ISSN 1064-2269, Journal of Communications Technology and Electronics, 2017, Vol. 62, No. 12, pp. 1484–1498. © Pleiades Publishing, Inc., 2017.
Original Russian Text © N.A. Kuznetsov, D.V. Myasnikov, K.V. Semenikhin, 2017, published in Informatsionnye Protsessy, 2017, Vol. 17, No. 1, pp. 19–42.

ANALYSIS AND SYNTHESIS OF CONTROL SYSTEMS

Optimization of Two-Phase Queuing System and Its Application to the Control of Data Transmission between Two Robotic Agents[1]

N. A. Kuznetsov[a, b, *], D. V. Myasnikov[a, **], and K. V. Semenikhin[a, b, c, *]**

[a]*Kotel'nikov Institute of Radio Engineering and Electronics, Russian Academy of Sciences, Moscow, 125009 Russia*
[b]*Moscow Institute of Physics and Technology, Dolgoprudnyi, Moscow oblast, 141700 Russia*
[c]*Moscow Aviation Institute (National Research University), Moscow, 125993 Russia*
**e-mail: kuznetsov@cplire.ru*
***e-mail: dmitry.myasnikov89@gmail.com*
****e-mail: siemenkv@gmail.com*
Received February 15, 2017

Abstract—A tandem queuing system that contains two single-channel stations with finite buffers and allows blocking of the first server is considered. The first station receives nonstationary Poisson packet flow that is processed at a controlled rate. In the case of the queue overflow in the first system, the input packet is lost. The second station does not allow overflow due to control of the acceptance probability (a decrease in such a probability leads to slowing of packet sending from the first station). The queuing system is described with the aid of the controlled Markov process . The optimal control problem is considered over a finite time horizon using the criterion of minimum average losses under the constraints on the total service time and energy consumption of the first station. Optimization algorithms are proposed for synthesis of the control law for nonstationary data flow in two-agent robotic system.

Keywords: optimal control, multiagent system, two-phase queueing system, controlled Markov process, constrained optimization

DOI: 10.1134/S1064226917120087

1. INTRODUCTION

Multiphase (tandem) queuing systems are used for mathematical modelling of the processing in which the arriving units are sequentially served at several stages. Sequential servicing is natural for query processing in help desk centers [1], multimedia data transfer using wireless channels [2], and data transmission control between elements of a multiagent robotic system [3].

Blocking policies are used to prevent overflow in the most important units of the multiphase system. Two-phase systems with blocking have been studied in [4–6] and stationary probabilities of states have been determined under different assumptions regarding the input flow and the distribution of the service time.

The threshold structure for the strategy that is optimal with respect to the minimum of the weighted average load has been determined in the first works on two-phase systems with feedback control [7, 8]. The theory of Markov decision processes (MDPs) has been employed in these works to derive the dynamic programming equation. This approach has been later generalized to more general models of controlled queuing networks [9–11]. Recent works [12, 13] on the optimization of two-phase queuing systems are devoted to the analysis of the optimal access control (explicit expressions for the threshold coefficients of the criterion are presented).

The analysis of published results shows that the controlled two-phase queuing systems are studied only in the steady-state mode and the control quality is determined using a single functional. Thus, the constrained optimization of tandem queues over a finite horizon has not been studied in the theory of controlled queuing networks.

The approach of constrained MDPs optimized in the steady-state mode has been developed in [14]. The optimization methods for solving finite horizon control problems have been proposed in [15, 16] for a class of nonstationary continuous-time Markov jump processes. Such methodology has been employed in [17] for optimization of a single-channel queuing system.

[1] The development of models and algorithms for control of nonstationary data flow in two-agent robotic system (Sections 2, 3, and 7) was supported by the Russian Science Foundation (project no. 16-11-00063). The analysis of the problem of constrained optimization of the two-phase queuing system (Sections 4–6) was supported by the Russian Foundation for Basic Research (projects nos. 15-37-20611-mol_a_ved and 16-07-00677-a).

In this work, we consider the controlled two-phase queuing system that contains two single-channel stations with finite buffers and allows blocking of the first server. Packets in a nonstationaty Poisson stream enter the first station. The packets are processed at a controlled rate at the first server. The input packet is lost if the first queue is full. The second station does not allow the overflow due to control of the acceptance probability (or, blocking probability). A decrease in the acceptance probability leads to delays in sending of packets from the first station. The queuing system is described by a controlled Markov jump process which is optimized on a finite-time interval using the criterion of minimum average losses under the constraints on the total processing time and energy consumption of the first station. We propose two algorithms to determine optimal control within the classes of centralized and decentralized strategies. The computer simulation results show a typical form of the optimal strategy in the data transmission control problem for a two-agent robotic system.

2. INFORMAL DESCRIPTION OF THE MODEL

We consider an open tandem queuing system that consists of two single-channel stations. They play the role of a transmitter and a base station. The transmitter receives the input flow of packets (data units) and sends them (with a random delay time) for the further processing to the base station.

