

Оптимальное управление передачей данных в мобильной двухагентной робототехнической системе¹

Н. А. Кузнецов, Д. В. Мясников, К. В. Семенихин

*Институт радиотехники и электроники им. В.А. Котельникова РАН, Москва, Россия,
Московский физико-технический институт (государственный университет), Москва, Россия*

Поступила в редколлегию 19.05.2016

Аннотация—Разработан метод оптимизации параметров инфотелекоммуникационной сети, отвечающей за процесс обработки и передачи данных в мобильной двухагентной робототехнической системе. Первый агент принимает простейший поток пакетов (блоков информации) и пересылает их второму агенту для последующей обработки; первый агент теряет входной пакет, если его буфер переполнен; второй агент не допускает переполнения своего буфера, сообщая об отказе и запрашивая повторную пересылку пакета. Задача минимизации потерь входных пакетов решена в стационарном режиме с учетом ограничений на время обработки и энергозатраты за счет управления скоростью пересылки и вероятностью отказа. Определены условия, при которых значения целевого функционала и функционалов, задающих ограничения, образуют выпуклое множество. Численный синтез управления реализован на основе метода модифицированной функции Лагранжа.

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

1. ВВЕДЕНИЕ

Мобильные робототехнические системы функционируют в условиях ограниченных энергетических и вычислительных ресурсов, пополнение которых невозможно в режиме реального времени. Поэтому для успешной эксплуатации таких систем необходимо эффективное перераспределение ее ресурсов с целью учета разнонаправленных требований к качеству выполнения поставленной задачи. Удовлетворение этих требований возможно за счет оптимизации нескольких показателей, таких как: энергопотребление, эффективность сбора данных, устойчивость связи, скорость передачи информации и т.д. Постановки и методики решения данной оптимизационной задачи для систем, образованных беспилотными летательными аппаратами и наземными средствами беспроводной связи, представлены в работах [1–3].

Если робототехническая система состоит из беспилотного летательного аппарата (БПЛА) и автономной базовой станции (АБС) и выполняет задачи слежения за наземной обстановкой, то основной целью управления данной системой является скорость ее реакции на критические изменения в состоянии контролируемых объектов. Для достижения этой цели необходимо максимизировать надежность приема входного потока информации. Поскольку передача и обработка видео данных связаны с серьезными телекоммуникационными и вычислительными нагрузками, для предотвращения коллизий система массового обслуживания АБС должна предоставлять раздельный доступ к своим ресурсам. Для сигнализации о переполнении буфера принимающего устройства обычно используют различные варианты алгоритма RED (Random Early Detection), приводящие к случайному отклонению некоторой доли входного

¹ Работа выполнена при поддержке Российского научного фонда (проект № 16-11-00063).

трафика [4]. При реализации этого механизма на базовой станции возникает необходимость повторной отсылки видео пакетов, что вызывает дополнительную нагрузку на телекоммуникационный узел БПЛА, а это, в свою очередь, повышает расход электроэнергии, потребляемой беспилотником для поддержания радиосвязи с базовой станцией.

Таким образом, описанная проблема может быть поставлена как задача минимизации вероятности потерь при приеме входной информации с учетом ограничений на средние показатели, описывающие время пересылки пакета, время его дальнейшей обработки и объем использованных энергоресурсов.

В данной работе эта сеть передачи данных рассматривается как тандем, т.е. последовательное соединение двух одноканальных систем массового обслуживания (СМО) с конечной очередью и простейшим входным потоком. Каждая из этих систем управляется отдельно и потому выступает в роли агента, самостоятельно принимающего решения. Первый агент (беспилотник) выбирает скорость передачи данных μ , иначе говоря, интенсивность обслуживания. Второй агент (базовая станция) назначает вероятность ρ , с которой происходит отказ в приеме пакета от первого агента. Каждый агент управляет выбором соответствующего параметра (μ или ρ) без учета состояния другого агента. Полученная сеть массового обслуживания изучается в стационарном режиме как управляемый марковский процесс, состояние которого определяется парой (ξ, η) , где ξ и η равно числу заявок в первой и второй СМО соответственно.

Необходимо отметить, что теории управляемых марковских процессов посвящено большое количество публикаций, среди которых отметим классическую монографию [5]. Методология условной оптимизации указанных процессов, рассматриваемых в стационарном режиме, описана в книге [6]. Применительно к СМО одна из первых задач данного типа была решена в работе [7]. Методика численного синтеза оптимального управления нестационарным многокомпонентным марковским процессом разработана в [8].

Основной теоретический результат, полученный в настоящей статье, состоит в определении достаточных условий, обеспечивающих выпуклость множества достижимости. Это множество по определению состоит из всех точек, координатами которых выступают целевой функционал и функционалы, задающие ограничения, при произвольном выборе управления. Благодаря этому факту проблема оптимального управления может быть преобразована в эквивалентную задачу выпуклого программирования.

Для численного синтеза оптимальных стратегий была разработана итерационная процедура, основанная на методе модифицированной функции Лагранжа. По результатам проведенных численных экспериментов установлен характерный вид оптимальных стратегий.

2. ОПИСАНИЕ СИСТЕМЫ

В работе изучается робототехническая система, состоящая из двух самостоятельно действующих агентов — БПЛА и АБС. Сеть передачи данных между этими агентами изображена на рис. 1. Будем рассматривать эту схему как сеть массового обслуживания типа тандем — последовательное соединение двух одноканальных систем. На вход СМО первого агента поступает пуассоновский поток пакетов, т.е. небольших порций данных, которые требуется передать на обработку в СМО второго агента. Интенсивность входного потока заранее известна и равна $\alpha > 0$.

Первая СМО содержит очередь на $M - 1$ мест для пакетов, ожидающих отправки на АБС, и прибор обслуживания, осуществляющий эту отсылку со скоростью μ . Если все места в очереди заняты, то вновь приходящий пакет теряется, что является крайне нежелательным событием. Для сокращения потерь пакетов первый агент может управлять интенсивностью μ , выбирая его в зависимости от количества пакетов, находящихся в системе.

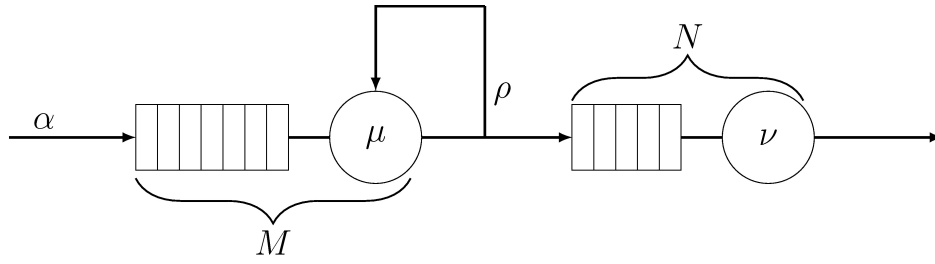


Рис. 1. Схема сети передачи данных от БПЛА к АБС.

