

Оптимизация двухфазной системы массового обслуживания и ее применение к управлению передачей данных между двумя агентами робототехнической системы¹

Н. А. Кузнецов^{*,**,a}, Д. В. Мясников^{*,b}, К. В. Семенихин^{*,**,***,c}

^{*}Институт радиотехники и электроники им. В.А. Котельникова РАН, Москва, Россия,

^{**}Московский физико-технический институт (государственный университет),
Долгопрудный, Москва

^{***}Московский авиационный институт (национальный исследовательский университет),
Москва, Россия

e-mail: ^akuznetsov@cplire.ru, ^bdmitry.mysnikov89@gmail.com, ^asiemenkv@gmail.com

Поступила в редколлегию 15.02.2017

Аннотация—Рассматривается управляемая сеть массового обслуживания типа тандем, включающая две одноканальные системы с конечной очередью и предусматривающая механизм блокировки первого сервера. Первая система принимает нестационарный пуассоновский поток пакетов, обработка которых производится с управляемой скоростью. Если очередь в первой системе переполнена, поступающий пакет теряется. Вторая система не допускает своей перегрузки за счет управления вероятностью приема (ее снижение приводит к замедлению отсылки пакета из первой системы). Данная сеть массового обслуживания описывается управляемым марковским процессом, оптимизация которого проводится на конечном промежутке времени на основе минимизации среднего числа потерь с учетом ограничений на время полного обслуживания и энергозатраты первой системы. Разработанные алгоритмы использованы для синтеза закона управления нестационарным потоком данных в двухагентной робототехнической системе.

КЛЮЧЕВЫЕ СЛОВА: оптимальное управление, многоагентная система, двухфазная система массового обслуживания, управляемый марковский процесс, условная оптимизация.

1. ВВЕДЕНИЕ

Многофазные (тандемные) системы массового обслуживания используются для моделирования процесса обработки, при котором входящие требования обслуживаются последовательно на нескольких этапах. Механизм последовательного обслуживания естественным образом возникает при обработке запросов в центрах информационной поддержки [1], при передаче мультимедийной информации по каналам беспроводной связи [2], при управлении потоком данных между элементами многоагентной робототехнической системы [3].

Для предотвращения перегрузки на наиболее важных узлах многофазной системы применяется механизм блокировки требований. Двухфазные системы с блокировкой изучались в [4–6], где при разных предположениях о входящем потоке и распределении времени обслуживания были найдены стационарные вероятности состояний.

¹ Модели и алгоритмы управления нестационарным потоком данных в двухагентной робототехнической системе (разделы 2, 3, 7) разработаны при финансовой поддержке Российского научного фонда (проект № 16-11-00063). Решение задачи условной оптимизации двухфазной системы массового обслуживания (разделы 4–6) получено при финансовой поддержке Российского фонда фундаментальных исследований (проекты № 15-37-20611-мол_а_вед и № 16-07-00677-а).

В первых публикациях [7, 8] об управлении двухфазными системами с обратной связью была определена порогового структура для стратегии, оптимальной по критерию минимума средневзвешенной загрузки узлов. В этих статьях на базе теории марковских процессов принятия решений были получены уравнения динамического программирования. Несколько позже эта методология была распространена на более общие модели управляемых сетей массового обслуживания [9–11]. В недавних работах [12, 13] об оптимизации двухфазных систем массового обслуживания была рассмотрена задача оптимального управления доступом, где указаны явные выражения для порогового значения коэффициентов критерия.

Анализ этих и близких публикаций показывает, что управляемые двухфазные системы массового обслуживания изучаются исключительно в стационарном режиме, а качество управления определяется только одним функционалом. Поэтому задача оптимального управления тандемной системой на конечном промежутке времени с учетом ограничений является актуальной проблемой оптимизации управляемых сетей массового обслуживания.

Теория управляемых марковских процессов принятия решений, оптимизируемых с учетом нескольких показателей качества в стационарном режиме, разработана в [14]. Методология условной оптимизации скачкообразных марковских процессов с непрерывным временем и на конечном горизонте изложена в [15, 16]. Применение разработанной методологии к оптимизации одноканальной системы массового обслуживания описано в [17].

В настоящей работе рассмотрена управляемая двухфазная система массового обслуживания, включающая две одноканальные системы с конечной очередью и предусматривающая механизм блокировки первого сервера. На вход первой системы поступает нестационарный пуассоновский поток пакетов, обработка которых производится с управляемой скоростью на первом сервере. Входящий пакет теряется, если очередь первой системы оказывается заполненной полностью. Вторая система не допускает перегрузки за счет управления вероятностью приема, иначе говоря, вероятностью блокировки. Снижение вероятности приема приводит к замедлению отсылки пакета из первой системы. Данная сеть массового обслуживания описывается управляемым марковским процессом, оптимизация которого проводится на конечном промежутке времени на основе минимизации среднего числа потерь с учетом ограничений на время полного обслуживания и энергозатраты первой системы. Разработаны два алгоритма для определения оптимальных управлений, на классе централизованных и децентрализованных стратегий. Приведены результаты численного эксперимента, демонстрирующие характерный вид оптимальных стратегий в задаче управления потоком данных в двуагентной робототехнической системе.

2. НЕФОРМАЛЬНОЕ ОПИСАНИЕ МОДЕЛИ

Рассмотрим открытую сеть массового обслуживания типа тандем, состоящую из двух одноканальных систем. Первую систему будем называть передатчиком, а вторую — базовой станцией. Передатчик принимает входной поток пакетов (блоков информации) и после некоторого случайного промежутка времени пересылает их для дальнейшей обработки на базовую станцию.

Число пакетов в обеих системах ограничено и максимально может быть равно M и N соответственно. Если очередь первой системы заполнена полностью, то принимаемый пакет теряется. Это событие считается крайне нежелательным, поэтому его вероятность необходимо сделать, как можно меньше. Для этого предусмотрена возможность регулировки двух параметров $\mu \geq 0$ и $\vartheta \in [0, 1]$, где μ обозначает интенсивность обработки пакета в первой системе, а $1 - \vartheta$ равно вероятности отказа второй системы принять пакет. Отказ в приеме пакета используется для предотвращения перегрузки базовой станции. В случае отказа пакет не теряется,

а остается на передатчике. Поэтому произведение $\mu\vartheta$ можно трактовать как интенсивность пересылки пакета с передатчика на базовую станцию.

Рассматриваемая сеть массового обслуживания должна обеспечивать определенный уровень среднего времени прохождения пакета от момента его прихода на передатчик до момента окончания обработки на сервере базовой станции с учетом того, что интенсивность этой обработки ν предполагается фиксированной величиной.

Для минимизации числа потерянных пакетов и сокращения времени полного обслуживания необходимо наращивать интенсивность обработки μ . Однако функционирование передатчика связано с жесткими энергетическими ограничениями, которые исключают неограниченное увеличение интенсивности μ .

Параметр ϑ также должен соответствовать некоторому промежуточному состоянию между двумя крайними ситуациями, когда все пакеты от передатчика принимаются ($\vartheta = 1$), или, когда прием пакетов на базовую станцию заблокирован ($\vartheta = 0$). Прием всех пакетов ведет к быстрому заполнению буфера базовой станции, что с некоторого момента приводит к ее продолжительной блокировке и, как следствие, к переполнению буфера передатчика, а затем и к массовой потере пакетов.

Для того чтобы учитывать нестационарность входного потока и реагировать на текущие изменения в функционировании сети, оба параметра μ и ϑ предполагаются зависящими от времени и состояния. Тем самым искомое управление строится по принципу обратной связи. Однако здесь необходимо выделить две принципиально разные постановки. В первой постановке управление интенсивностью обработки μ на передатчике и вероятностью приема ϑ базовой станцией осуществляется на основе *полной информации* о текущем состоянии всей сети. Такое управление будем называть *централизованным*, поскольку действия передатчика и базовой станции предполагаются согласованными. Во второй постановке интенсивность μ определяется только состоянием передатчика, а вероятность ϑ наоборот зависит лишь от состояния базовой станции. Тем самым принятие решений обеими системами производится в ситуации *неполной информации* о состоянии другой системы. Поэтому управление, соответствующее второй постановке, будем называть *децентрализованным*.

Случай неполной информации наиболее реалистично отражает проблему пересылки видеопотока или телеметрии на стационарный пункт сбора информации с передатчика, установленного на борту беспилотного летательного аппарата (БПЛА), который выполняет автономную миссию. Система массового обслуживания БПЛА занимается преобразованием входного потока разнородной информации в последовательность унифицированных пакетов, предусмотренных протоколом обмена данными с базовой станцией. Поэтому интенсивность обработки μ определяется временем, затраченным на формирование и отсылку пакета, а вероятность ϑ задается частотой прихода сообщения от базовой станции об успешной доставке пакета. При наличии отказов, т.е. при $\vartheta < 1$, среднее время отсылки пакета $1/\mu$ увеличивается на $(1/\vartheta - 1) \cdot 100\%$.

Таким образом, специфика описанной выше проблемы порождает необходимость оптимизации инфотелекоммуникационной сети типа тандем на фиксированном промежутке времени с учетом нестационарности входного потока данных и ограничений на потребление энергии.

3. ПОСТАНОВКА ЗАДАЧИ

Теперь перейдем к формальному описанию рассматриваемой сети массового обслуживания.

Пусть $X(t)$ и $Y(t)$ обозначают число пакетов, находящихся в момент времени t соответственно на передатчике и базовой станции. Тогда случайный процесс $Z(t) = (X(t), Y(t))$, описывающий текущее состояние сети, принимает значения из множества

$$\mathcal{Z} = \mathcal{X} \times \mathcal{Y}, \quad \text{где} \quad \mathcal{X} = \{0, 1, \dots, M\}, \quad \mathcal{Y} = \{0, 1, \dots, N\}.$$

Допустим, что пакеты, приходящие на передатчик, образуют нестационарный пуассоновский поток с известной непрерывной интенсивностью $\alpha(t)$. Рассмотрим сначала постоянное управление u :

$$u = (m, v) \in \mathcal{U}, \quad \mathcal{U} = [\underline{m}, \overline{m}] \times [\underline{v}, \overline{v}],$$

где $0 \leq \underline{m} \leq \overline{m} < \infty$ и $0 < \underline{v} \leq \overline{v} \leq 1$ — заданные границы изменения для интенсивности обработки на передатчике и для вероятности приема базовой станцией. В этой ситуации $Z(t)$ является неоднородным марковским процессом. Его генератор $A(t, u)$ представляет собой линейный оператор, действующий в пространстве $\mathbb{R}^{\mathcal{Z}}$, т.е. в пространстве вещественных функций h , определенных на множестве состояний $\mathcal{Z}$. Если элементы $h \in \mathbb{R}^{\mathcal{Z}}$ записывать в виде набора $h = \{h_z\}_{z \in \mathcal{Z}}$, то

$$(A(t, u)h)_z = \lim_{\delta \downarrow 0} \frac{\mathbb{E}\{h_{Z(t+\delta)} \mid Z(t) = z\} - h_z}{\delta} = \sum_{z' \in \mathcal{Z}} a_{z,z'}(t, u) h_{z'}$$