The number of packets in both systems is limited and the maximum numbers are M and N, respectively. If the first queue is full, the arriving packet is lost. Such an event is undesired, and the corresponding probability must be minimized. For this purpose, we can vary two parameters $\mu \geq 0$ and $\vartheta \in [0, 1]$, where μ is the service rate in the first station and $1 - \vartheta$ is the probability that the second station rejects the packet. The rejection is used to avoid overflow of the base station, and, in the case of rejection, the packet is not lost and is stored in the transmitter. Thus, the product of parameters μ and ϑ can be interpreted as the total rate for packet transmission from the transmitter to the base station.

The queuing network must provide a desired level of the total service time, including delays in transmission, waiting in queues, and processing at the base station. The server of the base station operates with a fixed processing rate ν.

Service rate μ must be increased to minimize the number of lost packets and decrease the total service time. However, the operability of the transmitter is related to strict constraints on energy consumption, so that an infinite increase in service rate μ becomes impossible.

Parameter ϑ must also correspond to a certain intermediate state between two extreme values corresponding to blocking of all the traffic from the transmitter ($\vartheta = 0$) and the acceptance of all packets ($\vartheta = 1$). In the latter case, the buffer of the base station is rapidly filled with big amount of packets which leads to overflow of the transmiter and, hence, significant losses.

We assume that both parameters (μ and ϑ) depend on time and state to take into account the nonstationary character of the input flow and be able to respond to changes of network functioning. Thus, the desired control is built in the feedback form. However, two basically different formulations must be considered. In the first formulation, service rate μ of the transmitter and acceptance probability of the base station ϑ are controlled using complete information on the current state of the network. Such control is called centralized, since it involves matched functioning of the transmitter and base station. In the second formulation, service rate μ is determined only by the state of the transmitter while probability ϑ depends only on the state of the base station. Thus, the decision making in each system is performed in the presence of incomplete information on the state of the other system. Thus, the corresponding control is decentralized.

The scenario with incomplete information is more realistic for the transmission of video-data flow or telemetry when the data are sent to a stationary monitor station from an on-board transmitter of an unmanned aerial vehicle (UAV). The UAV's queuing system transforms the input data flow into a series of standard packets in accordance with the data-exchange protocol of the base station. Thus, service rate μ is determined by the time that is spent by the formation and sending of the packet and probability ϑ is determined by the arrival rate of messages from the base station that confirms the successful delivery. In the presence of rejections ($\vartheta < 1$), average time of packet sending $1/\mu$ increases by $(1/\vartheta - 1) \times 100\%$.

Therefore, the infotelecommunication system described above needs to be optimized with specific requirements on its structure, nonstationary behavior of input data flow, and limitations on energy consumption.

3. FORMULATION OF THE PROBLEM

Below, we formally analyze the queuing system.

Let $X(t)$ and $Y(t)$ be the numbers of packets at the transmitter and base station, respectively, at moment t. Then, random process $Z(t) = (X(t), Y(t))$ that describes the current state of the network takes values from the set

$$\mathcal{Z} = \mathcal{X} \times \mathcal{Y}, \text{ where}$$
$$\mathcal{X} = \{0, 1, \ldots, M\}, \ \mathcal{Y} = \{0, 1, \ldots, N\}.$$

We assume that the packets that arrive to the transmitter form a nonstationary Poisson flow with known continuous intensity $\alpha(t)$. First, we consider a constant control u:

$$u = (m, v) \in \mathrm{U}, \quad \mathrm{U} = |\underline{m}, \overline{m}| \times |\underline{v}, \overline{v}|,$$

where $0 \le \underline{m} \le \overline{m} < \infty$ and $0 \le \underline{v} \le \overline{v} \le 1$ are the given ranges of the service rate at the transmitter and acceptance probability of the base station. In this case, $Z(t)$ is a nonhomogeneous Markov process with generator $A(t, u)$ represented by a linear operator that acts in space $\mathbb{R}^{\mathcal{Z}}$ (i.e., space of real functions h determined on $\mathcal{Z}$). If $h \in \mathbb{R}^{\mathcal{Z}}$ is defined by the set of entries $h = \{h_z\}_{z\in\mathcal{Z}}$, we have

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

where $a_{z,z'}(t,u)$ at $z \ne z'$ is the transition rate $z \to z'$ and $a_{z,z}(t,u)$ is opposite to the exit rate from state z, so that

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

Three variants of transitions are possible: acceptance of a packet by the transmitter, packet sending, and processing at the base station. For each state $(x, y) \in \mathcal{Z}$, we have

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

Then, we assume that control $U(t)$ is described using random process

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

in which service rate $\mu(t)$ and acceptance probability $\vartheta(t)$ are determined by the functions of time and current state:

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

Functions $m_z(t)$ and $v_z(t)$ are called strategies and represent Borel functions with values from intervals $[\underline{m}, \overline{m}]$ and $[\underline{v}, \overline{v}]$, respectively. The strategies are centralized, since they are parametrized using subscript (network state) $z \in \mathcal{Z}$. Notation $\mathcal{U}$ is used for the class of control (2) that employs complete information on the state of the system.