Вторая СМО также содержит конечную очередь, максимально включающую $N - 1$ пакетов, и прибор обслуживания с фиксированной интенсивностью обработки ν . Чтобы исключить потери пакетов при передачи между системами, достаточно обеспечить отсутствие переполнения в очереди второй СМО. Это становится возможным за счет выбора дополнительного параметра — вероятности отказа ρ . Если происходит отказ, то пакет, подлежащей пересылке во вторую систему, остается в первой. Поэтому итоговая скорость пересылки, иначе говоря, интенсивность обслуживания в первой СМО, будет равна $\mu(1 - \rho)$. Тем самым вероятность ρ рассматривается как управляемый параметр, чье значение назначает второй агент.

Для описания состояния рассматриваемой сети в текущий момент времени t введем двухкомпонентный случайный процесс $(\xi(t), \eta(t))$, такой что $\xi(t)$ и $\eta(t)$ обозначают число пакетов в первой и второй системе соответственно. Тем самым $\xi(t)$ — число пакетов, готовящихся к отправке на базовую станцию телекоммуникационной системой БПЛА, а $\eta(t)$ — число пакетов, обрабатываемых на самой АБС.

Если считать, что времена обслуживания на серверах обеих СМО образуют набор взаимно независимых экспоненциальных случайных величин, а параметры μ и ρ суть константы, то пара $(\xi(t), \eta(t))$ образует однородный марковский процесс с множеством состояний

$$\mathbf{N} = \mathbf{N}_1 \times \mathbf{N}_2, \quad \mathbf{N}_1 = \{0, 1, \dots, M\}, \quad \mathbf{N}_2 = \{0, 1, \dots, N\}$$

и интенсивностями переходов

$$\begin{cases} a_{(k,l),(k,l-1)}(\mu, \rho) = \nu, & l > 0, \\ a_{(k,l),(k-1,l+1)}(\mu, \rho) = \mu(1 - \rho), & k > 0, l < N, \\ a_{(k,l),(k+1,l)}(\mu, \rho) = \alpha, & k < M, \end{cases} \quad (1)$$

где $a_{(k,l),(m,n)}(\mu, \rho)$ обозначает интенсивность перехода из состояния (k, l) в (m, n) при $(m, n) \neq (k, l)$.

Если скорость пересылки и вероятность отказа являются функциями текущих состояний первой и второй СМО соответственно, то процесс $(\xi(t), \eta(t))$ также будет однородным марковским, но для определения его интенсивностей необходимо в (1) положить $\mu = \mu_k$, $\rho = \rho_l$, где $U = \{(\mu_k, \rho_l) : (k, l) \in \mathbf{N}\}$ — заданный набор чисел. Этот набор определяет стратегию управления рассматриваемой двухагентной системой.

Введем множество допустимых стратегий управлений U как

$$\mathcal{U} = \mathcal{U}_\mu \times \mathcal{U}_\rho, \quad (2)$$

где множества

$$\mathcal{U}_\mu = \left\{ \{\mu_k\}_{k=0, \overline{M}} : \mu_k \in [\mu_{min}, \mu_{max}] \right\}, \quad \mathcal{U}_\rho = \left\{ \{\rho_l\}_{l=0, \overline{N}} : \rho_l \in [\rho_{min}, \rho_{max}] \right\},$$

содержат допустимые управления интенсивностью пересылки и вероятностью отказа, подчиняющихся нижним и верхним границам $0 \leq \rho_{\min} \leq \rho_{\max} < 1$ и $0 < \mu_{\min} \leq \mu_{\max}$.

3. ПОСТАНОВКА ЗАДАЧИ

Качество работы рассматриваемой инфокоммуникационной сети будем описывать в стационарном режиме с использованием нескольких функционалов.

Зависимость от времени для процесса $(\xi(t), \eta(t))$ далее опускается, поскольку он будет рассматриваться в стационарном режиме.

Начнем с самого важного функционала, который определяет надежность передачи информации. В силу конечности буфера может возникнуть ситуация, когда поток данных, поступающих на БПЛА, невозможно принять. Объем потерянных данных необходимо сделать как можно меньше. Поэтому соответствующий функционал следует определить, как долю потерянного входного трафика:

$$J_0 = P\{\xi = M\}. \quad (3)$$

Следующий функционал характеризует оперативность передачи данных. Это особенно актуально в ситуации, когда управление БПЛА осуществляется оператором на основе данных, которые должны поступать на АБС в режиме реального времени. Согласно формуле Вальда для определения среднего времени полного обслуживания в системе $M_a|M_b|1$ необходимо прибавить к среднему числу заявок единицу и поделить эту сумму на интенсивность обслуживания. Учитывая то, что интенсивность μ_k зависит от состояния первой системы, а отклонение пакетов при пересылке на вторую систему происходит с вероятностью ρ_l , получим функционал вида

$$J_1 = E\left\{\frac{\xi + 1}{\mu_\xi(1 - \rho_\eta)}\right\}. \quad (4)$$

Среднее время полного обслуживания во второй СМО описывается функционалом

$$J_2 = \frac{E\eta + 1}{\nu}. \quad (5)$$

Важной характеристикой работы телекоммуникационной системы, установленной на БПЛА, является ее энергоемкость. При слишком интенсивной передаче данных запас аккумуляторных батарей беспилотника может быстро исчерпаться, в результате чего аппарат не выполнит запланированную миссию. Будем считать, что объем потребляемой энергии пропорционален интенсивности отправки пакетов μ_ξ , а во время простоя энергия не расходуется. Тогда соответствующий функционал будет равен

$$J_3 = E\{\mu_\xi I\{\xi > 0\}\}. \quad (6)$$

Стационарную вероятность нахождения сети в состоянии (k, l) обозначим $\pi_{k,l}$. Тогда с учетом условия нормировки

$$\sum_{(k,l) \in N} \pi_{k,l} = 1$$

систему уравнений баланса для управляемого марковского процесса (ξ, η) можно представить в виде

$$\sum_{(k,l) \in N} a_{(k,l),(i,j)}(\mu_k, \rho_l) \pi_{k,l} = 0, \quad (i, j) \in N, \quad (7)$$

где обозначено

$$a_{(k,l),(k,l)}(\mu, \rho) = - \sum_{(m,n) \in \mathbb{N} \setminus \{(k,l)\}} a_{(k,l),(m,n)}(\mu, \rho). \quad (8)$$

Теперь для функционалов (3)–(6) можно указать их явную зависимость от управления $U \in \mathcal{U}$ и стационарного распределения $\pi = \{\pi_{k,l}\}_{(k,l) \in \mathbb{N}}$:

$$\begin{aligned} J_0(\pi, U) &= \mathbb{P}\{\xi = M\} = \sum_{l \in \mathbb{N}_2} \pi_{M,l}, \\ J_1(\pi, U) &= \mathbb{E}\left\{ \frac{\xi + 1}{\mu_\xi(1 - \rho_\eta)} \right\} = \sum_{(k,l) \in \mathbb{N}} \frac{k + 1}{\mu_k(1 - \rho_l)} \pi_{k,l}, \\ J_2(\pi, U) &= \frac{\mathbb{E}\eta + 1}{\nu} = \sum_{(k,l) \in \mathbb{N}} \frac{l + 1}{\nu} \pi_{k,l}, \\ J_3(\pi, U) &= \mathbb{E}\{\mu_\xi \mathbb{I}\{\xi > 0\}\} = \sum_{(k,l) \in \mathbb{N}: k > 0} \mu_k \pi_{k,l}. \end{aligned}$$

