + j \leq s$, справедливо неравенство (13).

Определение 4. Функция $g(i)$ называется A -выпуклой справа ($A \geq 0$) в интервале $[r, \infty)$, если для некоторого натурального числа r , которое обладает тем свойством, что оно находится левее точки абсолютного мини-

² Фантастическим в [16] было предположение о том, что склад может не только подавать заказы на пополнение запасов, но и возвращать товар его поставщикам.

³ В настоящее время подвергнутый критике на ряде семинаров термин “фантазийная задача” заменен на новый — “задача управления запасами с возвратами”.

му функции $g(i)$, и для всех натуральных чисел i и j , удовлетворяющих неравенствам $p \leq i, i + j$, справедливо неравенство (13).

4. Исследование свойств оптимальных стратегий переключения каналов

Чтобы воспользоваться введенными выше понятиями A -выпуклости, а также A -выпуклости слева и справа для выявления свойств оптимальной стратегии переключения каналов, убедимся в том, что для введенных выше функционалов (5), (6) и (11) эти свойства имеют место. Рассмотрим сначала «близорукий» случай, когда число шагов $n = 1$.

4.1. Свойства «близоруких» стратегий переключения каналов

Рассмотрим одношаговую задачу переключения каналов, которая представляет собой задачу, описываемую уравнением (5) или уравнением (11) при $n = 1$. В этом случае целевой функционал состоит в минимизации функции (7) при включении каналов или функции (8) при выключении каналов. В обоих случаях эти функционалы представляют собой сумму возрастающей линейной функции от назначаемого к включению числа каналов u и взвешенной с коэффициентом d средней длины очереди $\bar{L}_{\text{очереди}}(\lambda_i, u)$. Изучим свойства характеристики $\bar{L}_{\text{очереди}}(\lambda_i, u)$ как функции переменной u .

Из чисто вероятностных соображений следует, что функция $\bar{L}_{\text{очереди}}(\lambda_i, u)$, как функция от u , является монотонно убывающей функцией u , при этом такой, что скорость этого убывания с ростом u также монотонно убывает. При этом очевидно, что для любого значения λ_i предел $\lim_{u \rightarrow \infty} \bar{L}_{\text{очереди}}(\lambda_i, u) = 0$, и вполне понятно, что по мере роста числа каналов от минимально возможного значения $\underline{u}_i$ величина $\bar{L}_{\text{очереди}}(\lambda_i, u)$ уменьшается, причем с ростом u это сокращение средней длины очереди $\bar{L}_{\text{очереди}}(\lambda_i, u)$ само становится все меньше, поскольку введение еще одного дополнительного рабочего канала на фоне большего числа рабочих каналов приносит выигрыш, «доля» которого $1/(u + 1)$ постоянно снижается. Однако оказывается, что строгое доказательство этого факта требует довольно утомительных выкладок (по крайней мере авторам статьи не удалось технически упростить соответствующие доказательства).

Изменим обозначения, сняв индекс i состояния СМО и заменив величину ρ_i из формулы (4) величиной $\rho_0 = \lambda/\mu$. В результате формула (4) приобретает вид

$$(14) \quad \bar{L}_{\text{очереди}}(\lambda, u) = \left[\sum_{k=1}^{u-1} \frac{(\rho_0)^k}{k!} + \frac{(\rho_0)^u}{(u-1)!(u-\rho_0)} \right]^{-1} \frac{(\rho_0)^{u+1}}{(u-1)!(u-\rho_0)^2}.$$

При $u - 1 \geq u_{\text{крит}}(\lambda) = \underline{u} = [\rho_0] + 1$ рассмотрим первую разность функции $\bar{L}_{\text{очереди}}(\lambda_i, u)$, как функции от u :

$$(15) \quad \Delta \bar{L}_{\text{очереди}}(\lambda, u) = \bar{L}_{\text{очереди}}(\lambda, u) - \bar{L}_{\text{очереди}}(\lambda, u - 1).$$

Утверждение 1. При $u - 1 \geq u_{\text{крит}}(\lambda) = \underline{u} = [\rho_0] + 1$ подсчитываемая по формуле (15) первая разность $\bar{L}_{\text{очереди}}(\lambda, u)$ отрицательна.