с учетом того, что $a_{z,z'}(t, u)$ при $z \neq z'$ обозначает интенсивность перехода $z \rightarrow z'$, а $a_{z,z}(t, u)$ — число, противоположное интенсивности выхода из состояния z , т.е.

$$a_{z,z}(t, u) = - \sum_{z' \in \mathcal{Z} \setminus \{z\}} a_{z,z'}(t, u).$$

Для рассматриваемого процесса возможны только три варианта переходов: прием пакета передатчиком; его пересылка; обработка на базовой станции. Поэтому для каждого состояния $(x, y) \in \mathcal{Z}$ имеем

$$\begin{aligned} a_{(x,y),(x+1,y)}(t, u) &= \alpha(t), \quad x < M, \\ a_{(x,y),(x-1,y+1)}(t, u) &= mv, \quad x > 0, y < N, \\ a_{(x,y),(x,y-1)}(t, u) &= \nu, \quad y > 0. \end{aligned}$$

Теперь предположим, что управление $U(t)$ описывается случайным процессом

$$U(t) = (\mu(t), \vartheta(t)), \tag{1}$$

в котором интенсивность обработки $\mu(t)$ и вероятность приема $\vartheta(t)$ определяются функциями времени и текущего состояния сети:

$$\mu(t) = m_{Z(t)}(t), \quad \vartheta(t) = v_{Z(t)}(t). \tag{2}$$

Функции $m_z(t)$ и $v_z(t)$, называемые далее *стратегиями*, суть борелевские функции со значениями в множествах $[\underline{m}, \overline{m}]$ и $[\underline{v}, \overline{v}]$ соответственно. Поскольку указанные стратегии параметризованы индексом $z \in \mathcal{Z}$, они являются *централизованными*. Класс управлений (2), использующих полную информацию о состоянии сети, будем обозначать $\mathcal{U}$.

Если же передатчик и базовая станция управляются на основе информации о состоянии только своей системы массового обслуживания, то соответствующие стратегии называются *децентрализованными* и обозначаются в виде $m_x(t), v_y(t)$, где $x \in \mathcal{X}$ и $y \in \mathcal{Y}$. В этом случае процесс (1) определяется по правилу

$$\mu(t) = m_{X(t)}(t), \quad \vartheta(t) = v_{Y(t)}(t). \tag{3}$$

Рассматриваемые управления U относятся к классу *марковских* управлений, поскольку $U(t)$ определяется только текущим состоянием управляемого процесса $Z(t)$ без учета его предыдущей эволюции. Кроме того, $Z(t)$ остается марковским процессом, но его генератор теперь определяется другим линейным отображением:

$$g_z = \sum_{z' \in Z} a_{z,z'}(t, m_z(t), v_z(t)) h_{z'} \quad \forall z \in Z.$$

При использовании децентрализованной стратегии в указанном представлении необходимо заменить $m_z(t)$ на $m_x(t)$ и $v_z(t)$ на $v_y(t)$.

Перейдем к формулировке оптимизационной задачи.

Как упоминалось выше, основной критерий оптимизации рассматриваемой сети массового обслуживания — это минимум среднего числа потерянных пакетов на конечном промежутке $[0, T]$. Так как потеря пакета передатчиком происходит только в случае переполнения его очереди, целевой функционал принимает вид

$$J_0[U] = \int_0^T P\{X(t) = M\} \alpha(t) dt. \quad (4)$$

Функционал, характеризующий время полного обслуживания, можно задать следующим образом:

$$S[U] = \frac{1}{T} \int_0^T E\{X(t) + Y(t)\} dt \Big/ \frac{1}{T} \int_0^T P\{X(t) < M\} \alpha(t) dt. \quad (5)$$

Для обоснования данного соотношения рассмотрим систему массового обслуживания, в которой соответствующие процессы являются стационарными и эргодическими (подробно см. [18, § 5.8]). Тогда справедлива формула Литтла:

$$\text{среднее время полного обслуживания} = \frac{\text{среднее число заявок в системе}}{\text{интенсивность входного потока}}.$$

Соотношение (5) повторяет данную формулу с учетом того, что в нестационарном случае среднее необходимо брать и по пространству и по времени, а при прореживании входного потока его интенсивность должна быть умножена на вероятность отсутствия потерь.

Теперь определим функционал, характеризующий энергозатраты передатчика:

$$E[U] = \int_0^T E\{\mu(t) I\{X(t) > 0\}\} dt. \quad (6)$$

Суть данного выражения легко объяснить, если предположить, что мощность, потребляемая сервером передатчика, пропорциональна его интенсивности μ .

Итак, рассматриваемая в статье задача оптимального управления состоит в следующем:

$$J_0[U] \rightarrow \min_{U \in \mathcal{U}} \quad \text{при ограничениях} \quad S[U] \leq \bar{S}, \quad E[U] \leq \bar{E}, \quad (7)$$

где $\bar{S}$ и $\bar{E}$ — верхние границы на время полного обслуживания и энергопотребления.

Отметим, что (7) — задача с полной информацией о состоянии управляемой сети, поэтому искомое оптимальное управление $\hat{U}(t)$ будет определяться централизованной стратегией.

Детализация формулировки задачи оптимизации на классе децентрализованных стратегий будет дана несколько позже.

4. ОПТИМАЛЬНОЕ УПРАВЛЕНИЕ ОТНОСИТЕЛЬНО СВЕРТКИ КРИТЕРИЕВ

Преобразуем задачу оптимизации управления с полной информацией (7) к следующей эквивалентной формулировке:

$$J_0[U] \rightarrow \min_{U \in \mathcal{U}}: \quad J_1[U] \leq 0, \quad J_2[U] \leq 0, \quad (8)$$

в которой функционалы $J_1[U]$ и $J_2[U]$ являются интегральными и содержат в себе указанные выше границы, т.е.

$$J_1[U] = \int_0^T \mathbf{E} \{ X(t) + Y(t) - \bar{S} \mathbf{I} \{ X(t) < M \} \alpha(t) \} dt, \quad (9)$$

$$J_2[U] = \int_0^T \mathbf{E} \{ \mu(t) \mathbf{I} \{ X(t) > 0 \} - \bar{E}/T \} dt. \quad (10)$$

Рассмотрим сначала задачу безусловной оптимизации линейной комбинации введенных выше функционалов

$$\langle \lambda, J[U] \rangle = \lambda_0 J_0[U] + \lambda_1 J_1[U] + \lambda_2 J_2[U] \rightarrow \min_{U \in \mathcal{U}}, \quad (11)$$

где $J[U] = \text{col}[J_0[U], J_1[U], J_2[U]]$ — векторный критерий, $\lambda = \text{col}[\lambda_0, \lambda_1, \lambda_2]$ — вектор неотрицательных коэффициентов, а U пробегает класс управлений с полной информацией $\mathcal{U}$.

В дальнейшем решение задачи (11) будет рассматриваться как подготовительный шаг для последующего синтеза оптимального управления с учетом ограничений (8). Тем не менее определение стратегий, оптимальных в задаче без ограничений, имеет самостоятельный интерес, так как позволяет заранее проанализировать их чувствительность по отношению к выбору весовых коэффициентов λ_l .

Поскольку каждый из функционалов $J_l[\cdot]$ допускает запись в виде интеграла от математического ожидания, то же представление будет справедливо и для их свертки

$$\langle \lambda, J[U] \rangle = \int_0^T \mathbf{E} \langle \lambda, g(t, Z(t), U(t)) \rangle dt \quad (12)$$

при подходящем выборе функции $g(t, z, u)$. Запишем этот функционал в виде

$$\langle \lambda, J[U] \rangle = \int_0^T \sum_{l=0}^2 \sum_{z \in \mathcal{Z}} \lambda_l f_{l,z}(t, m_z(t), v_z(t)) \pi_z(t) dt, \quad (13)$$

с учетом того, что $U(t)$ — управление, определяемое стратегиями $m_z(t)$ и $v_z(t)$ согласно (2), $\pi_z(t) = \mathbf{P}\{Z(t) = z\}$ — вероятности состояний сети, а функции $f_{l,z}(t, m, v)$ при $z = (x, y)$ равны

$$f_{l,(x,y)}(t, m, v) = \begin{cases} \mathbf{I}\{x = M\} \alpha(t), & l = 0, \\ x + y - \bar{S} \alpha(t) \mathbf{I}\{x < M\}, & l = 1, \\ m \mathbf{I}\{x > 0\} - \bar{E}/T, & l = 2. \end{cases} \quad (14)$$

В [15] разработан метод синтеза управления, оптимального на классе всех предсказуемых стратегий для управляемого марковского процесса с конечным числом состояний. Соответствующее оптимальное управление будет марковским, поэтому оно же будет доставлять оптимум в задаче на классе управлений $U \in \mathcal{U}$ вида (2).

Для построения управления, оптимального относительно свертки критериев

$$\tilde{U}(\cdot, \lambda) \in \arg \min_{U \in \mathcal{U}} \langle \lambda, J[U] \rangle \quad (15)$$

достаточно выполнить следующие шаги:

- 1) представить минимизируемый функционал в интегральной форме

$$\langle \lambda, J[U] \rangle = \int_0^T \langle F^*(t, u) \lambda, \pi(t) \rangle dt$$

для случая постоянного управления $U(t) \equiv u$, где u — это произвольная точка из множества значений управляющих воздействий U ;

- 2) задать функцию $W(\cdot) = \{W_z(\cdot)\}_{z \in \mathcal{Z}}$ со значениями в $\mathbb{R}^{\mathcal{Z}}$

$$W(t, \phi, u, \lambda) = A(t, u) \phi + F^*(t, u) \lambda, \quad t \in [0, T], \phi \in \mathbb{R}^{\mathcal{Z}}, u \in U; \quad (16)$$

- 3) решить параметрическую задачу минимизации

$$\tilde{u}_z(t, \phi, \lambda) \in \arg \min_{u \in U} W_z(t, \phi, u, \lambda); \quad (17)$$

- 4) определить решение $\phi(t, \lambda) = \{\phi_z(t, \lambda)\}_{z \in \mathcal{Z}}$ системы уравнений динамического программирования

$$\dot{\phi}_z(t, \lambda) = - \min_{u \in U} W_z(t, \phi(t, \lambda), u, \lambda), \quad t \in [0, T], \quad \phi_z(T, \lambda) = 0. \quad (18)$$

Тогда искомое управление и оптимальное значение функционала определяются по правилу

$$\tilde{U}(t, \lambda) = \tilde{u}_{Z(t)}(t, \phi(t, \lambda), \lambda) \quad \text{и} \quad \min_{U \in \mathcal{U}} \langle \lambda, J[U] \rangle = \langle \phi(0, \lambda), \pi(0) \rangle. \quad (19)$$

Укажем представление для функции

$$W_z(t, \phi, u, \lambda) = \sum_{z' \in \mathcal{Z}} a_{z, z'}(t, u) \phi_{z'} + \sum_{l=0}^2 \lambda_l f_{l, z}(t, u). \quad (20)$$