В общем виде критерии можно записать как

$$J_m(\pi, U) = \sum_{(k,l) \in \mathbb{N}} f_{k,l}^m(\mu_k, \rho_l) \pi_{k,l}, \quad (9)$$

где функции $f_{k,l}^m(\cdot)$ определены по правилу

$$f_{k,l}^m(\mu, \rho) = \begin{cases} \mathbb{I}\{k = M\}, & m = 0, \\ (k + 1)/(\mu(1 - \rho)), & m = 1, \\ (l + 1)/\nu, & m = 2, \\ \mu \mathbb{I}\{k > 0\}, & m = 3. \end{cases} \quad (10)$$

Итак, задача управления передачей данных в рассматриваемой инфокоммуникационной сети состоит в нахождении допустимого управления $\hat{U} \in \mathcal{U}$ и распределения $\hat{\pi}$, которые бы доставляли минимум

$$\begin{cases} J_0(\pi, U) \rightarrow \min_{U \in \mathcal{U}, \pi \geq 0} \text{ при ограничениях:} \\ J_m(\pi, U) \leq \bar{J}_m, \quad m = 1, 2, 3, \\ \sum_{(k,l) \in \mathbb{N}} a_{(k,l),(i,j)}(\mu_k, \rho_l) \pi_{k,l} = 0 \quad \forall (i, j) \in \mathbb{N}, \\ \sum_{(k,l) \in \mathbb{N}} \pi_{k,l} = 1, \end{cases} \quad (11)$$

где $\bar{J}_m$ — заданные верхние границы. Искомый набор $\hat{U} = \{(\hat{\mu}_k, \hat{\rho}_l)\}_{(k,l) \in \mathbb{N}}$ будем называть *оптимальным управлением*.

Если вероятности отказа заданы изначально, т.е. $\rho = \rho^o$, где $\rho^o = \{\rho_l^o\}_{l \in \mathbb{N}_2}$, то задача оптимального управления сетью передачей данных сводится к нахождению такого управления

$\hat{U} = \{(\hat{\mu}_k, \rho_l^o)\}_{(k,l) \in \mathbb{N}}$, в котором оптимизации подлежит только интенсивность пересылки $\{\mu_k\}$. Эту задачу сформулируем в виде

$$\begin{cases} J_0(\pi, U) \rightarrow \min_{U \in \mathcal{U}_\mu \times \{\rho^o\}, \pi \geq 0} & \text{при ограничениях:} \\ J_m(\pi, U) \leq \bar{J}_m, \quad m = 1, 2, 3, \\ \sum_{(k,l) \in \mathbb{N}} a_{(k,l),(i,j)}(\mu_k, \rho_l^o) \pi_{k,l} = 0 \quad \forall (i,j) \in \mathbb{N}, \\ \sum_{(k,l) \in \mathbb{N}} \pi_{k,l} = 1. \end{cases} \quad (12)$$

4. АНАЛИЗ ОПТИМИЗАЦИОННОЙ ПОСТАНОВКИ

Решение задач (11), (12) может быть найдено аналитически в особом случае, когда $M = N = \infty$, $\mu \equiv \text{const}$, $\rho \equiv \text{const}$, а интенсивность входного потока α меньше $\mu(1 - \rho)$ и ν . В этом случае тандем является сетью Джексона, узлы которой в стационарном режиме можно рассматривать как независимые СМО $M_\alpha | M_{\mu(1-\rho)} | 1$ и $M_\alpha | M_\nu | 1$ соответственно, благодаря чему система уравнений баланса (7) будет иметь аналитическое решение.

Наличие обратной связи, т.е. зависимости параметров μ_k, ρ_l от состояния сети (k, l) , не позволяет представить стационарное распределение π в явном виде. Поэтому для управляемой сети оптимизационные постановки (11) и (12) не допускают аналитического решения. Тем не менее, для любого управления, определяемого набором параметров $\mu_k > 0$ и $\rho_l < 1$, стационарное распределение π существует и единственно. Это связано с тем, что значения марковского процесса (ξ, η) образуют конечный класс сообщающихся состояний. Поэтому система (7) с учетом условия нормировки задает стационарное распределение π , как функцию допустимого управления $U \in \mathcal{U}$.

4.1. Выпуклость множества достижимости в расширенной постановке

Рассмотрим задачу, аналогичную оптимизационной постановке (11), для произвольного управляемого марковского процесса $\zeta(t)$ с конечным множеством состояний $\mathbb{N}$.

Для анализа расширенной постановки понадобятся дополнительные обозначения:

$\mathbf{m}, \mathbf{n}$ — элементы множества $\mathbb{N}$;

$\mathbb{R}^{\mathbb{N}}$ — линейное пространство, состоящее из наборов $x = \{x_n\}_{n \in \mathbb{N}}$, где $x_n \in \mathbb{R}^{\mathbb{N}}$;

$\mathcal{U}$ — множество управляющих параметров;

$a_{\mathbf{m}, \mathbf{n}}(u)$ — интенсивность перехода $\mathbf{m} \rightarrow \mathbf{n}$ при постоянном управлении $u \in \mathcal{U}$, где $m \neq n$;

U — допустимое управление, задаваемое набором $\{u_n\}_{n \in \mathbb{N}}$, в котором u_n — произвольный элемент множества $\mathcal{U}$ (класс допустимых управлений обозначается $\mathcal{U}$);

$A(U)$ — генератор марковского процесса, управляемого стратегией U ; генератор определяется как линейный оператор, действующий в пространстве $\mathbb{R}^{\mathbb{N}}$ по правилу

$$(A(U)x)_{\mathbf{m}} = \lim_{t \downarrow 0} \frac{1}{t} (\mathbb{E}\{x_{\zeta(t)} \mid \zeta(0) = \mathbf{m}\} - x_{\mathbf{m}}) = \sum_{\mathbf{n} \in \mathbb{N}} a_{\mathbf{m}, \mathbf{n}}(u_{\mathbf{m}}) x_{\mathbf{n}}, \quad x \in \mathbb{R}^{\mathbb{N}},$$

где коэффициенты $a_{\mathbf{m}, \mathbf{n}}(u)$ при $\mathbf{m} = \mathbf{n}$ задаются через интенсивности переходов аналогично (8);

$A^*(U)$ — оператор, сопряженный к генератору, т.е.

$$(A^*(U)x)_{\mathbf{m}} = \sum_{\mathbf{n} \in \mathbb{N}} a_{\mathbf{n}, \mathbf{m}}(u_{\mathbf{n}}) x_{\mathbf{n}}, \quad x \in \mathbb{R}^{\mathbb{N}};$$

$\pi(U)$ — стационарное распределение, однозначно определяемое управлением $U \in \mathcal{U}$ из системы уравнений баланса и условия нормировки

$$\pi(U) = \pi: \quad \pi \in \mathbb{R}^N, \quad A^*(U)\pi = 0, \quad \sum_{n \in N} \pi_n = 1.$$

Расширенная постановка задачи оптимизации управляемого марковского процесса состоит в том, чтобы найти управление $\hat{U} \in \mathcal{U}$, доставляющее минимум

$$J_0(U) \rightarrow \min_{U \in \mathcal{U}} \quad \text{при ограничениях} \quad J_m(U) \leq \bar{J}_m, \quad m = \overline{1, p}, \quad (13)$$

где $\bar{J}_m$ — известные верхние границы, а функционалы $J_m(\cdot)$ допускают представление, аналогичное (9):

$$J_m(U) = \sum_{n \in N} f_n^m(u_n) \pi_n(U), \quad m = \overline{0, p}. \quad (14)$$

Определим векторный критерий

$$J(U) = \text{col}[J_0(U), J_1(U), \dots, J_p(U)]$$

и множество достижимости

$$J = \{y \in \mathbb{R}^{p+1}: \exists U \in \mathcal{U}: y \geq J(U)\}. \quad (15)$$

Элементы пространства, в котором принимает значения векторный критерий, будем записывать в виде вектор-столбцов $y = \text{col}[y_0, y_1, \dots, y_p]$.