Доказательство. Используя формулы (14) и (15) и опуская промежуточные выкладки, запишем:

$$\begin{aligned} \Delta \bar{L}_{\text{очереди}}(\lambda, u) &= \\ &= \left[\sum_{k=1}^{u-1} \frac{(\rho_0)^k}{k!} + \frac{(\rho_0)^u}{(u-1)!(u-\rho_0)} \right]^{-1} \frac{(\rho_0)^{u+1}}{(u-1)!(u-\rho_0)^2} - \\ &- \left[\sum_{k=1}^{u-2} \frac{(\rho_0)^k}{k!} + \frac{(\rho_0)^u}{(u-2)!(u-1-\rho_0)} \right]^{-1} \frac{(\rho_0)^u}{(u-2)!(u-1-\rho_0)^2} = \\ &= -C_1(\rho_0, u) + C_2(\rho_0, u) [\rho_0(u-1-\rho_0)^2 - (u-1)(u-\rho_0)^2] \sum_{k=1}^{u-2} \frac{(\rho_0)^k}{k!}. \end{aligned}$$

В последнем выражении величины $C_1(\rho_0, u)$ и $C_2(\rho_0, u)$ положительны, а для множителя в квадратных скобках можно записать:

$$\begin{aligned} \rho_0(u-1-\rho_0)^2 - (u-1)(u-\rho_0)^2 &< \rho_0(u-\rho_0)^2 - (u-1)(u-\rho_0)^2 = \\ &= (\rho_0 - u + 1)(u - \rho_0)^2 < 0. \end{aligned}$$

Последнее неравенство записано в силу первого условия в утверждении 1. Из этих выкладок следует утверждение 1.

Введем вторую разность функции $\bar{L}_{\text{очереди}}(\lambda, u)$, как функция от u :

$$\begin{aligned} (16) \quad \Delta^{(2)} \bar{L}_{\text{очереди}}(\lambda, u) &= \Delta \bar{L}_{\text{очереди}}(\lambda, u) - \Delta \bar{L}_{\text{очереди}}(\lambda, u-1) = \\ &= \bar{L}_{\text{очереди}}(\lambda, u) - 2\bar{L}_{\text{очереди}}(\lambda, u-1) - \bar{L}_{\text{очереди}}(\lambda, u-2). \end{aligned}$$

Выполнив аналогичные выкладки, можно доказать следующее утверждение.

Утверждение 2. При $u - 2 \geq u_{\text{крит}}(\lambda) = \underline{u} = [\rho_0] + 1$ рассчитываемая по формуле (16) вторая разность функции $\bar{L}_{\text{очереди}}(\lambda, u)$ положительна.

Из утверждений 1 и 2 следует, что средняя длина очереди $\bar{L}_{\text{очереди}}(\lambda, u)$ является выпуклой вниз функцией переменной u . Из последнего утверждения и свойств 2 и 3 вытекает, что функции $G_{\text{включ}}^{(1)}(\lambda, u)$ из формулы (7) и $G_{\text{выключ}}^{(1)}(\lambda, u)$ из формулы (8) выпуклы вниз. Последнее в силу формул (5) и (11) и по аналогии с рассуждениями при доказательстве факта A -выпуклости в главе 7 монографии [18], казалось бы, должно было привести к выводу о том, что функция $C_1^*(\lambda, m)$ является A -выпуклой. Однако такой вывод оказывается поспешным, поскольку структура функции $C_1^*(\lambda, m)$ более сложна, чем те конструкции, которые были рассмотрены в классической публикации по теории управления запасами и производством [17]: она “склеена” из функции средних затрат при включении дополнительных рабочих каналов (через функцию $G_{\text{включ}}^{(1)}(\lambda, u)$) и функции средних затрат при отключении части