На рис. 1 отмечены характерные случаи возможных переходов из данного состояния (x, y) . В зависимости от этих случаев выражение для функции $W_{x, y}(t, \phi, (m, v), \lambda)$ принимает вид:

- а) $\alpha \phi_{1,0} - \alpha \phi_{0,0} - \lambda_1 \bar{S} \alpha + \lambda_2 (-\bar{E}/T)$, если $x = y = 0$;
- б) $\alpha \phi_{x+1,0} + mv \phi_{x-1,1} - (\alpha + mv) \phi_{x,0} + \lambda_1 (x - \bar{S} \alpha) + \lambda_2 (m - \bar{E}/T)$, если $0 < x < M, y = 0$;
- в) $mv \phi_{M-1,1} - mv \phi_{M,0} + \lambda_0 \alpha + \lambda_1 M + \lambda_2 (m - \bar{E}/T)$, если $x = M, y = 0$;
- г) $\alpha \phi_{x+1,y} + \nu \phi_{x,y-1} - (\alpha + \nu) \phi_{x,y} + \lambda_1 (x + y - \bar{S} \alpha) + \lambda_2 (m I\{x > 0\} - \bar{E}/T)$, если $x = 0, y > 0$ или $x < M, y = N$;
- е) $\alpha \phi_{x+1,y} + mv \phi_{x-1,y+1} + \nu \phi_{x,y-1} - (\alpha + mv + \nu) \phi_{x,y} + \lambda_1 (x + y - \bar{S} \alpha) + \lambda_2 (m - \bar{E}/T)$, если $0 < x < M, 0 < y < N$;

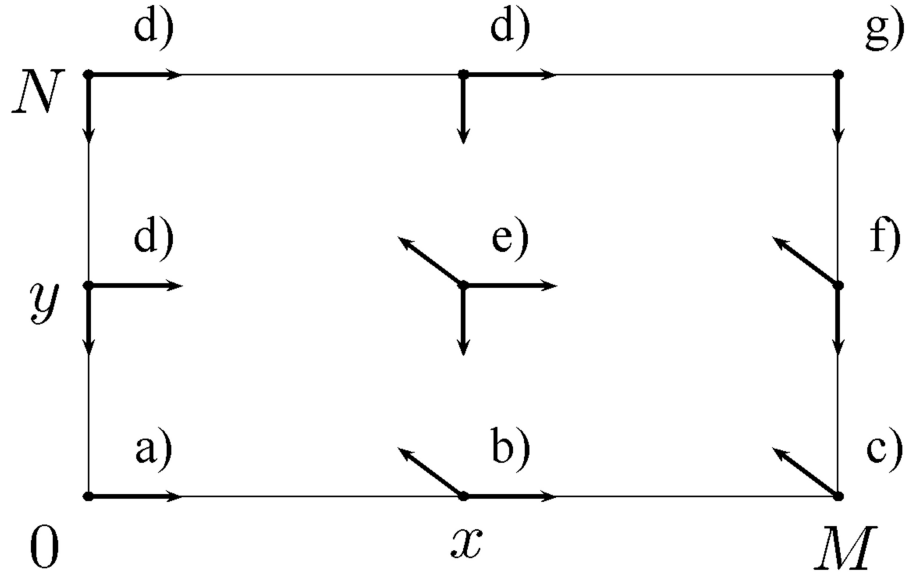


Рис. 1. Переходы между состояниями сети.

f) $mv\phi_{M-1,y+1} + \nu\phi_{M,y-1} - (mv + \nu)\phi_{M,y} + \lambda_0\alpha + \lambda_1(M + y) + \lambda_2(m - \bar{E}/T)$, если $x = M$, $0 < y < N$;

g) $\nu\phi_{M,N-1} - \nu\phi_{M,N} + \lambda_0\alpha + \lambda_1(M + N) + \lambda_2(m - \bar{E}/T)$, если $x = M$, $y = N$.

Для краткости зависимость интенсивности $\alpha(t)$ от времени была опущена.

Чтобы определить вид оптимальной стратегии, представим $W_{x,y}(t, \phi, (m, v), \lambda)$ как функцию переменных m, v , скрыв в обозначении «...» зависимость от других переменных:

$$W_{x,y}(t, \phi, (m, v), \lambda) = \begin{cases} m\lambda_2 I\{x > 0\} + \dots, & \text{a), d), g);} \\ m(v(\phi_{x-1,y+1} - \phi_{x,y}) + \lambda_2) + \dots, & \text{b), c), e), f).} \end{cases} \quad (21)$$

Тогда искомая оптимальная стратегия (17) будет иметь вид

$$\tilde{u}_{x,y}(t, \phi, \lambda) = \begin{cases} (\underline{m}, \bar{v}), & x = 0 \text{ или } y = N; \\ \text{MV}(\phi_{x-1,y+1} - \phi_{x,y}; \lambda_2), & x > 0, y < N, \end{cases} \quad (22)$$

где $\text{MV}(a; b)$ — обозначение для решения задачи минимизации

$$m(av + b) \rightarrow \min_{m,v}: \quad \underline{m} \leq m \leq \bar{m}, \quad \underline{v} \leq v \leq \bar{v}, \quad (23)$$

параметризованной числами a, b . В качестве этого решения можно взять:

$$\text{MV}(a; b) = (\tilde{m}, \tilde{v}): \quad \tilde{m} = \begin{cases} \underline{m}, & a\tilde{v} + b \geq 0, \\ \bar{m}, & a\tilde{v} + b < 0, \end{cases} \quad \tilde{v} = \begin{cases} \underline{v}, & a > 0, \\ \bar{v}, & a \leq 0. \end{cases} \quad (24)$$

Теперь в силу теоремы 3 из [15] получаем следующий результат.

Теорема 1. Задача Коши для системы обыкновенных дифференциальных уравнений (18) имеет единственное решение $\phi(t, \lambda) = \{\phi_z(t, \lambda)\}_{z \in \mathcal{Z}}$, которое совпадает с функцией Беллмана, т.е.

$$\phi_z(t, \lambda) = \inf_{U \in \mathcal{U}} \int_t^T \mathbb{E}\{\langle \lambda, g(\tau, Z(\tau), U(\tau)) \rangle \mid Z(t) = z\} d\tau \quad \forall t \in [0, T] \quad \forall z \in \mathcal{Z},$$

где $g(t, z, u)$ — функция из представления (12).

В частности, оптимум в задаче (11) равен

$$\min_{U \in \mathcal{U}} \langle \lambda, J[U] \rangle = \langle \phi(0, \lambda), \pi(0) \rangle,$$

где $\pi(0)$ — начальное распределение процесса $Z(t) = (X(t), Y(t))$, $t \in [0, T]$.

Тогда оптимальное управление в задаче (11) строится согласно выражениям (19) и (22), а именно,

$$\tilde{U}(t, \lambda) = \begin{cases} (\underline{m}, \bar{v}), & \text{если } X(t) = 0 \text{ или } Y(t) = N, \\ \text{MV}(\phi_{x-1, y+1}(t, \lambda) - \phi_{x, y}(t, \lambda); \lambda_2), & \text{если } X(t) = x > 0 \text{ и } Y(t) = y < N, \end{cases}$$

где $\text{MV}(a; b)$ — обозначение, введенное в (24).

Структуру указанного выше управления $\tilde{U}(t, \lambda) = (\mu(t), \vartheta(t))$ можно объяснить следующим образом.

Разность $a = \phi_{x-1, y+1}(t, \lambda) - \phi_{x, y}(t, \lambda)$ определяет затраты на пересылку пакета с передатчика на базовую станцию. Если $a \leq 0$, то пересылка пакета оказывается более выгодной по сравнению с его задержкой, поэтому согласно (24) оптимальная вероятность приема $\vartheta(t)$ полагается равной верхней границе $\bar{v}$. Если же $a > 0$, то, наоборот, более предпочтительно удержание пакета на передатчик, что приводит к минимизации вероятности приема $\vartheta(t) = \underline{v}$.

Для оптимальной стратегии $\mu(t)$ выбор между двумя альтернативами $\underline{m}$ и $\bar{m}$ происходит на основе сравнения величины a , умноженной на вероятность приема $\vartheta(t)$, и коэффициента λ_2 , определяющего важность ограничения на энергопотребление. Если $a\vartheta(t) + \lambda_2 \geq 0$, то экономия энергии является более существенным фактором по сравнению с затратами, связанными с задержкой пакета на передатчик. Поэтому интенсивность обработки будет находиться на минимальном уровне $\mu(t) = \underline{m}$. При $a\vartheta(t) + \lambda_2 < 0$, наоборот, необходимость пересылки пакета оказывается важнее энергозатрат, откуда $\mu(t) = \bar{m}$.

5. СИНТЕЗ УПРАВЛЕНИЯ В ЗАДАЧЕ С ОГРАНИЧЕНИЯМИ

Рассмотрим задачу (8) минимизации целевого функционала $J_0[U]$ на классе управлений с полной информацией $\mathcal{U}$ с учетом ограничений $J_l[U] \leq 0$, $l = 1, 2$.

Для построения оптимального управления в этой задаче воспользуемся подходом, предложенным в [15]. Для этого необходимо выполнить следующие шаги:

- 1) сформулировать исходную задачу с ограничениями в виде эквивалентной минимаксной постановки для свертки критериев

$$\hat{U} \in \arg \min_{U \in \mathcal{U}} \max_{\lambda \in \Lambda} \langle \lambda, J[U] \rangle, \quad (25)$$

где Λ — выпуклое компактное множество, содержащее векторы коэффициентов $\lambda = \text{col}[1, \lambda_1, \lambda_2]$ с неотрицательными координатами;

2) определить решение двойственной задачи

$$\hat{\lambda} \in \arg \max_{\lambda \in \Lambda} \min_{U \in \mathcal{U}} \langle \lambda, J[U] \rangle; \quad (26)$$

3) взять в качестве искомого управления то, которое является оптимальным относительно свертки критериев с найденными выше коэффициентами, т.е. $\hat{U}(t) = \tilde{U}(t, \hat{\lambda})$.

Отметим, что минимаксная постановка (25) заведомо будет эквивалентной задаче с ограничениями (11), если вместо операции «max» взять «sup» и считать, что Λ содержит произвольные векторы вида $\lambda = \text{col}[1, \lambda_1, \lambda_2]$, $\lambda_1 \geq 0$, $\lambda_2 \geq 0$. Чтобы выбрать множество Λ ограниченным, нужно обеспечить принадлежность ему любого вектора $\hat{\lambda}$, удовлетворяющего условиям Куна—Таккера. Для этого согласно [17] достаточно указать управление $U^o \in \mathcal{U}$, на котором выполнялось бы условие Слейтера: $J_l[U^o] < 0$, $l = 1, 2$, и положить

$$\Lambda = \{\lambda \in \mathbb{R}^3: \lambda_0 = 1, 0 \leq \lambda_1 \leq c_1, 0 \leq \lambda_2 \leq c_2\}, \quad (27)$$

где числа c_1, c_2 удовлетворяют $c_l \geq -J_0[U^o]/J_l[U^o] > 0$.