Согласно определению множество достижимости J образовано значениями всех функционалов $J_m(U)$ при условии, что U пробегает класс допустимых управлений $\mathcal{U}$. Множество J называют также оболочкой Эджуорта—Парето, поскольку точки, эффективные по Парето в многокритериальной оптимизационной задаче $J(U) \rightarrow \min_{U \in \mathcal{U}}$, остаются таковыми после перехода к расширенной задаче $y \rightarrow \min_{y \in J}$.

Следующая теорема указывает условия, достаточные для выпуклости множества достижимости.

Теорема 1. *Допустим, что состояния марковского процесса с интенсивностями переходов $\{\inf_{u \in U} a_{m,n}(u)\}_{m \neq n}$ являются сообщающимися.*

Множество достижимости (15) будет выпуклым, если выпукло каждое из множеств

$$\{(x, y) \in \mathbb{R}^N \times \mathbb{R}^{p+1}: \exists u \in U: x_m = a_{n,m}(u) \quad \forall m \in N, y_m \geq f_n^m(u) \quad \forall m = \overline{0, p}\}, \quad n \in N. \quad (16)$$

Доказательство. Первое условие теоремы обеспечивает тот факт, что каждому управлению $U \in \mathcal{U}$ соответствует единственное стационарное распределение $\pi(U)$.

Для доказательства теоремы достаточно проверить, что для любых $U', U'' \in \mathcal{U}$ и $t \in (0, 1)$ определено допустимое управление U , такое что

$$J_m(U) \leq (1-t)J_m(U') + tJ_m(U''), \quad m = \overline{0, p}, \quad (17)$$

$$A^*(U)\pi(U) = (1-t)A^*(U')\pi(U') + tA^*(U'')\pi(U''). \quad (18)$$

Обозначим для краткости $\pi' = \pi(U')$, $\pi'' = \pi(U'')$ и определим новое распределение

$$\pi = (1 - t)\pi' + t\pi''. \quad (19)$$

Для выполнения соотношений (17), (18) достаточно так подобрать управляющие параметры $u_n \in U$, чтобы были выполнены условия

$$f_n^m(u_n)\pi_n \leq (1 - t)\pi'_n f_n^m(u'_n) + t\pi''_n f_n^m(u''_n), \quad m = \overline{0, p}, \quad (20)$$

$$a_{n,m}(u_n)\pi_n = (1 - t)a_{n,m}(u'_n)\pi'_n + ta_{n,m}(u''_n)\pi''_n, \quad n, m \in N, \quad (21)$$

$$\pi = \pi(U), \quad \text{где } U = \{u_n\}_{n \in N}. \quad (22)$$

Из указанных условий (22) можно исключить, поскольку оно автоматически следует из второго равенства. Действительно, после суммирования по n равенство (21) превращается в уравнение баланса $A^*(U)\pi = 0$.

Систему условий (20), (21) можно представить в эквивалентном виде

$$f_n^m(u_n) \leq t'_n f_n^m(u'_n) + t''_n f_n^m(u''_n), \quad m = \overline{0, p}, \quad (23)$$

$$a_{n,m}(u_n) = t'_n a_{n,m}(u'_n) + t''_n a_{n,m}(u''_n), \quad m \in N, \quad (24)$$

если обозначить $t'_n = (1 - t)\pi'_n/\pi_n$, $t''_n = t\pi''_n/\pi_n$ и считать, что n — произвольное состояние, такое что $\pi_n > 0$. Случай $\pi_n = 0$ исключается из рассмотрения, так как он влечет $\pi'_n = \pi''_n = 0$, что превращает соотношения (20), (21) в тождества.

В силу $t'_n + t''_n = 1$ существование параметра u_n , удовлетворяющего условиям (23) и (24), вытекает из выпуклости множества (16), соответствующего состоянию n .

Таким образом, теорема 1 доказана.

Установленный результат аналогичен теореме 6 из [9] и теореме 2 из [10], которые сформулированы для нестационарной модели управляемого марковского процесса.

Заметим, что в силу (19) из приведенного доказательства следует выпуклость семейства распределений, соответствующих допустимым управлениям.

Следствие 1. В условиях теоремы 1 стационарные распределения $\pi(U)$, $U \in \mathcal{U}$, образуют выпуклое множество.

Благодаря теореме 1 постановку (13) можно трактовать как задачу выпуклого программирования, если перейти к оптимизации в пространстве критериев. Итак, справедливо

Следствие 2. Пусть выполнены условия теоремы 1. Тогда значение векторного критерия $J(\hat{U})$ на управлении $\hat{U}$, оптимальном в смысле (13), является решением задачи выпуклого программирования

$$y_0 \rightarrow \min_{y \in J} \quad \text{при } y_m \leq \bar{J}_m, \quad m = \overline{1, p}. \quad (25)$$

Обратно, для всякого решения $\hat{y}$ задачи (25) найдется оптимальное управление $\hat{U}$ из (13), такое что $J(\hat{U}) = \hat{y}$.

4.2. Выпуклость множества достижимости в двухагентной модели

Вернемся к исходной модели, в которой управляемый марковский процесс (ξ, η) определяет состояние сети передачи данных между двумя агентами: беспилотником и базовой станцией. Качество работы указанной сети описывается четырьмя функционалами (3)–(6), для которых по аналогии с (14) всюду далее будем использовать обозначения $J_m(U)$, $m = \overline{0, 3}$. Рассмотрим вопрос о выпуклости множества достижимости, задаваемого соответствующим векторным критерием $J(U) = \text{col}[J_0(U), J_1(U), J_2(U), J_3(U)]$.

К сожалению, непосредственное применение теоремы 1 к исходной задаче (11) невозможно, поскольку она не является частным случаем расширенной постановки (13). Действительно, в силу (2) допустимое управление $U \in \mathcal{U}$ определяется парой стратегий $\{\mu_k\}$ и $\{\rho_l\}$, каждая из которых зависит только от состояния своей системы, но не от состояния (k, l) сети в целом.

Тем не менее, изложенный выше метод доказательства теоремы 1 позволяет получить следующий результат.