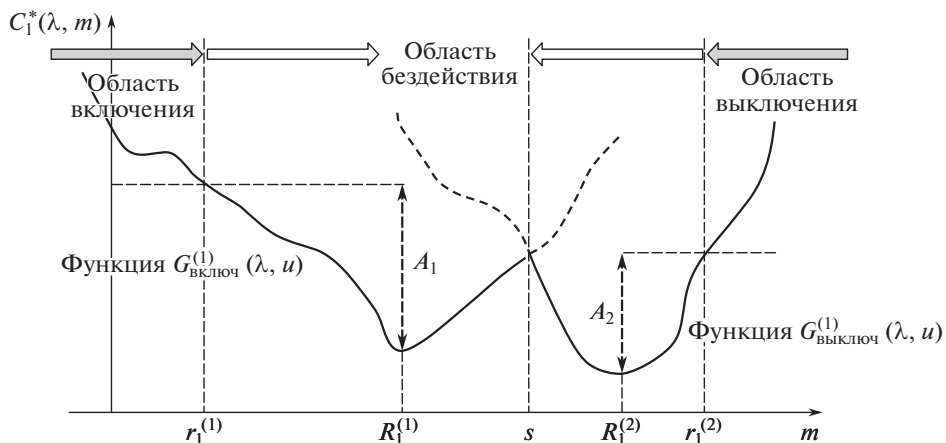


Рис. 2. Вид функции $C_1^*(\lambda, m)$.

рабочих каналов (через функцию $G_{\text{выключ}}^{(1)}(\lambda, u)$). Чтобы окончательно разобраться в характере этой “склейки”, выделим точки абсолютных минимумов функций $G_{\text{включ}}^{(1)}(\lambda, u)$ и $G_{\text{выключ}}^{(1)}(\lambda, u)$, обозначив их через $R_1^{(1)}$ и $R_1^{(2)}$ соответственно. Из формул (7) и (8) очевидно, что $R_1^{(1)} < R_1^{(2)}$. Таким образом, в силу формулы (11) для случая $n = 1$ график функции $C_1^*(\lambda, m)$ может быть представлен таким, как это показано на рис. 2.

Из рис. 2 видно, что в силу упорядочения точек абсолютных минимумов $R_1^{(1)}$ и $R_1^{(2)}$ функций $G_{\text{включ}}^{(1)}(\lambda, u)$ и $G_{\text{выключ}}^{(1)}(\lambda, u)$ существует точка s пересечения соответствующих фрагментов графиков (пунктиром на рис. 2 показаны нереализуемые части графиков). Эта точка играет роль точки s и, одновременно, точки p из определений 3 и 4. Таким образом, становится понятным, что до точки s функция $C_1^*(\lambda, m)$ является A_1 -выпуклой слева, а после точки s — A_2 -выпуклой справа. Тогда, следуя логике формирования оптимальных двухуровневых стратегий в теории управления запасами [17, 18], оптимальное правило переключения каналов может быть формализовано в виде:

$$(17) \quad u = \begin{cases} R_1^{(1)}, & \text{если } m \leq r_1^{(1)} (\text{включение}), \\ m, & \text{если } r_1^{(1)} < m \leq r_1^{(2)}, \\ R_1^{(2)}, & \text{если } m \geq r_1^{(2)} (\text{отключение}). \end{cases}$$

При этом параметры $r_1^{(1)}$ и $r_1^{(2)}$ в формуле (17)⁴ представляют собой решения следующих уравнений (см. рис. 2):

$$(18) \quad A_1 + G_{\text{включ}}^{(1)}(\lambda, R_1^{(1)}) = G_{\text{включ}}^{(1)}(\lambda, r_1^{(1)}) \quad \text{для } r_1^{(1)} < R_1^{(1)},$$

$$(19) \quad A_2 + G_{\text{выключ}}^{(1)}(\lambda, R_1^{(2)}) = G_{\text{выключ}}^{(1)}(\lambda, r_1^{(2)}) \quad \text{для } r_1^{(2)} < R_1^{(2)}.$$

⁴ Применение формулы (17) связано с некоторыми оговорками, суть которых заключается в том, что в некоторых случаях, перечисленных в [13], параметр $R_{1,i}^{(2)}$ назначается равным $\underline{u}_i$. Имеются и другие нюансы.