Двойственная задача (26) представляет собой задачу выпуклого программирования. Как и в [17], для ее решения можно использовать метод условного градиента [19, 22] или квазиньютоновский алгоритм [20], специально предназначенный для оптимизации на координатном параллелепипеде.

Для обоснования последнего шага достаточно убедиться в том, что оптимальное управление $\tilde{U}(t, \lambda)$ непрерывно зависит от вектора множителей λ (см. теорему 4 из [15]). Однако согласно (22)–(24) при прохождении через нуль соответствующих коэффициентов оптимальные значения интенсивности обработки $\tilde{m}$ и вероятности приема $\tilde{v}$ меняются скачком. Поэтому гарантировать непрерывную зависимость оптимальной стратегии $\tilde{u}_z(t, \phi, \lambda)$ от λ невозможно.

Тем не менее, применение изложенной выше схемы 1)–3) возможно за счет перехода к регуляризованной минимаксной задаче [21]:

$$\hat{U}^\varepsilon \in \arg \min_{U \in \mathcal{U}} \max_{\lambda \in \Lambda} \langle \lambda, J[U] \rangle + \Sigma^\varepsilon[U], \quad (28)$$

где $\Sigma^\varepsilon[U]$ обозначает стабилизирующий функционал

$$\Sigma^\varepsilon[U] = \frac{1}{2} \int_0^T \mathbb{E} \{ \varepsilon_1 (\mu(t) - \underline{m})^2 + \varepsilon_2 \mu(t) (\bar{v} - \vartheta(t))^2 \} dt, \quad (29)$$

а $\varepsilon_1 > 0$ и $\varepsilon_2 > 0$ — параметры регуляризации. Стабилизатор выбран в интегральной форме по аналогии с функционалом $\langle \lambda, J[U] \rangle$, а также из расчета на то, чтобы минимум $\Sigma^\varepsilon[U]$ достигался на постоянной стратегии $(\underline{m}, \bar{v})$, обеспечивающей минимальное потребление энергии и наибольшую вероятность приема.

Оптимальное управление в регуляризованной задаче без ограничений

$$\tilde{U}^\varepsilon(\cdot, \lambda) \in \arg \min_{U \in \mathcal{U}} \langle \lambda, J[U] \rangle + \Sigma^\varepsilon[U] \quad (30)$$

строится по теореме 1 с учетом того, что функция $\{W_z(\cdot)\}$, задающая правую часть уравнения динамического программирования (18), должна быть заменена ее регуляризованной версией

$$W_z^\varepsilon(t, \phi, u, \lambda) = W_z(t, \phi, u, \lambda) + (\varepsilon_1 (m - \underline{m})^2 + \varepsilon_2 m (\bar{v} - v)^2)/2, \quad u = (m, v), \quad (31)$$

а оптимальная стратегия

$$\tilde{u}_z^\varepsilon(t, \phi, \lambda) \in \arg \min_{u \in U} W_z^\varepsilon(t, \phi, u, \lambda) \quad (32)$$

определена теперь однозначно из решения задачи, аналогичной (23)

$$\varepsilon_1(m - \underline{m})^2/2 + m(av + b + \varepsilon_2(\bar{v} - v)^2/2) \rightarrow \min_{m, v}: \quad \underline{m} \leq m \leq \bar{m}, \quad \underline{v} \leq v \leq \bar{v}.$$

Нетрудно проверить, что решение этой задачи — есть пара $(\tilde{m}, \tilde{v})$, такая что

$$\tilde{m} = \begin{cases} \underline{m}, & \tilde{c} \geq 0, \\ \underline{m} - \tilde{c}/\varepsilon_1, & -\varepsilon_1(\bar{m} - \underline{m}) \leq \tilde{c} \leq 0, \\ \bar{m}, & \tilde{c} \leq -\varepsilon_1(\bar{m} - \underline{m}), \end{cases} \quad \tilde{v} = \begin{cases} \underline{v}, & a \geq \varepsilon_2(\bar{v} - \underline{v}), \\ \bar{v} - a/\varepsilon_2, & 0 \leq a \leq \varepsilon_2(\bar{v} - \underline{v}), \\ \bar{v}, & a \leq 0, \end{cases} \quad (33)$$

где $\tilde{c} = a\tilde{v} + b + \varepsilon_2(\bar{v} - \tilde{v})^2/2$. Если указанную пару $(\tilde{m}, \tilde{v})$ обозначить $MV^\varepsilon(a; b)$, то стратегия (32) принимает вид

$$\tilde{u}_{x,y}^\varepsilon(t, \phi, \lambda) = \begin{cases} (\underline{m}, \bar{v}), & x = 0 \text{ или } y = N; \\ MV^\varepsilon(\phi_{x-1,y+1} - \phi_{x,y}; \lambda_2), & x > 0, y < N. \end{cases} \quad (34)$$

Благодаря регуляризации, оптимальная стратегия (34) непрерывно зависит от выбора λ и, следовательно, схема 1)–3) действительно приводит к решению минимаксной задачи (28). Более того, в силу единственности оптимального управления (30) соответствующая двойственная задача

$$\hat{\lambda}^\varepsilon \in \arg \max_{\lambda \in \Lambda} \underline{L}^\varepsilon(\lambda), \quad \underline{L}^\varepsilon(\lambda) = \min_{U \in \mathcal{U}} \langle \lambda, J[U] \rangle + \Sigma^\varepsilon[U], \quad (35)$$

представляет собой задачу гладкого выпуклого программирования.

Сформулированная ниже теорема устанавливает свойства регуляризованной задачи и ее связь с исходной задачей оптимального управления с ограничениями.

Теорема 2. Пусть дано управление $U^0 \in \mathcal{U}$, удовлетворяющее условию Слейтера: $J_l[U^0] < 0$, $l = 1, 2$; множество Λ определено по правилу (27) с использованием чисел c_1, c_2 таких, что $c_l \geq -(J_0[U^0] + \Sigma^\varepsilon[U^0])/J_l[U^0]$.

Тогда справедливы следующие утверждения:

- 1) оптимальное управление в регуляризованной задаче без ограничений (30) вычисляется в виде $\tilde{U}^\varepsilon(t, \lambda) = \tilde{u}_{X(t), Y(t)}^\varepsilon(t, \phi^\varepsilon(t, \lambda), \lambda)$ с помощью стратегии (34) и функции $\phi^\varepsilon(t, \lambda)$, определяемой из решения задачи Коши

$$\dot{\phi}_z^\varepsilon(t, \lambda) = -\min_{u \in U} W_z^\varepsilon(t, \phi^\varepsilon(t, \lambda), u, \lambda), \quad t \in [0, T], \quad \phi_z^\varepsilon(T, \lambda) = 0, \quad (36)$$

где $W^\varepsilon(\cdot)$ — функция, заданная в (31);

- 2) целевая функция в двойственной задаче (35) определяется выражением $\underline{L}^\varepsilon(\lambda) = \langle \phi^\varepsilon(0, \lambda), \pi(0) \rangle$, является выпуклой, дифференцируемой и имеет градиент

$$\nabla \underline{L}^\varepsilon(\lambda) = J[\tilde{U}^\varepsilon(\cdot, \lambda)]; \quad (37)$$

- 3) управление $\tilde{U}^\varepsilon(t, \hat{\lambda}^\varepsilon)$, соответствующее решению $\hat{\lambda}^\varepsilon$ регуляризованной двойственной задачи (35), подчиняется ограничениям (11) и удовлетворяет

$$J_0[\hat{U}] \leq J_0[\tilde{U}^\varepsilon(\cdot, \hat{\lambda}^\varepsilon)] \leq J_0[\hat{U}] + \Sigma^\varepsilon[\hat{U}], \quad (38)$$

где $\hat{U}(t)$ — управление, оптимальное в исходной задаче с ограничениями (11).

6. ОПТИМИЗАЦИЯ НА КЛАССЕ ДЕЦЕНТРАЛИЗОВАННЫХ СТРАТЕГИЙ

Теперь рассмотрим задачу управления в ситуации отсутствия полной информации о состоянии сети. Для этого оптимизационную постановку (7) на классе $\mathcal{U}_d$ децентрализованных управлений (3) запишем в виде:

$$J_0[U] \rightarrow \min_{U \in \mathcal{U}}: \quad J_1[U] \leq 0, \quad J_2[U] \leq 0, \quad (39)$$

где функционалы $J_1[U]$ и $J_2[U]$ определены выражениями (9) и (10).

Согласно (3) любое управление $U(t)$ из $\mathcal{U}_d$ определяется стратегией

$$\mathbf{u}(t) = \{m_x(t), v_y(t) : x \in \mathcal{X}, y \in \mathcal{Y}\}, \quad (40)$$

в котором $m_x(t)$ и $v_y(t)$ обозначают соответственно интенсивность обработки на передатчике и вероятность приема пакета базовой станцией в том случае, если в момент времени t сеть находится в состоянии (x, y) . Будем предполагать, что стратегия $\mathbf{u}(t)$ образует кусочно-непрерывную функцию, заданную на промежутке $[0, T]$ и принимающую значения в множестве $\mathbf{U} = [\underline{m}, \bar{m}]^{\mathcal{X}} \times [\underline{v}, \bar{v}]^{\mathcal{Y}}$. Данный класс функций будем обозначать $\mathcal{K}$.

По аналогии с (13) запишем каждый из функционалов $J_l[U]$ запишем в виде

$$J_l(\pi, \mathbf{u}) = \int_0^T \sum_{(x,y) \in \mathcal{Z}} f_{l,(x,y)}(t, m_x(t), v_y(t)) \pi_{x,y}(t) dt, \quad l = 0, 1, 2, \quad (41)$$

указав в качестве аргументов распределение состояний сети $\pi(t) = \{\pi_{x,y}(t)\}$ и стратегию (40). Тогда от (39) можно перейти к задаче оптимального управления

$$J_0(\pi, \mathbf{u}) \rightarrow \min_{\pi, \mathbf{u}}: \quad J_1(\pi, \mathbf{u}) \leq 0, \quad J_2(\pi, \mathbf{u}) \leq 0, \quad (42)$$

для детерминированной системы

$$\dot{\pi}_{x,y}(t) = \sum_{(x',y') \in \mathcal{Z}} a_{(x',y'),(x,y)}(t, m_{x'}(t), v_{y'}(t)) \pi_{x',y'}(t), \quad (x, y) \in \mathcal{Z}, \quad (43)$$

рассматриваемой на промежутке $[0, T]$ с фиксированным начальным условием π^0 и кусочно-непрерывными управлениями (40).

Систему дифференциальных уравнений (43) будем записывать кратко

$$\dot{\pi}(t) = \{\bar{A}(t, \mathbf{u}(t))\}^* \pi(t), \quad (44)$$

с использованием обозначения $\bar{A}(t, \mathbf{v})$, которое при фиксированных $t \in [0, T]$ и $\mathbf{v} \in \mathbf{U}$, определяет линейный оператор, действующий в пространстве $\mathbb{R}^{\mathcal{Z}}$. Ясно, что $\bar{A}(t, \mathbf{u}(t))$ совпадает с генератором управляемого марковского процесса, соответствующего указанной стратегии $\mathbf{u}(\cdot)$.