Теорема 2. *Множество достижимости*

$$J = \{y \in \mathbb{R}^4 : \exists U \in \tilde{\mathcal{U}} : y \geq J(U)\}$$

выпукло, если класс допустимых управлений $\tilde{\mathcal{U}}$ имеет вид

$$\tilde{\mathcal{U}} = \{(\mu_{k,l}, \rho_l) : \mu_{k,l} \in (0, \mu_{\max}], \rho_l \in [\rho_{\min}, \rho_{\max}], (k, l) \in \mathbb{N}\},$$

где $0 \leq \rho_{\min} \leq \rho_{\max} < 1$.

Доказательство. Рассуждая, как при доказательстве теоремы 1, приходим к системе условий

$$\begin{cases} f_{k,l}^m(\mu_{k,l}, \rho_l) \leq t'_{k,l} f_{k,l}^m(\mu'_{k,l}, \rho'_l) + t''_{k,l} f_{k,l}^m(\mu''_{k,l}, \rho''_l), \\ a_{(k,l),(i,j)}(\mu_{k,l}, \rho_l) = t'_{k,l} a_{(k,l),(i,j)}(\mu'_{k,l}, \rho'_l) + t''_{k,l} a_{(k,l),(i,j)}(\mu''_{k,l}, \rho''_l), \end{cases} \quad (26)$$

которые необходимо удовлетворить при любых $m = \overline{0, 3}$ и $(k, l), (i, j) \in \mathbb{N}$ за счет выбора управления $U = \{(\mu_{k,l}, \rho_l)\} \in \tilde{\mathcal{U}}$ по заданным

$$U' = \{(\mu'_{k,l}, \rho'_l)\} \in \tilde{\mathcal{U}}, \quad U'' = \{(\mu''_{k,l}, \rho''_l)\} \in \tilde{\mathcal{U}}, \quad t'_{k,l}, t''_{k,l} \geq 0, \quad t'_{k,l} + t''_{k,l} = 1.$$

В силу (1) и (10) условия (26) можно представить в следующем виде:

$$\frac{1}{\mu_{k,l}(1 - \rho_l)} \leq t'_{k,l} \frac{1}{\mu'_{k,l}(1 - \rho'_l)} + t''_{k,l} \frac{1}{\mu''_{k,l}(1 - \rho''_l)}, \quad (27)$$

$$\mu_{k,l} \leq t'_{k,l} \mu'_{k,l} + t''_{k,l} \mu''_{k,l}, \quad (28)$$

$$\mu_{k,l}(1 - \rho_l) = t'_{k,l} \mu'_{k,l}(1 - \rho'_l) + t''_{k,l} \mu''_{k,l}(1 - \rho''_l), \quad \text{где } (k, l) \in \mathbb{N}. \quad (29)$$

Выберем искомые стратегии по правилу

$$\rho_l = \min\{\rho'_l, \rho''_l\}, \quad \mu_{k,l} = t'_{k,l} \mu'_{k,l} \gamma'_l + t''_{k,l} \mu''_{k,l} \gamma''_l, \quad \gamma'_l = \frac{1 - \rho'_l}{1 - \rho_l}, \quad \gamma''_l = \frac{1 - \rho''_l}{1 - \rho_l}.$$

Тогда равенство (29) выполнено автоматически, неравенство (28) следует из того, что коэффициенты γ'_l, γ''_l не больше единицы, а соотношение (27) вытекает из (29) и выпуклости $1/x$, как функции положительного аргумента x . При этом условие $0 < \mu_{k,l} \leq \mu_{\max}$ имеет место в силу $0 < \gamma'_l, \gamma''_l \leq 1$, а включение $\rho_l \in [\rho_{\min}, \rho_{\max}]$ справедливо по построению.

Итак, теорема 2 доказана.

Тем самым выпуклость множества достижимости установлена для случая, когда интенсивность обслуживания μ зависит от состояния обоих агентов, а вероятность отказа ρ является функцией состояния второго агента.

Для исходных постановок (11), (12) представленный выше метод доказательства уже не реализуем. Причина этого состоит в том, что для обеспечения $(M+1)(N+1)$ равенств (29) не хватает набора переменных $\{\mu_k, \rho_l, (k, l) \in \mathbb{N}\}$, число которых равно лишь $M+N+2$.

Поэтому выпуклость множества достижимости в постановках (11), (12), а также их эквивалентность задачам выпуклого программирования остается открытым вопросом.

5. ЧИСЛЕННЫЕ СХЕМЫ

Цель данного раздела описать численный метод решения исходной задачи оптимального управления (11). Для этого запишем ее с учетом упрощенных обозначений для функционалов (14):

$$J_0(U) \rightarrow \min_{U \in \mathcal{U}} \quad \text{при} \quad J_m(U) \leq \bar{J}_m, \quad m = 1, 2, 3. \quad (30)$$

Из представления (9) функционалов $J_m(U)$, $m = \overline{0, 3}$, получаем соотношения для их частных производных по стратегиям μ_k и ρ_l :

$$\frac{\partial J_m(U)}{\partial \mu_k} = \sum_{j \in \mathbb{N}_2} \frac{\partial f_{k,j}^m(\mu_k, \rho_j)}{\partial \mu_k} \pi_{k,j} + \sum_{(i,j) \in \mathbb{N}} f_{i,j}^m(\mu_i, \rho_j) \frac{\partial \pi_{i,j}}{\partial \mu_k}, \quad k \in \mathbb{N}_1, \quad (31)$$

$$\frac{\partial J_m(U)}{\partial \rho_l} = \sum_{i \in \mathbb{N}_1} \frac{\partial f_{i,l}^m(\mu_i, \rho_l)}{\partial \rho_l} \pi_{i,l} + \sum_{(i,j) \in \mathbb{N}} f_{i,j}^m(\mu_i, \rho_j) \frac{\partial \pi_{i,j}}{\partial \rho_l}, \quad l \in \mathbb{N}_2, \quad (32)$$

где π обозначает стационарное распределение, соответствующее управлению $U = \{(\mu_k, \rho_l)\}$, а производные функций $f_{i,j}^m$, $(i, j) \in \mathbb{N}$, имеют следующий вид:

$$\frac{\partial f_{i,j}^m(\mu, \rho)}{\partial \mu} = \begin{cases} 0, & m = 0, 2, \\ -\frac{i+1}{\mu^2(1-\rho)}, & m = 1, \\ I\{i > 0\}, & m = 3, \end{cases} \quad \frac{\partial f_{i,j}^m(\mu, \rho)}{\partial \rho} = \begin{cases} 0, & m = 0, 2, 3, \\ \frac{i+1}{\mu(1-\rho)^2}, & m = 1. \end{cases}$$

Продифференцируем уравнения баланса (7) и условие нормировки по управляющим параметрам. В результате этого получим две независимые системы уравнений для определения частных производных стационарных вероятностей по интенсивности обслуживания μ_p

$$\sum_{l \in \mathbb{N}_2} \frac{\partial a_{(p,l),(i,j)}(\mu_p, \rho_l)}{\partial \mu_p} \pi_{p,l} + \sum_{(k,l) \in \mathbb{N}} a_{(k,l),(i,j)}(\mu_k, \rho_l) \frac{\partial \pi_{k,l}}{\partial \mu_p} = 0 \quad \forall (i, j) \in \mathbb{N}, \quad \sum_{(k,l) \in \mathbb{N}} \frac{\partial \pi_{k,l}}{\partial \mu_p} = 0, \quad (33)$$