Осталось заметить, что, на самом деле, четыре параметра “близорукой” стратегии управления запасами должны быть снабжены индексом i состояния, в котором СМО находится на первом шаге, т.е. фактически их надо обозначать как $r_{1,i}^{(1)}$, $R_{1,i}^{(1)}$, $r_{1,i}^{(2)}$ и $R_{1,i}^{(2)}$, рассчитывая эти параметры для всех $i \in \overline{1, k}$.

4.2. Многошаговые динамические задачи о переключении каналов

Вернемся к решению многошаговой задачи управления переключениями каналов, которая описывается уравнениями (5)–(6) или (11). В этом случае функции $G_{\text{вкл}}^{(1)}(\lambda, u)$ и $G_{\text{выкл}}^{(1)}(\lambda, u)$ уступают место функциям $G_{\text{вкл}}^{(n)}(\lambda, u)$ и $G_{\text{выкл}}^{(n)}(\lambda, u)$, $n = 2, 3, \dots$, задаваемым формулами (9) и (10). Повторим запись целевого функционала в уравнении (6):

$$C_n^*(\lambda_i, m) = \min_{u \geq u_i} \left\{ C^{(1)}(\lambda_i, m, u) + \alpha \sum_{j=1}^l p_{ij} C_{n-1}^*(\lambda_j, u) \right\}.$$

Выражение в фигурных скобках в правой части этого уравнения, независимо от того, какая альтернатива оценивается (включение или отключение), содержит не меняющийся при смене альтернативы член $\alpha \sum_{j=1}^l p_{ij} C_{n-1}^*(\lambda_j, u)$. Одношаговый добавок $C^{(1)}(\lambda_i, m, u)$, являющийся, как показано в подразделе 4.1, выпуклой вниз функцией от u , наоборот, варьирует при смене альтернативы. Из этого можно сделать вывод, что есть все основания считать, что выражение в фигурных скобках в правой части формулы (6) (также двухвариантное) является одновременно A_1 -выпуклым слева и A_2 -выпуклым справа. Вопрос только в том, как будут упорядочены точки абсолютного минимума функций $G_{\text{вкл}}^{(n)}(\lambda, u)$ и $G_{\text{выкл}}^{(n)}(\lambda, u)$, которые обозначим как $R_{n,i}^{(1)}$ и $R_{n,i}^{(2)}$.

Утверждение 3. Для любого n справедливо неравенство $R_{n,i}^{(1)} < R_{n,i}^{(2)}$.

Доказательство. Используя формулы (9) и (10), которые можно переписать в виде

$$G_{\text{вкл}}^{(n)}(\lambda, u) = uc_1 + \beta^{(n)}(\lambda_i, u) \quad \text{и} \quad G_{\text{выкл}}^{(n)}(\lambda, u) = u(c_1 - c_2) + \beta^{(n)}(\lambda_i, u),$$

устанавливаем искомое неравенство для любого n .

Теперь выскажем предположение математической индукции о том, что это для любого номера n функции $G_{\text{вкл}}^{(n)}(\lambda, u)$ и $C_n^*(\lambda, m)$ являются A_1 -выпуклыми слева и A_2 -выпуклыми справа. Дальнейшее продвижение заключается в том, что следует установить факт A_1 -выпуклости слева и A_2 -выпуклости справа функции $G_{\text{вкл}}^{(n+1)}(\lambda, u)$. Тогда из уравнения (11) и утверждения 3 будет следовать, что A_1 -выпуклой слева и A_2 -выпуклой справа будет функция $C_{n+1}^*(\lambda, m)$. В свою очередь из факта A_1 -выпуклости слева и A_2 -выпуклости справа функции $G_{\text{вкл}}^{(n+1)}(\lambda, u)$ вытекает (по аналогии с подразделом 4.1), что