Аналогичное обозначение придется использовать для краткой записи вектора, составленного из значений функционалов (41):

$$J(\pi, \mathbf{u}) = \int_0^T \bar{F}(t, \mathbf{u}(t)) \pi(t) dt. \quad (45)$$

Если зафиксировать аргументы $t \in [0, T]$ и $\mathbf{v} \in \mathbf{U}$, то $\bar{F}(t, \mathbf{v})$ представляет собой линейный оператор из $\mathbb{R}^{\mathcal{Z}}$ в $\mathbb{R}^3$.

Принцип максимума для детерминированных систем позволяет сформулировать необходимое условие оптимальности для рассматриваемого класса управляемых сетей массового обслуживания.

Теорема 3. Если кусочно-непрерывная стратегия

$$\hat{\mathbf{u}}(t) = \{\hat{m}_x(t), \hat{v}_y(t) : x \in \mathcal{X}, y \in \mathcal{Y}\}$$

задает оптимальное управление $\hat{U}(t) = (\hat{m}_{X(t)}(t), \hat{v}_{Y(t)}(t))$ по неполной информации (39), то существуют вектор и функция

$$\hat{\lambda} = \text{col}[\hat{\lambda}_0, \hat{\lambda}_1, \hat{\lambda}_2], \quad \hat{\lambda}_l \geq 0, \quad l = 0, 1, 2, \quad \phi(t) = \{\phi_{x,y}(t) : x \in \mathcal{X}, y \in \mathcal{Y}\},$$

не равные одновременно нулю, т.е.

$$\sum_{l=0}^2 |\hat{\lambda}_l| + \max_{t \in [0, T]} \sum_{(x,y) \in \mathcal{Z}} |\phi_{x,y}(t)| > 0,$$

и такие что:

1) в каждый момент t стратегия $\hat{\mathbf{u}}(t)$ является точкой минимума

$$\hat{\mathbf{u}}(t) \in \arg \min_{\mathbf{v} \in \mathbf{U}} \langle \pi(t), \bar{W}(t, \phi(t), \mathbf{v}, \hat{\lambda}) \rangle, \quad (46)$$

где функция $\bar{W}(\cdot)$ имеет вид

$$\bar{W}(t, \phi, \mathbf{v}, \lambda) = \bar{A}(t, \mathbf{v})\phi + \{\bar{F}(t, \mathbf{v})\}^* \lambda, \quad t \in [0, T], \quad \mathbf{v} \in \mathbf{U}, \quad \phi \in \mathbb{R}^{\mathcal{Z}}, \quad \lambda \in \mathbb{R}^3, \quad (47)$$

а пара $\pi(t), \phi(t)$ на отрезке $[0, T]$ образует решение двухточечной краевой задачи

$$\begin{cases} \dot{\pi}(t) = \{\bar{A}(t, \hat{\mathbf{u}}(t))\}^* \pi(t), & \pi(0) = \pi^o, \\ \dot{\phi}(t) = -\bar{W}(t, \phi(t), \hat{\mathbf{u}}(t), \hat{\lambda}), & \phi(T) = 0; \end{cases} \quad (48)$$

2) выполнено условие дополняющей нежесткости

$$\hat{\lambda}_l J_l(\pi, \hat{\mathbf{u}}) = 0, \quad l = 1, 2. \quad (49)$$

Рассмотрим отдельно задачу минимизации (46). Для краткости вместо функций $\pi(t), \phi(t)$ возьмем независимые переменные $\pi, \phi \in \mathbb{R}^{\mathcal{Z}}$. Тогда с учетом $\mathbf{v} = \{m_x, v_y : x \in \mathcal{X}, y \in \mathcal{Y}\}$ и обозначения (16) минимизируемая в (46) функция принимает вид

$$\begin{aligned} \sum_{x,y \in \mathcal{Z}} \pi_{x,y} W(t, \phi_{x,y}, (m_x, v_y), \lambda) = \\ = \sum_{x,y \in \mathcal{Z}} \pi_{x,y} \left\{ \sum_{x',y' \in \mathcal{Z}} a_{(x,y),(x',y')}(t, m_x, v_y) \phi_{x',y'} + \sum_{l=0}^2 \lambda_l f_{l,(x,y)}(t, m_x, v_y) \right\}. \end{aligned}$$

Данное выражение с помощью представления (21) можно переписать в виде

$$\sum_{x>0} m_x \left\{ \sum_{y<N} \pi_{x,y} (v_y (\phi_{x-1,y+1} - \phi_{x,y}) + \lambda_2) + \pi_{x,N} \lambda_2 \right\} + \dots$$

К сожалению, явная минимизация данного выражения невозможна. Его билинейная структура гарантирует лишь то, что искомый минимум по $\mathbf{v}$ будет достигаться в вершинах параллелепипеда $\mathbf{U}$. Однако такой перебор реализуем лишь при малом количестве состояний. Кроме того, для нахождения $\mathbf{u}(t)$ требуются значения вероятностей $\pi_{x,y}(t)$ и двойственных

переменных $\phi_{x,y}(t)$. Поэтому двухточечная краевая задача (48) не может быть сведена к двум отдельным задачам Коши. Еще одно препятствие, возникающее на пути применения теоремы 3, состоит в том, что нельзя гарантировать положительность коэффициента λ_0 . Таким образом, в отличие от уравнений динамического программирования, представленных в теореме 1, условия (46)–(49), определяемые принципом максимума, не позволяют сформулировать практически реализуемый алгоритм оптимального управления с учетом ограничений.

Для синтеза численного метода оптимизации рассмотрим класс стационарных управлений

$$U(t) = (m_{X(t)}, v_{Y(t)}), \quad \mathbf{u} = \{m_x, v_y : x \in \mathcal{X}, y \in \mathcal{Y}\}, \quad (50)$$

где стратегия $\mathbf{u}$ пробегает множество $\mathbf{U}$. Класс стационарных управлений обозначим $\mathcal{S}$. Поскольку $\mathbf{u}$ представляет собой точку в конечномерном пространстве, для функционалов (41) удобно использовать обозначение в виде $J_l(\mathbf{u})$, опуская зависимость от распределения $\pi(t)$, однозначно определяемого из системы дифференциальных уравнений

$$\dot{\pi}(t) = \{\bar{A}(t, \mathbf{u})\}^* \pi(t), \quad \pi(0) = \pi^o. \quad (51)$$

Для численного решения задачи

$$J_0(\mathbf{u}) \rightarrow \min_{\mathbf{u} \in \mathbf{U}} : \quad J_1(\mathbf{u}) \leq 0, \quad J_2(\mathbf{u}) \leq 0, \quad (52)$$

воспользуемся методом модифицированной функции Лагранжа [19, 22]. Модифицированная функция Лагранжа для задачи (52) имеет вид

$$M^\varepsilon(\mathbf{u}, \lambda) = J_0(\mathbf{u}) + \frac{1}{2} \sum_{l=1}^2 \varepsilon_l ((\lambda_l + J_l(\mathbf{u})/\varepsilon_l)_+^2 - \lambda_l^2),$$

где $\lambda = \text{col}[1, \lambda_1, \lambda_2]$ — вектор множителей Лагранжа, $\varepsilon_1, \varepsilon_2$ — параметры регуляризации, $(\cdot)_+$ — положительная часть числа.

Указанный численный метод состоит в последовательном применении градиентного спуска (подъема) по переменным $\mathbf{u}$ и λ соответственно. Однако в приведенном ниже алгоритме шаг градиентного метода по стратегии $\mathbf{u}$ заменен численным решением задачи минимизации. Детали реализации этого шага описаны после формулировки алгоритма.

Алгоритм 1. Задать уровень погрешности $\delta_\lambda > 0$, скорость убывания нормы градиента $\delta_N \in (0, 1)$ и установить начальные значения: параметров регуляризации $\varepsilon_1^{(0)}, \varepsilon_2^{(0)} > 0$; нормы градиента $N_\lambda^{(0)} = +\infty$; стратегии $\mathbf{u}^{(0)} = \{m_k, v_q : k \in \mathcal{X}, q \in \mathcal{Y}\}$, где $m_k = \underline{m}$, $v_q = \bar{v}$; вектора множителей $\lambda^{(0)} = \text{col}[1, 0, 0]$; номера итерации $s = 0$.

1. Получить стратегию $\mathbf{u}^{(s+1)}$ как решение задачи

$$M^\varepsilon(\mathbf{u}, \lambda^{(s)}) \rightarrow \min_{\mathbf{u} \in \mathbf{U}}, \quad (53)$$

где $\varepsilon = (\varepsilon_1, \varepsilon_2)$, $\varepsilon_l = \varepsilon_l^{(s)}$, $l = 1, 2$.

2. Сделать градиентный шаг по λ

$$\lambda_l^{(s+1)} = \left(\lambda_l^{(s)} + J_l(\mathbf{u}^{(s+1)})/\varepsilon_l \right)_+, \quad l = 1, 2.$$

3. Вычислить норму градиента в полученной точке:

$$N_{\lambda}^{(s+1)} = \left| \nabla_{\lambda} M^{\varepsilon}(\mathbf{u}^{(s+1)}, \lambda^{(s+1)}) \right|,$$

где частные производные имеют вид

$$\frac{\partial M^{\varepsilon}(\mathbf{u}^{(s+1)}, \lambda^{(s+1)})}{\partial \lambda_l} = \begin{cases} J_l(\mathbf{u}^{(s+1)}), & \text{если } J_l(\mathbf{u}^{(s+1)}) \geq -\varepsilon_l \lambda_l^{(s+1)}, \\ -\varepsilon_l \lambda_l^{(s+1)}, & \text{иначе,} \end{cases} \quad l = 1, 2.$$

4. Закончить итерации, если выполнен критерий останова $N_{\lambda}^{(s+1)} \leq \delta_{\lambda}$.

5. Скорректировать значения параметров регуляризации

$$\varepsilon_l^{(s+1)} = \begin{cases} \varepsilon_l, & \text{если } N_{\lambda}^{(s+1)} \leq \delta_N N_{\lambda}^{(s)}, \\ \delta_N \varepsilon_l, & \text{иначе,} \end{cases} \quad l = 1, 2.$$

6. Увеличить s на единицу и перейти к шагу 1.