и по вероятности отказа ρ_q

$$\sum_{k \in \mathbb{N}_1} \frac{\partial a_{(k,q),(i,j)}(\mu_k, \rho_q)}{\partial \rho_q} \pi_{k,q} + \sum_{(k,l) \in \mathbb{N}} a_{(k,l),(i,j)}(\mu_k, \rho_l) \frac{\partial \pi_{k,l}}{\partial \rho_q} = 0 \quad \forall (i, j) \in \mathbb{N}, \quad \sum_{(k,l) \in \mathbb{N}} \frac{\partial \pi_{k,l}}{\partial \rho_q} = 0. \quad (34)$$

Для нахождения указанных частных производных сначала необходимо найти стационарное распределение из (7), а затем для каждого $p \in \mathbb{N}_1$ и $q \in \mathbb{N}_2$ отдельно решить (33) и (34),

как системы линейных неоднородных уравнений относительно набора неизвестных $\frac{\partial \pi}{\partial \mu_p}$ и $\frac{\partial \pi}{\partial \rho_q}$ соответственно.

Для численного решения задачи (30) воспользуемся методом модифицированной функции Лагранжа [11]. Введем модифицированную функцию Лагранжа

$$M_\varepsilon(U, \lambda) = J_0(U) + \frac{\varepsilon}{2} \sum_{m=1}^3 ((\lambda_m + (J_m(U) - \bar{J}_m)/\varepsilon)_+^2 - \lambda_m^2),$$

где $\lambda = \text{col}[\lambda_1, \lambda_2, \lambda_3]$ — вектор множителей Лагранжа, ε — параметр регуляризации, $(\cdot)_+$ — положительная часть числа.

Норма градиента $M_\varepsilon(U, \lambda)$ по $U \in \mathcal{U}$ определяет удаленность текущего решения от точки оптимума. Так как множество $\mathcal{U}$ представляет собой координатный параллелепипед, в формулировке критерия останова градиентного метода необходимо учесть знак частной производной при попадании на границу. Например, при минимизации выпуклой функции $f(v)$ на отрезке $[a, b]$ критерием останова будет $|S_{a,b}(g, v)| < \delta$, где δ — заданный порог, $S_{a,b}(g, v)$ — «подправленное» значение производной $g = f'(v)$:

$$S_{a,b}(g, v) = \begin{cases} g, & v = a, g < 0 \quad \text{или} \quad v = b, g > 0 \quad \text{или} \quad v \in (a, b); \\ 0, & v = a, g \geq 0 \quad \text{или} \quad v = b, g \leq 0. \end{cases}$$

С учетом введенных обозначений алгоритм численного решения задачи (30) примет следующий вид.

Алгоритм 1. Задать величину шага γ , параметр регуляризации ε , уровни погрешности δ_μ и δ_ρ , начальные приближения $U^{(0)} = \{(\mu_k^{(0)}, \rho_l^{(0)}) : (k, l) \in \mathbf{N}\}$, $\mu_k^{(0)} = (\mu_{\min} + \mu_{\max})/2$, $\rho_l^{(0)} = (\rho_{\min} + \rho_{\max})/2$ и $\lambda_m^{(0)} = 1$, $m = 1, 2, 3$, а также положить $s = 0$.

1. Найти стационарное распределение $\pi = \pi(U^{(s)})$ из (7), подставить его в системы уравнений (33), (34) и решить относительно $\frac{\partial \pi}{\partial \mu_p}$ и $\frac{\partial \pi}{\partial \rho_q}$ для $(p, q) \in \mathbf{N}$.
2. Найти $\frac{\partial J_m(U^{(s)})}{\partial \mu_k}$, $\frac{\partial J_m(U^{(s)})}{\partial \rho_l}$ по формулам (31)–(32).
3. Вычислить частные производные модифицированной функции Лагранжа:

$$\frac{\partial M_\varepsilon(U^{(s)}, \lambda^{(s)})}{\partial \mu_k} = \frac{\partial J_0(U^{(s)})}{\partial \mu_k} + \sum_{m=1}^3 \left(\lambda_m^{(s)} + \frac{1}{\varepsilon} (J_m(U^{(s)}) - \bar{J}_m) \right)_+ \frac{\partial J_m(U^{(s)})}{\partial \mu_k}, \quad k \in \mathbf{N}_1,$$

$$\frac{\partial M_\varepsilon(U^{(s)}, \lambda^{(s)})}{\partial \rho_l} = \frac{\partial J_0(U^{(s)})}{\partial \rho_l} + \sum_{m=1}^3 \left(\lambda_m^{(s)} + \frac{1}{\varepsilon} (J_m(U^{(s)}) - \bar{J}_m) \right)_+ \frac{\partial J_m(U^{(s)})}{\partial \rho_l}, \quad l \in \mathbf{N}_2,$$

$$\frac{\partial M_\varepsilon(U^{(s)}, \lambda^{(s)})}{\partial \lambda_m} = \begin{cases} J_m(U^{(s)}), & \text{если } J_m(U^{(s)}) \geq \bar{J}_m - \varepsilon \lambda_m^{(s)}, \\ -\varepsilon \lambda_m^{(s)}, & \text{иначе,} \end{cases} \quad m = 1, 2, 3.$$

4. Закончить итерации, если выполнен следующий критерий останова:

$$\left| S_{\mu_{\min}, \mu_{\max}} \left(\frac{\partial M_\varepsilon(U^{(s)}, \lambda^{(s)})}{\partial \mu_k}, \mu_k \right) \right| \leq \delta_\mu, \quad \left| S_{\rho_{\min}, \rho_{\max}} \left(\frac{\partial M_\varepsilon(U^{(s)}, \lambda^{(s)})}{\partial \rho_l}, \rho_l \right) \right| \leq \delta_\rho \quad \forall (k, l) \in \mathbf{N}.$$

5. Положить

$$U^{(s+1)} = P_{\mathcal{U}}[U^{(s)} - \gamma \nabla_U M_\varepsilon(U^{(s)}, \lambda^{(s)})], \quad \lambda^{(s+1)} = (\lambda^{(s)} + \gamma \nabla_\lambda M_\varepsilon(U^{(s)}, \lambda^{(s)}))_+,$$

где $P_{\mathcal{U}}$ — оператор проектирования на множество $\mathcal{U}$, увеличить s на единицу и перейти к шагу 1.

6. ЧИСЛЕННЫЙ ЭКСПЕРИМЕНТ

Теперь перейдем к описанию результатов численного эксперимента, который был проведен с целью оптимизации стратегий управления в рамках постановок (11) и (12).