оптимальная стратегия переключения каналов будет определяться правилом:

$$(20) \quad u = \begin{cases} R_{n+1,i}^{(1)}, & \text{если } m \leq r_{n+1,i}^{(1)} \text{ (включение)}, \\ m, & \text{если } r_{n+1,i}^{(1)} < m \leq r_{n+1,i}^{(2)}, \\ R_{n+1,i}^{(2)}, & \text{если } m \geq r_{n+1,i}^{(2)} \text{ (отключение)}. \end{cases}$$

При этом параметры $r_{n+1,i}^{(1)}$ и $r_{n+1,i}^{(2)}$ в формуле (20) представляют собой решения следующих уравнений (см. рис. 2):

$$(21) \quad A_1 + G_{\text{включ}}^{(n+1)}(\lambda, R_{n+1,i}^{(1)}) = G_{\text{включ}}^{(n+1)}(\lambda, r_{n+1,i}^{(1)}) \quad \text{для } r_{n+1,i}^{(1)} < R_{n+1,i}^{(1)},$$

$$(22) \quad A_2 + G_{\text{выключ}}^{(n+1)}(\lambda, R_{n+1,i}^{(2)}) = G_{\text{выключ}}^{(n+1)}(\lambda, r_{n+1,i}^{(2)}) \quad \text{для } r_{n+1,i}^{(2)} < R_{n+1,i}^{(2)}.$$

Нетрудно видеть, что доказательство сформулированного выше утверждения математической индукции идейно повторяет доказательство аналогичного утверждения (для двухуровневых стратегий) в теории управления запасами [18] с поправкой на дискретность и отмеченную выше альтернативность (включение и выключение каналов) целевого функционала.

Как отмечалось выше, немаловажным фактом, который позволяет не опасаться “подводных камней” в форме скользящих режимов, является упорядочение точек абсолютного минимума альтернативных функционалов в форме неравенств $R_{n,i}^{(1)} < R_{n,i}^{(2)}$, $n \in \overline{1, N}$, $i \in \overline{1, k}$.

5. Заключение

Рассмотрена многошаговая задача оптимального переключения каналов в многолинейной системе массового обслуживания. Исследованы стратегии переключения каналов в дискретном времени при Марковском описании процесса изменения интенсивности входящего потока. Критерием для выбора стратегий переключения является минимизация суммарных средних затрат на многошаговом периоде планирования. Предполагается, что процедура переключения (включения или отключения) каналов приводит к расходам, которые состоят из фиксированных платежей и затрат, величина которых зависит от числа включаемых или отключаемых рабочих каналов. Доказано, что оптимальные стратегии переключения каналов оказываются параметрическими и на каждом шаге оптимальное решение задачи о переключениях зависит только от четырех характеристических параметров. Указаны способы расчета этих параметров.

СПИСОК ЛИТЕРАТУРЫ

1. Рыков В.В. Управляемые системы массового обслуживания // Теория вероятностей. Математическая статистика. Теоретическая кибернетика. 1975. Т. 12. С. 43–153.
2. Неймарк Ю.И. Динамические системы и управляемые процессы. М.: Наука, 1978.