Решение вспомогательной задачи (53) находилось с помощью квазиньютоновского алгоритма [20], приспособленного для оптимизации гладкой выпуклой функции на координатном параллелепипеде. Для реализации этого алгоритма необходимо вычислять градиент $\nabla_{\mathbf{u}} M^{\varepsilon}(\mathbf{u}, \lambda)$, координаты которого имеют вид

$$\begin{aligned} \frac{\partial M^{\varepsilon}(\mathbf{u}, \lambda)}{\partial m_k} &= \frac{\partial J_0(\mathbf{u})}{\partial m_k} + \sum_{l=1}^2 \left(\lambda_l + \frac{1}{\varepsilon_l} J_l(\mathbf{u}) \right)_+ \frac{\partial J_l(\mathbf{u})}{\partial m_k}, \quad k \in \mathcal{X}, \\ \frac{\partial M^{\varepsilon}(\mathbf{u}, \lambda)}{\partial v_q} &= \frac{\partial J_0(\mathbf{u})}{\partial v_q} + \sum_{l=1}^2 \left(\lambda_l + \frac{1}{\varepsilon_l} J_l(\mathbf{u}) \right)_+ \frac{\partial J_l(\mathbf{u})}{\partial v_q}, \quad q \in \mathcal{Y}. \end{aligned}$$

Выражения для частных производных функционалов $J_l(\mathbf{u})$, $l = 0, 1, 2$ по стратегиям m_k, v_q , $k \in \mathcal{X}$, $q \in \mathcal{Y}$ следуют из их представления с учетом (14):

$$\begin{aligned} \frac{\partial J_l(\mathbf{u})}{\partial m_k} &= \int_0^T \left\{ \mathbf{I}\{k > 0, l = 2\} \sum_{y \in \mathcal{Y}} \pi_{k,y}(t) + \sum_{(x,y) \in \mathcal{Z}} f_{l,(x,y)}(m_x, v_y) \frac{\partial \pi_{x,y}(t)}{\partial m_k} \right\} dt, \\ \frac{\partial J_l(\mathbf{u})}{\partial v_q} &= \int_0^T \sum_{(x,y) \in \mathcal{Z}} f_{l,(x,y)}(m_x, v_y) \frac{\partial \pi_{x,y}(t)}{\partial v_q} dt, \end{aligned}$$

где распределение $\pi(t) = \{\pi_{x,y}(t)\}$ определяется из системы (51) и соответствует стратегии $\mathbf{u} = \{m_k, v_q, k \in \mathcal{X}, q \in \mathcal{Y}\}$.

Если продифференцировать уравнения системы (51) по управляющим параметрам, то указанная дифференциальная система дополняется новыми уравнениями для частных производных вероятностей состояний по интенсивности обработки m_k

$$\begin{aligned} \frac{\partial \dot{\pi}_{x,y}(t)}{\partial m_k} &= \sum_{(x',y') \in \mathcal{Z}} a_{(x',y'),(x,y)}(t, m_{x'}, v_{y'}) \frac{\partial \pi_{x',y'}(t)}{\partial m_k} + \\ &\quad + \mathbf{I}\{x = k - 1, y > 0\} v_{y-1} \pi_{k,y-1}(t) - \\ &\quad - \mathbf{I}\{x = k, y < N, k > 0\} v_y \pi_{k,y}(t), \quad \frac{\partial \pi_{x,y}(0)}{\partial m_k} = 0, \end{aligned}$$

и по вероятности приема v_q

$$\begin{aligned} \frac{\partial \pi_{x,y}(t)}{\partial v_q} = & \sum_{(x',y') \in \mathcal{Z}} a_{(x',y'),(x,y)}(t, m_{x'}, v_{y'}) \frac{\partial \pi_{x',y'}(t)}{\partial v_q} + \\ & + I\{x < M, y = q + 1\} m_{x+1} \pi_{x+1,q}(t) - \\ & - I\{y = q, q < N, x > 0\} m_x \pi_{x,q}(t), \quad \frac{\partial \pi_{x,y}(0)}{\partial v_q} = 0. \end{aligned}$$

Теперь полученных соотношений вполне достаточно для реализации численного решения вспомогательной задачи (53).

7. ОПТИМИЗАЦИЯ ПЕРЕДАЧИ ДАННЫХ В ДВУХАГЕНТНОЙ РОБОТОТЕХНИЧЕСКОЙ СИСТЕМЕ

Предположим, что в течение некоторого времени T ведется прием, передача и обработка данных в двухагентной робототехнической системе, образованной передатчиком на БПЛА и приемником базовой станцией. Входной поток данных является нестационарным и имеет интенсивность $\alpha(t)$, изображенную на рис. 2.

Параметры сети передачи данных между передатчиком и приемником были выбраны следующим образом:

$$\begin{aligned} T = 100, \quad A = \int_0^T \alpha(t) dt \approx 167, \quad \max_{t \in [0, T]} \alpha(t) = 2, \quad \nu = 2,5, \\ \underline{m} = 0,5, \quad \bar{m} = 4, \quad \underline{v} = 0,05, \quad \bar{v} = 1, \quad M = 10, \quad N = 15, \end{aligned}$$

где A — ожидаемое число пакетов во входном потоке.

Количество состояний рассматриваемой сети равно $(M + 1)(N + 1) = 176$. Для того чтобы стационарное распределение не было локализовано в состояниях, соответствующих перегрузке базовой станции, интенсивность обработки была взята больше интенсивности входного потока, т.е. $\nu > \alpha(t)$.

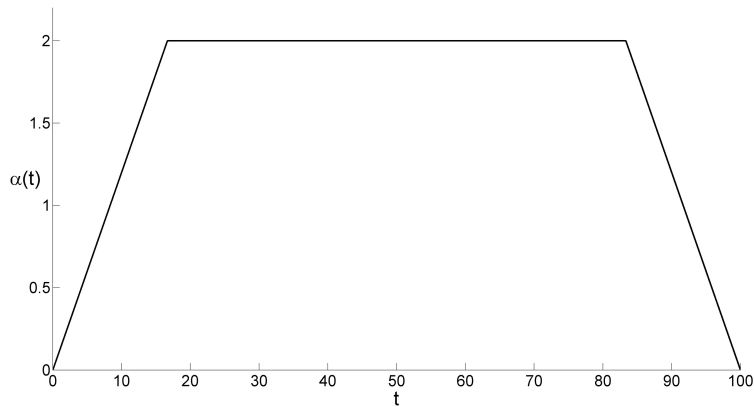


Рис. 2. Интенсивность входного потока $\alpha(t)$.

Рассмотрим следующую стационарную стратегию $\mathbf{u}^o = \{m_x^o, v^o\}$:

$$m_x^o = \underline{m} \quad \text{при} \quad x \leq 7 \quad \text{и} \quad m_x^o = \bar{m} \quad \text{при} \quad x > 7, \quad v^o = \bar{v} = 1. \quad (54)$$

Поскольку при использовании (54) реализуется только два варианта, назовем эту стратегию двухточечной.

Значения функционалов, описывающих полное время обработки данных и энергозатраты БПЛА при использовании двухточечной стратегии, оказались равны:

$$S(\mathbf{u}^o) = 5,7685, \quad E(\mathbf{u}^o) = 159,3577.$$

Назначим верхние границы $\bar{S}, \bar{E}$ из того расчета, чтобы они оказались примерно на 1% больше полученных значений:

$$\bar{S} = 5,8262, \quad \bar{E} = 160,9513.$$

При использовании двухточечной стратегии $\mathbf{u}^o$ среднее число потерянных пакетов, а также значения преобразованных функционалов (9), (10) равны

$$J_0(\mathbf{u}^o) = 8,3402, \quad J_1(\mathbf{u}^o) = -9,1331, \quad J_2(\mathbf{u}^o) = -1,5936.$$

С помощью теоремы 2 была найдена стратегия $\hat{\mathbf{u}}(t) = \{\hat{m}_{x,y}(t), \hat{v}_{x,y}(t) : x \in \mathcal{X}, y \in \mathcal{Y}\}$, оптимальная на классе централизованных нестационарных управлений. При решении двойственной задачи был использован квазиньютоновский алгоритм (его подробное описание см. в [17]).

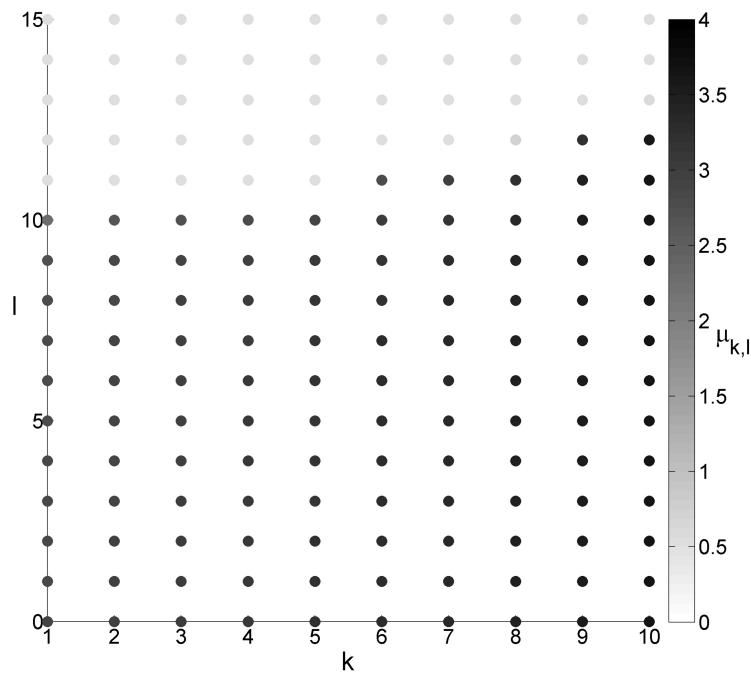


Рис. 3. Усредненная по времени оптимальная интенсивность обработки $\hat{m}_{x,y}$.

Оптимальная вероятность приема пакета оказалась тождественно равной единице: $\hat{v}_{x,y} \equiv 1$. На рис. 3 представлена усредненная по времени оптимальная интенсивность обработки $\hat{m}_{x,y}$, относительно которой можно сделать два вывода: увеличение числа пакетов x в СМО передатчика ведет к росту интенсивности $\hat{m}_{x,y}$; при приближении y к состоянию перегрузки базовой станции интенсивность $\hat{m}_{x,y}$ падает до минимально допустимой.

Для наглядного описания того, как стратегия $\hat{m}_{x,y}(t)$ эволюционирует во времени, на рис. 4 и 5 представлен результат ее усреднения по состояниям каждой из двух систем:

$$\hat{m}(t|X=x) = \mathbb{E}\{\hat{m}_{Z(t)}(t) | X(t)=x\}, \quad \hat{m}(t|Y=y) = \mathbb{E}\{\hat{m}_{Z(t)}(t) | Y(t)=y\}.$$

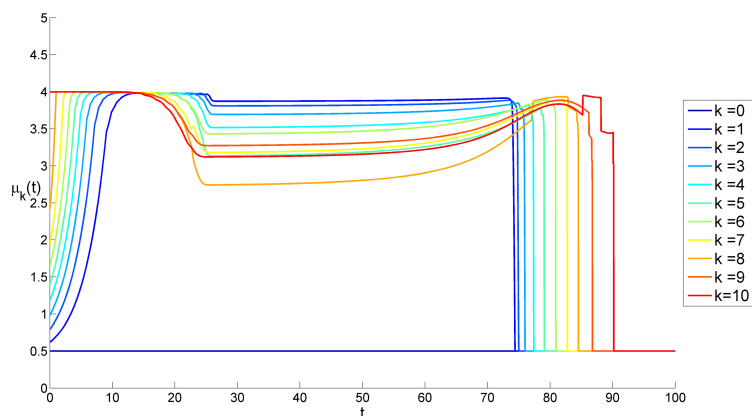


Рис. 4. Оптимальная интенсивность обработки, усредненная по состоянию базовой станции $\hat{m}(t|X=x)$.