Параметры сети передачи данных между БПЛА и АБС были выбраны следующим образом:

$$\begin{cases} \alpha = 2, & \nu = 2,5, & M = 10, & N = 15, \\ \mu_{min} = 0,5, & \mu_{max} = 4, & \rho_{min} = 0, & \rho_{max} = 0,95, \\ \bar{J}_1 = 2,50, & \bar{J}_2 = 1,55, & \bar{J}_3 = 2,15. \end{cases} \quad (35)$$

Общее количество состояний в данной модели составило $(M + 1)(N + 1) = 176$. Для того чтобы стационарное распределение не было локализовано в состояниях, соответствующих перегрузке базовой станции, интенсивность обработки ν была взята больше интенсивности входного потока α .

В таблице 1 представлены значения функционалов на управлении $\bar{U}$, определяемом постоянными стратегиями $\mu_k \equiv \bar{\mu} = 2,4$ и $\rho_l \equiv \bar{\rho} = 0,1$. Каждая из верхних границ $\bar{J}_m$ была выбрана так, чтобы она оказалась примерно на 3 % больше значения соответствующего функционала на управлении $\bar{U}$.

Таблица 1. Верхние границы для функционалов

m	$J_m(\bar{U})$	$\bar{J}_m/J_m(\bar{U})$	$\bar{J}_m$
0	0,0608	—	—
1	2,4369	1,0259	2,50
2	1,5025	1,0316	1,55
3	2,0922	1,0276	2,15

Решение задачи (12) было получено в результате применения алгоритма 1 к модели (35) при условии, что вероятность отказа определялась следующей фиксированной стратегией:

$$\rho_l^o = \begin{cases} \rho_{min}, & l \leq N/3, \\ \rho_{min} + (\rho_{max} - \rho_{min}) \frac{l - N/3}{N/3}, & N/3 < l < 2N/3, \\ \rho_{max}, & l \geq 2N/3. \end{cases}$$

На рис. 2 представлена оптимальная интенсивность обслуживания $\{\hat{\mu}_k\}$ вместе с заданной выше вероятностью отказа $\{\rho_l^o\}$. Как можно видеть, $\hat{\mu}_k > \nu$ для любого k , т.е. при оптимальном управлении беспилотник обрабатывает пакеты быстрее базовой станции. Неравенство $\hat{\mu}_0 < \hat{\mu}_k$, $k = \bar{1}, \bar{M}$, означает, что при простое системы интенсивность обработки назначается минимально возможной. При увеличении μ функционалы J_0, J_1 убывают, а J_2, J_3 , напротив, возрастают. Поэтому изменение направления монотонности μ_k при $k = 4$ означает, что начиная с этого числа пакетов более весомыми оказываются первые два критерия, учитывающие частоту потерь входных данных и время их пересылки на базовую станцию. Однако, как следует из таблицы 2, ограничение на J_1 было выполнено с существенным запасом, поэтому рост $\hat{\mu}_k$ при $k = \bar{4}, \bar{M}$ вплоть до μ_{max} обусловлен требованием минимизации целевого функционала J_0 .

Стоит отметить, что в сравнении с неуправляемой сетью, имеющей параметры $\bar{\mu}$ и $\bar{\rho}$, оптимизация интенсивности обслуживания $\{\hat{\mu}_k\}$ позволила сократить долю потерь входного трафика более, чем в четыре раза.

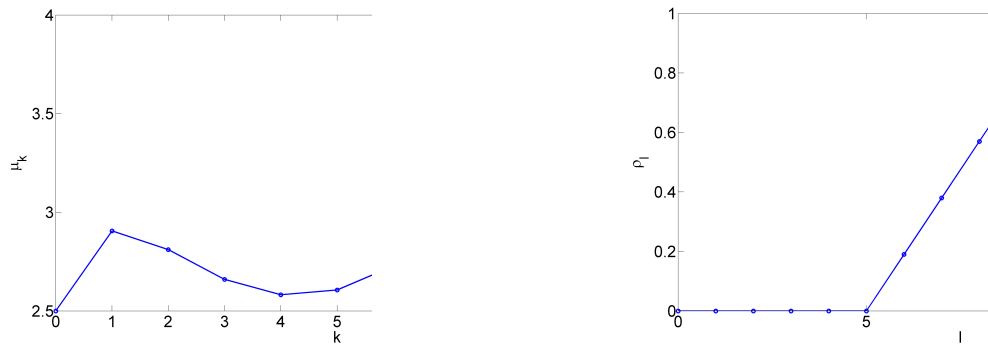


Рис. 2. Оптимальная интенсивность обслуживания $\hat{\mu}_k$ (слева) при заданной вероятности отказа ρ_l^o (справа).

Теперь рассмотрим решение задачи (11), в которой оптимизации подлежит как интенсивность обслуживания $\{\mu_k\}$, так и вероятность отказа $\{\rho_l\}$.

В результате применения алгоритма 1 были получены стратегии $\{\hat{\mu}_k\}$ и $\{\hat{\rho}_l\}$, изображенные на рис. 3. В сравнении с задачей (12) вид оптимальной интенсивности $\{\hat{\mu}_k\}$ в рассматриваемом случае не изменился. Оптимальная вероятность отказа $\{\hat{\rho}_l\}$ также сохранила структуру зависимости стратегии $\{\rho_l^o\}$. Поэтому оптимизация по $\{\rho_l\}$ привела лишь к незначительному уменьшению вероятности потерь J_0 (см. таблицу 2).

Присутствие колебаний в найденной стратегии $\{\hat{\rho}_l\}$ требует проанализировать алгоритм на предмет устойчивости его работы. Для этого были построены несколько субоптимальных стратегий, являющихся по отношению к найденной различными вариантами сглаженной аппроксимации. На рис. 3 изображена одна из них $\{\tilde{\rho}_l\}$, вычисленная с помощью операции скользящего среднего по трем значениям. Как видно из таблицы 2, данная стратегия не привела к уменьшению целевого функционала J_0 , но позволила уменьшить значение J_1 . Однако указанное улучшение стало возможным за счет нарушения ограничений, т.е. для управления $\tilde{U} = \{(\hat{\mu}_k, \tilde{\rho}_l)\}$ оказалось, что $J_m(\tilde{U}) > \bar{J}_m$ для $m = 2, 3$. Другие варианты субоптимальных стратегий также не уменьшают долю потерянных пакетов и приводят к тому, что часть ограничений перестают выполняться. Таким образом, проведенный анализ подтверждает оптимальность вероятности отказа, найденной с помощью алгоритма 1.

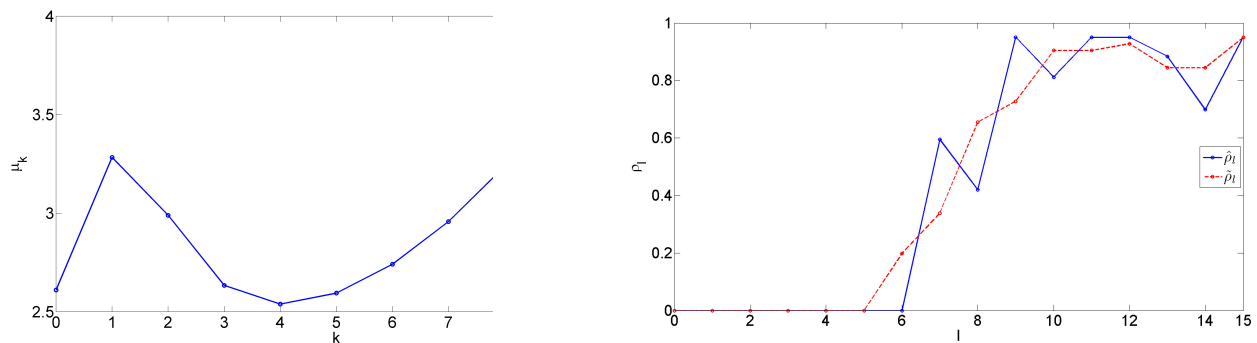


Рис. 3. Оптимальная интенсивность обслуживания $\hat{\mu}_k$ (слева) и вероятность отказа (справа): оптимальная $\hat{\rho}_l$ (сплошная линия) и субоптимальная $\tilde{\rho}_l$ (штриховая линия).