3. Голышева Н.М., Федоткин М.А. Циклическое управление конфликтными потоками в условиях гибели и рождения очередей критических размеров // *АиТ*. 1990. № 4. С. 68–75.
Golyшева N.M., Fedotkin M.M. A Conflict Flows Cyclic Control in Conditions of Critical Queues Deaths and Births // Autom. Remote Control. 1990. V. 51. No. 4. P. 479–484.
4. Вишневецкий В.М. Теоретические основы проектирования компьютерных сетей. М.: Техносфера, 2003.
5. Захаров П.П. Разработка автоматизированного программного комплекса для исследования качества и эффективности функционирования моделей технических систем и управляемых систем массового обслуживания. Дисс. канд. техн. наук. М.: МИЭМ, 2006.
6. Ivanov R., Mukhtarov A., Pershin O. A Problem of Optimal Location of Given Set of Base Stations in Wireless Networks with Linear Topology // *Distributed Computer and Communication Networks. 22nd Int. Conf., DCCN 2019, Moscow, Russia, September 23–27, 2019, Revised Selected Papers. Springer, 2019. P. 53–64.*
7. Галкина О.С., Андреев С.Д., Кучерявый Е.А. Оптимальное управление мощностью передачи устройства в гетерогенных сетях связи. // Матер. 21-й Междунар. науч. конф. “Распределенные компьютерные и телекоммуникационные сети: управление, вычисление, связь” (DCCN-2018, Москва). М.: ИПУ РАН, 2018. С. 79–83.
8. Вишневецкий В.М., Дудин А.Н. Системы массового обслуживания с коррелированными входными потоками и их применение для моделирования телекоммуникационных сетей // *АиТ*. 2017. № 8. С. 3–59.
Vishnevskii V.M., Dudin A.N. Queueing Systems with Correlated Arrival Flows and Their Applications to Modeling Telecommunication Networks // Autom. Remote Control. 2017. V. 74. No. 8. P. 1361–1403.
9. Mandel A. Econometric Models of Controllable Multiple Queuing Systems // *Proc. 18th Int. Conf. “Distributed Computer and Communication Networks” (DCCN 2015, Moscow, Russia). Geneva: Springer, 2016. P. 296–304.*
10. Мандель А.С., Барладян И.И., Токмакова А.Б. Многолинейная СМО с изменением числа рабочих каналов: стационарный случай. Ч. 1 // Материалы 18-й Международной научной конференции “Распределенные компьютерные и телекоммуникационные сети: управление, вычисление, связь” (DCCN-2015, Москва). М.: ИПУ РАН, 2015. С. 345–354.
11. Мандель А.С., Махукова В.В. Многолинейная СМО с изменением числа рабочих каналов: нестационарный случай. Ч. 2 // Матер. 18-й Междунар. науч. конф. “Распределенные компьютерные и телекоммуникационные сети: управление, вычисление, связь” (DCCN-2015, Москва). М.: ИПУ РАН, 2015. С. 355–361.
12. Mandel A., Bakulin K. Models of Controllable Multiple Queuing Systems for Channel Switching Myopic Strategies // *Proc. 20th Int. Conf. “Distributed Computer and Communication Networks” (DCCN 2017, Moscow, Russia). M.: Техносфера, 2017. С. 534–542.*
13. Mandel A., Laptin V. Myopic Channel Switching Strategies for Stationary Mode: Threshold Calculation Algorithms / Vishnevskiy V., Kozyrev D. (eds). *Distributed Computer and Communication Networks. DCCN 2018. Communications in Computer and Information Science. V. 919. Geneva: Springer, 2018.*
https://doi.org/10.1007/978-3-319-99447-5_35

14. *Mandel A., Laptin V.* Channel Switching Threshold Strategies for Multichannel Controllable Queuing Systems // Communications in Computer and Information Science. 2020. V. 1337. P. 259–270. [link.springer.com/chapter
https://doi.org/ 10.1007/978-3-030-66242-4_21](https://link.springer.com/chapter/https://doi.org/10.1007/978-3-030-66242-4_21)
15. *Гнеденко Б.В., Коваленко И.Н.* Введение в теорию массового обслуживания. М.: Наука, 1966.
16. *Mandel A., Granin S.* Investigation of Analogies between the Problems of Inventory Control and the Problems of the Controllable Queuing Systems // Proc. 2018 Eleventh Int. Conf. “Management of large-scale system development” (MLSD’2018). IEEE, 2018. P. 1–4. [https://doi.org/ 10.1109/MLSD.2018.8551852](https://doi.org/10.1109/MLSD.2018.8551852)
17. *Arrow K., Karlin S., Scarf H.* Studies in the Mathematical Theory of Inventory and Production. Stanford University Press, 1958.
18. *Хедли Дж., Уайтин Т.* Анализ систем управления запасами. М.: Наука, 1969.

Статья представлена к публикации членом редколлегии А.А. Лазаревым.

Поступила в редакцию 20.01.2021

После доработки 20.03.2021

Принята к публикации 30.06.2021