Из рис. 4 можно сделать вывод, что стратегия $\hat{m}_{x,y}(t)$ меняется вслед за изменением интенсивности входного потока $\alpha(t)$. Сначала $\hat{m}_{x,y}(t)$ возрастает до верхней границы $\bar{m}$, причем для более загруженных состояний x , скорость роста выше. Затем оптимальная интенсивность обработки остается почти постоянной до смены режима входного потока, после чего скачком оказывается равной нижней границе $\underline{m}$. Для более загруженных состояний x переход на минимальный уровень происходит позже.

Рис. 5 подтверждает зависимость стратегия $\hat{m}_{x,y}(t)$ от эволюции входного потока. Здесь также можно выделить три периода и соответствующие им три уровня интенсивности $\hat{m}_{x,y}(t)$: максимальный, промежуточный и минимальный. Однако зависимость от загрузки базовой станции иная. Для трех наиболее загруженных состояний y оптимальная интенсивность обработки оказывается заметно ниже, чем для других состояний.

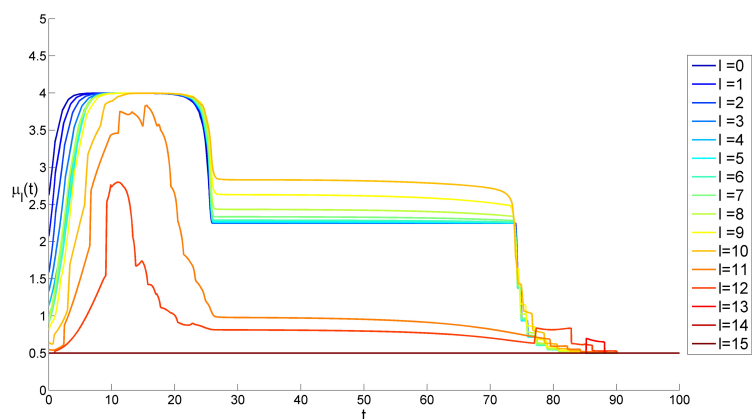


Рис. 5. Оптимальная интенсивность обработки, усредненная по состоянию СМО передатчика $\hat{m}(t|Y=y)$.

Необходимо отметить, что на вид графиков, изображенных на рис. 4 и 5 влияют соответствующие вероятности состояний $\hat{\pi}_{x,y}(t)$. На рис. 6, 7 для СМО передатчика и базовой станции представлены частные распределения состояний, соответствующие оптимальной стратегии $\hat{\mathbf{u}}(t)$:

$$\hat{\pi}_x^X(t) = \sum_{y \in \mathcal{Y}} \hat{\pi}_{x,y}(t), \quad \hat{\pi}_y^Y(t) = \sum_{x \in \mathcal{X}} \hat{\pi}_{x,y}(t).$$

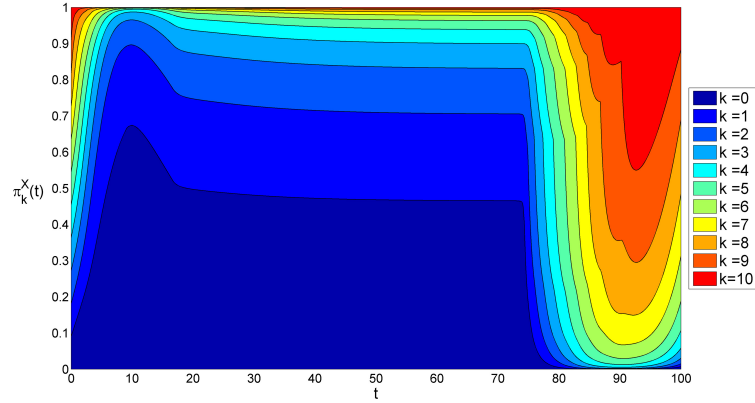


Рис. 6. Вероятности состояний СМО передатчика $\hat{\pi}_x^X(t)$ при использовании оптимального централизованного управления $\hat{\mathbf{u}}$.

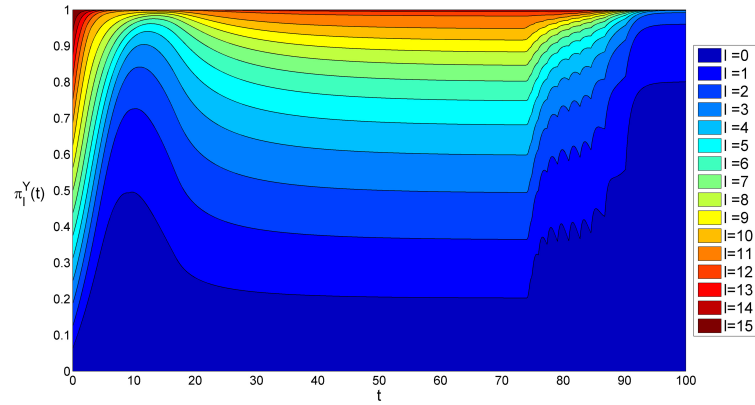


Рис. 7. Вероятности состояний СМО базовой станции $\hat{\pi}_y^Y(t)$ при использовании оптимального централизованного управления $\hat{\mathbf{u}}$.

Решение $\tilde{\mathbf{u}} = \{\tilde{m}_x, \tilde{v}_y: x \in \mathcal{X}, y \in \mathcal{Y}\}$ задачи оптимального управления на классе стационарных децентрализованных стратегий было получено численно с помощью алгоритма 1. Как и в случае централизованного управления, оптимальной стратегией вероятности приема оказывается $\tilde{v}_y \equiv 1$.

Зависимость оптимальной интенсивности обработки $\tilde{m}_x$ от загрузки передатчика x представлена на рис. 8. Она почти полностью повторяет двухточечную стратегию (54): отличие

есть только при $x = 7$. В сравнении со стратегией, построенной по полной информации, децентрализованная стратегия $\tilde{m}_x$ в условиях отсутствия данных о состоянии базовой станции демонстрирует значительно более активную реакцию на изменение загрузки передатчика.

Необходимо отметить, что двухточечная стратегия была взята в качестве начального приближения в алгоритме 1. С целью проверки нечувствительности итерационной схемы к выбору начальной точки были рассмотрены также и другие варианты начального приближения. Поскольку эти расчеты привели к практически идентичным результатам, вид оптимальной децентрализованной стратегии $\tilde{m}_x$ можно признать вполне обоснованным.

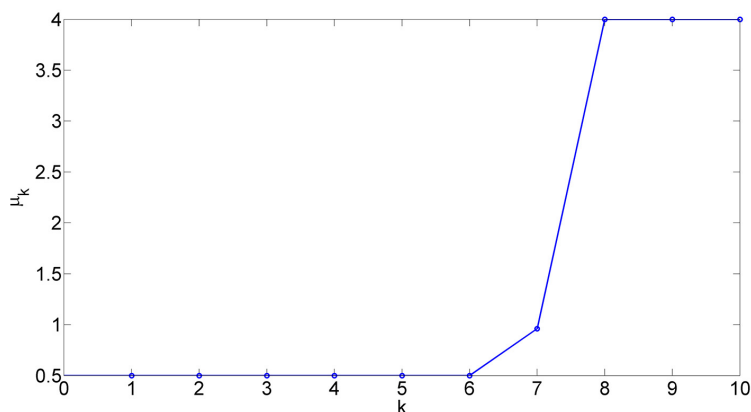


Рис. 8. Оптимальная децентрализованная стратегия интенсивности обработки $\tilde{m}_x$.

На рис. 9, 10 представлены частные распределения состояний СМО для передатчика и базовой станции при использовании оптимального децентрализованного управления $\tilde{\mathbf{u}}$. Для СМО передатчика состояния повышенной загрузки оказываются наиболее вероятными, что сильно отличает рассматриваемый случай от ситуации, в которой использовалась стратегия $\hat{m}_{x,y}(t)$ с полной информацией (ср. с рис. 6). Вместе с тем распределение загрузки базовой станции в этих двух случаях отличается мало (ср. с рис. 7).

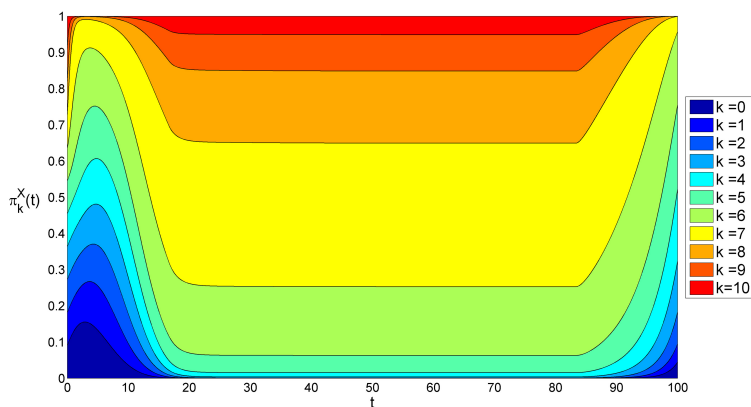


Рис. 9. Вероятности состояний СМО передатчика $\pi_k^x(t)$ при использовании оптимального децентрализованного управления $\tilde{\mathbf{u}}$.

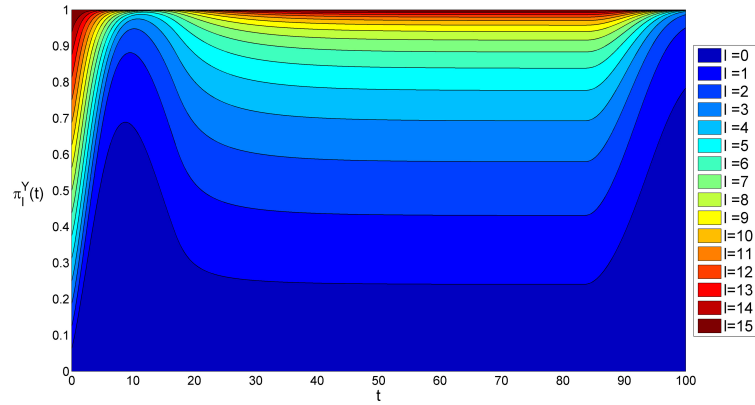


Рис. 10. Вероятности состояний базовой станции $\tilde{\pi}_y^Y(t)$ при использовании оптимального децентрализованного управления $\tilde{\mathbf{u}}$.