На рис. 4 приведены оптимальные стратегии для сети, в которой размер буфера второй системы N уменьшен с 15 до 10. Структура зависимости найденных стратегий остается в целом такой же. Тем не менее, сокращение мест для ожидания во второй системе привело к тому, что оптимальная вероятность отказа $\{\hat{\rho}_l\}$ уже не содержит колебаний, ее переход от ρ_{min} к ρ_{max}

происходит более резко, а пересылка пакетов теперь занимает большее время по сравнению (см. таблицу 2).

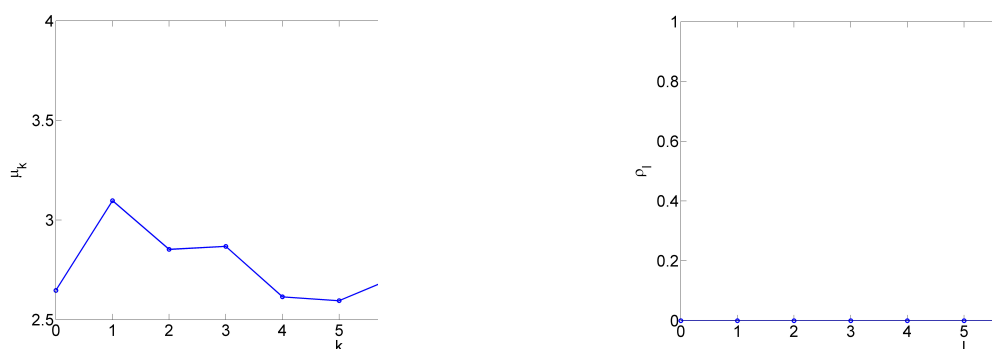


Рис. 4. Оптимальные стратегии при уменьшенном размере буфера второй системы: интенсивность обслуживания $\hat{\mu}_k$ (слева) и вероятность отказа $\hat{\rho}_l$ (справа).

На основе полученных результатов можно сделать вывод о том, что разработанный метод оптимизации параметров инфокоммуникационной сети позволяет определить наилучший закон управления передачей данных в рамках имеющихся ограничений.

Таблица 2. Значения функционалов на нескольких управлениях вместе с верхними границами

m	$J_m(\hat{U})$ при заданной ρ^o	$J_m(\hat{U})$	$J_m(\tilde{U})$	$J_m(\hat{U})$ при сокращенном буфере	$\bar{J}_m$
0	0,0143	0,0139	0,0139	0,0139	—
1	1,7710	2,1004	1,9895	2,3691	2,50
2	1,5500	1,5500	1,5518	1,5500	1,55
3	2,1488	2,1500	2,1531	2,1500	2,15

7. ЗАКЛЮЧЕНИЕ

В статье рассмотрена задача оптимизации параметров сети, обеспечивающей процесс обработки и передачи данных в мобильной двухагентной робототехнической системе. Данная задача состоит в минимизации потерь входных пакетов в стационарном режиме с учетом ограничений на время обработки и энергозатраты за счет управления скоростью пересылки между первым и вторым агентами и вероятностью отказа в приеме пакета со стороны второго агента.

В работе получены достаточные условия выпуклости множества достижимости, определяемого значениями векторного критерия в задаче оптимального управления стационарным марковским процессом с конечным множеством состояний. Для управляемой сети передачи данных выпуклость указанного множества доказана в предложении, что интенсивность обработки в первой системе зависит от состояния всей сети, а вероятность отказа определяется только вторым агентом.

Если же оба агента назначают свои стратегии без учета состояния другого агента, то выпуклость множества достижимости в такой постановке остается открытой проблемой. Тем не менее, численный синтез оптимальных решений в этой задаче можно осуществить с помощью метода модифицированной функции Лагранжа. Этот факт продемонстрирован на серии численных экспериментов, по результатам которых определена структура оптимальных стратегий.

СПИСОК ЛИТЕРАТУРЫ

1. Choi D. H., Kim S. H., Sung D. K. Energy-efficient maneuvering and communication of a single UAV-based relay. *IEEE Trans. Aerospace & Electronic Systems*, 2014, vol. 50, no. 3, pp. 2319–2326.
2. Ho D.-T., Grotli E. I., Sujit P. B. *et al.* Optimization of wireless sensor network and UAV data acquisition. *J. Intelligent & Robotic Systems*, 2015, vol. 78, no. 1, pp. 159–179.
3. Cesare K., Skeeel R., Yoo S.-H. *et al.* Multi-UAV exploration with limited communication and battery. *Proc. IEEE International Conf. Robotics and Automation (ICRA'2015)*. Seattle, USA: IEEE, 2015, pp. 2230–2235.
4. Floyd S., Jacobson V. Random early detection gateways for congestion avoidance. *IEEE/ACM Transactions on Networking*, 1993, vol. 1, no. 4, pp. 397–413.
5. Дынкин Е. Б., Юшкевич А. А. *Управляемые марковские процессы и их приложения*. М.: Наука, 1975.
6. Altman E. *Constrained Markov decision processes*. Boca Raton, FL: Chapman & Hall/CRC, 1999.
7. Hordijk A., Spieksma F. Constrained admission control to a queueing system. *Advances in Applied Probability*, 1989, vol. 21, no. 2, pp. 409–431.
8. Miller A., Miller B., Popov A., Stepanyan K. Towards the development of numerical procedure for control of connected Markov chains. *Proc. 5th Australian Control Conf. (AUCC'2015)*. Gold Coast, Australia: IEEE, 2015, pp. 336–341.
9. 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.
10. 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.
11. Поляк Б. Т. *Введение в оптимизацию*. М.: Наука, 1980.

Optimal control of data transmission in mobile multiagent robotic systems

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

In this article, we consider a controlled queueing network providing data transmission in a mobile two-agent robotic system. The agents interact as follows: the first agent receives the Poisson flow of packets and retransmits them to the second agent; the first agent rejects incoming packets if the buffer is full, whereas the second agent notifies the first one on the overflow and requests the packet to be retransmitted. The optimization method is developed to obtain the policies for the retransmission rate and the rejection probability that minimize input packet losses under the constraints on processing time and energy consumption. We define sufficient conditions for the objective function and constraints to form a convex set. The augmented Lagrangian method is then used to numerically synthesize the optimal control policies.

KEYWORDS: optimal control, multiagent system, queueing network, Markov decision process, optimization under constraints.