Не выходя за рамки класса (50), оптимальную децентрализованную стратегию $\tilde{\mathbf{u}}$ можно сравнить со следующей усредненной стратегией:

$$\check{\mathbf{u}} = \{\check{m}_x, 1\}: \quad \check{m}_x = \frac{1}{T} \int_0^T \mathbb{E}\{\hat{m}_{Z(t)}(t) \mid X(t) = x\} dt, \quad x \in \mathcal{X}, \quad (55)$$

где усреднению (по времени и по состоянию базовой) подвергается оптимальная интенсивность обработки $\hat{m}_{x,y}(t)$. Указанный способ приводит к децентрализованной стратегии, изображенной на рис. 11. Полученная интенсивность $\check{m}_x$ оказывается практически постоянной, что кардинально отличается от оптимальной стратегии $\tilde{m}_x$ (ср. с рис. 8).

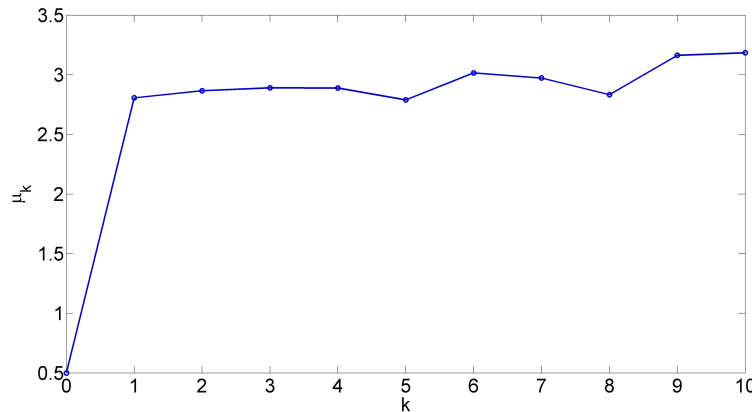


Рис. 11. Интенсивность обработки $\check{\mu}_x$ при использовании усредненной стратегии.

В таблице 1 представлены значения функционалов на следующих четырех управлениях: $\mathbf{u}^o$ — двухточечная стратегия (54); $\hat{\mathbf{u}}(t)$ — оптимальная стратегия, построенная в соответствии с теоремой 2; $\tilde{\mathbf{u}}$ — оптимальная децентрализованная стратегия, полученная в результате применения алгоритма 1; $\check{\mathbf{u}}$ — усредненная стратегия, найденная по правилу (55).

Обе численно синтезированные оптимальные стратегии $\hat{\mathbf{u}}(t)$ и $\tilde{\mathbf{u}}$ удовлетворяют заданным ограничениям. При этом лимит энергетических ресурсов оказался исчерпан практически пол-

Таблица 1. Значения функционалов для нескольких управлений

$\mathbf{u}$	$J_0(\mathbf{u})$	$J_1(\mathbf{u})$	$J_2(\mathbf{u})$	$S(\mathbf{u})$	$E(\mathbf{u})$
$\mathbf{u}^o$	8,3402	-9,1331	-1,5936	5,7685	159,3577
$\hat{\mathbf{u}}$	3,5164	-424,0999	-0,0078	3,2267	160,9435
$\tilde{\mathbf{u}}$	7,3230	-30,8468	-0,0007	5,6326	160,9506
$\check{\mathbf{u}}$	1,0189	-494,6514	11,6171	2,8400	172,5684
Ограничения:		0	0	5,8262	160,9513

ностью: $J_2(\mathbf{u}) \approx 0$ и $E(\mathbf{u}) \approx \bar{E}$. Однако ограничение на время прохождения пакета по всей сети передачи данных оказалось выполненным с большим запасом, так как $J_2(\mathbf{u}) \ll 0$ или $S(\mathbf{u}) \ll \bar{S}$.

По критерию минимума среднего числа потерь $J_0(\mathbf{u})$ оптимальная стратегия $\hat{\mathbf{u}}(t)$, использующая полную информацию, оказалась значительно лучше, как двухточечной стратегии $\mathbf{u}^o$, так и децентрализованной стратегии $\tilde{\mathbf{u}}$, оптимальной в своем классе. Вместе с тем, усредненная стратегия $\check{\mathbf{u}}$ позволяет терять наименьшее число пакетов, однако ценою этого стало серьезное превышение энергозатрат: $J_2(\check{\mathbf{u}}) \gg 0$ и $E(\check{\mathbf{u}}) \gg \bar{E}$.

В заключение сравним полученные результаты с выводами, сделанными в [3] относительно похожей модели. Согласно результатам [3] вероятность приема пакета не должна быть равна единице тождественно (как указано выше): она должна падать до нижней границы при сильной загрузке базовой станции. Следствием такого различия является разный выбор функционалов. Дело в том, что в [3] использовались два функционала, отвечающих за время прохождения пакета в СМО передатчика и базовой станции. Для оптимальных стратегий, найденных в [3], ограничение на среднее время полного обслуживания в первой системе оказалось выполнено с запасом, зато лимит времени пребывания во второй системе был исчерпан полностью за счет активного управления вероятностью приема. В модели же, исследованной выше, эти функционалы были объединены в один, поэтому управление всей сетью проводилось только за счет выбора интенсивности обработки в первой системе.

8. ЗАКЛЮЧЕНИЕ

В статье рассмотрена управляемая двухфазная система массового обслуживания, включающая две одноканальные системы с конечной очередью и предусматривающая механизм блокировки первого сервера. Данная сеть массового обслуживания описывается управляемым марковским процессом, оптимизация которого проводится на конечном промежутке времени на основе минимизации среднего числа потерь с учетом ограничений на время полного обслуживания и энергозатраты первой системы. Разработаны два метода для определения оптимальных управлений на классе централизованных и децентрализованных стратегий. Получены явные выражения для оптимальной интенсивности обслуживания и вероятности приема пакета в случае оптимизации по свертке критериев на классе централизованных стратегий. Оптимизация с учетом ограничений проведена на основе метода двойственной оптимизации. Для случая децентрализованного управления получены необходимые условия оптимальности. Оптимизация децентрализованного управления реализована на классе стационарных стратегий за счет разработанной итерационной процедуры. Приведены результаты численного эксперимента, демонстрирующие характерный вид оптимальных стратегий в задаче управления потоком данных в двухагентной робототехнической системе.

СПИСОК ЛИТЕРАТУРЫ

1. Дудин С. А., Дудина О. С. Модель функционирования центра информационной и технической поддержки как двухфазная система массового обслуживания // *Проблемы передачи информации*. 2013, т. 49, № 1, стр. 66–82.
2. Вишнеvский В. М., Ларионов А. А., Семенова О. В. Оценка производительности высокоскоростной беспроводной тандемной сети с использованием каналов сантиметрового и миллиметрового диапазона радиоволн в системах управления безопасностью дорожного движения // *Проблемы управления*. 2013, № 4, стр. 50–56.
3. Кузнецов Н. А., Мясников Д. В., Семенихин К. В. Оптимальное управление передачей данных в мобильной двуагентной робототехнической системе // *Информационные процессы*. 2016, т. 16, № 2, стр. 137–151.
4. Gomez-Corral A. A tandem queue with blocking and Markovian arrival process // *Queueing Systems*. 2002, vol. 41, pp. 343–370.
5. Бройер Л., Дудин А. Н., Клименок В. И., Царенков Г. В. Двухфазная система $BMAP/G/1/N \rightarrow PH/1/M - 1$ с блокировкой // *Автоматика и телемеханика*. 2004, № 1, стр. 117–130.
6. Bocharov P. P., Manzo R., Pechinkin A. V. Analysis of a two-phase queueing system with a Markov arrival process and blocking // *Journal of Mathematical Sciences*. 2006, vol. 132, no. 5, pp. 578–589.
7. Rosberg Z., Varaiya P. P., Walrand J. Optimal control of service in tandem queues // *IEEE Trans. Automat. Control*. 1982, vol. 27, no. 3, pp. 600–610.
8. Hajek B. Optimal control of a two interacting service stations // *IEEE Trans. Automat. Control*. 1984, vol. 29, no. 6, pp. 491–499.
9. Stidham S., Weber R. A survey of Markov decision models for control of networks of queues // *Queueing Systems*. 1993, vol. 13, no. 1, pp. 291–314.
10. Kitaev M. Y., Rykov V. V. *Controlled queueing systems*. Boca Raton: FL: CRC, 1995.
11. Koole G. Convexity in tandem queues // *Probability in the Engineering and Informational Sciences*. 2004, vol. 18, no. 1, pp. 13–31.
12. Silva D. F., Zhang B., Ayhan H. Optimal admission control for tandem loss systems with two stations // *Operations Research Letters*. 2013, vol. 41, no. 4, pp. 351–356.
13. Zhang B., Ayhan H. Optimal admission control for tandem queues with loss // *IEEE Trans. Automat. Control*. 2013, vol. 58, no. 1, pp. 163–167.
14. Altman E. *Constrained Markov decision processes*. Boca Raton, FL: Chapman & Hall/CRC, 1999.
15. Миллер Б. М., Миллер Г. Б., Семенихин К. В. Методы синтеза оптимального управления марковским процессом с конечным множеством состояний при наличии ограничений // *Автоматика и телемеханика*. 2011, № 2, стр. 111–130.
16. Миллер Б. М., Миллер Г. Б., Семенихин К. В. Регуляризация задачи оптимального управления марковским процессом с конечным числом состояний при наличии ограничений // *Автоматика и телемеханика*. 2016, № 9, стр. 96–123.
17. Мясников Д. В., Семенихин К. В. Управление параметрами одноканальной системы массового обслуживания при наличии ограничений // *Известия РАН. Теория и системы управления*. 2016, № 1, стр. 66–85.
18. Гнеденко Б. В., Коваленко И. Н. *Введение в теорию массового обслуживания*. М.: Наука, 1987.
19. Поляк Б. Т. *Введение в оптимизацию*. М.: Наука, 1980.
20. Kim D., Sra S., Dhillon I. S. Tackling box-constrained optimization via a new projected quasi-Newton approach // *SIAM J. Scientific Computing*. 2010, vol. 32, pp. 3548–3563.
21. Тихонов А. Н., Арсенин В. Я. *Методы решения некорректных задач*. М.: Наука, 1979.

22. Bertsekas D. P. *Nonlinear Programming. 2nd ed.* Belmont, MA: Athena Scientific, 1999.

Two-phase queueing system optimization and its applications to data transmission control for two-agent robot system

N. A. Kuznetsov, D. V. Myasnikov, K. V. Semenikhin

We consider the controlled tandem queueing network, comprised of two single-server systems with finite buffers. This queueing network allows blocking, i.e. if the second server's buffer is full, the first one is blocked and stops service. The first station receives non-stationary Poisson flow of packets, with service rate being one of controlled parameters. If the first buffer is full, any incoming packet is dropped. The second station allows to control the packet acceptance probability (decrease in this probability results in increase of packet delay on the first station). The queueing network is described by the controlled Markov process. Optimization is performed on a finite time interval by minimizing the average number of dropped packets with constraints on the full service time in the network and the energy consumption in the first system. We propose numerical schemes for optimization of non-stationary data flow control in two-agent robot systems.

KEYWORDS: optimal control, two-phase queueing system, controlled Markov process, multi-agent systems, constrained optimization.