If the queuing systems of the transmitter and the base station are controlled using information on its own states only, the strategies are decentralized and denoted by $m_x(t)$ and $v_y(t)$, where $x \in \mathcal{X}$ and $y \in \mathcal{Y}$. In this case, process (1) is determined using the rule

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

Since $U(t)$ is determined by the current state of controlled process $Z(t)$ with disregard of the preceding evolution U is a Markov control. In addition, $Z(t)$ remains to be a Markov process but its generator is given by a different linear mapping:

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

Note that $m_x(t)$ and $v_y(t)$ must be replaced by $m_z(t)$ and $v_z(t)$, respectively, when the decentralized strategy is applied in the above representation.

We formulate the optimization problem.

As was mentioned, the optimization of the queuing system under study is aimed at the minimization of the number of lost packets at finite interval $[0, T]$. A packet is lost by the transmitter in the case of overflow. Thus, the objective functional is represented as

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

The total service time can be characterized by

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

In the particular case, when the corresponding processes are stationary and ergodic (see §5.8 in [18] for details), we have the Little's law:

Average sojourn time = Average number of units in the system/Arrival rate.

Relationship (5) corresponds to this formula with allowance for the fact that the space and time averaging must be used for the nonstationary system and the arrival rate must be multiplied by the probability of the absence of losses taking into account thinning of the input flow.

The functional that characterizes the energy consumption of the transmitter is given by

$$E[U] = \frac{1}{T}\int_0^T \mathrm{E}\{\mu(t) I\{X(t) > 0\}\}\, dt. \tag{6}$$

Such an expression can easily be interpreted on the assumption that the power consumption of the server of transmitter is proportional to rate μ.

Thus, the optimal control problem is represented as

$$J_0[U] \to \min_{U\in\mathcal{U}} \quad \text{under constraints } S[U] \le \overline{S}, \; E[U] \le \overline{E}, \tag{7}$$

where $\overline{S}$ and $\overline{E}$ are the given upper bounds for the total service time and energy consumption.

Note that (7) is the optimal control problem with complete information on the network. Thus, the optimal control $\hat{U}(t)$ is determined by a centralized strategy.

Below, we present detailed formulation of the optimization problem for a class of decentralized strategies.

4. OPTIMAL CONTROL WITH RESPECT TO THE AUGMENTED CRITERION

The following equivalent representation is possible for optimization problem with complete data (7):

$$J_0[U] \to \min_{U\in \mathcal{U}} : \quad J_1[U] \le 0, \; J_2[U] \le 0. \tag{8}$$

Here, $J_1[U]$ and $J_2[U]$ are integral functionals that contain the above bounds:

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

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

First, we consider unconstrained optimization of the augmented functional:

$$\begin{aligned}\langle\lambda, J[U]\rangle = \lambda_0 J_0[U] + \lambda_1 J_1[U] \\ + \lambda_2 J_2[U] \to \min_{U\in\mathcal{U}},\end{aligned} \tag{11}$$

where $J[U] = \mathrm{col}[J_0[U], J_1[U], J_2[U]]$ is the vector criterion, $\lambda = \mathrm{col}[\lambda_0, \lambda_1, \lambda_2]$ is the vector of nonnegative coefficients, and U is taken from the control class with complete data $\mathcal{U}$.

Below, the solution of problem (11) is considered as a preliminary stage for the subsequent synthesis of the optimal control subject to constraints (8). However, it is expedient to determine strategies that are optimal in the problem without constraints, since such an approach allows preliminary analysis of the sensitivity with respect to weight coefficients λ_l.

Each functional $J_l[\cdot]$ can be represented as an integral of mathematical expectation. Thus, the same representation must be valid for the augmented functional:

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

provided that appropriate functions $g(t, z, u)$ are chosen. Such a functional is represented as

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

with allowance for the fact that $U(t)$ is the control determined by strategies $m_z(t)$ and $v_z(t)$ in accordance with expression (2), $\pi_z(t) = \mathrm{P}\{Z(t) = z\}$ is the state probability, and functions $f_{l,z}(t, m, v)$ at $z = (x, y)$ are given by

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

The optimal predictable control of the Markov process with a finite number of states can be constructed using the method developed in [15]. This control is Markov and provides an optimal solution for the class of controls $U \in \mathcal{U}$ (expression (2)).

For the construction of the optimal control with respect to the augmented criterion

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

the following procedures must be implemented.

(i) The integral representation must be used for the minimized functional

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

for constant control $U(t) \equiv u$, where u is an arbitrary point from the set of values of control action U.

(ii) Function $W(\cdot) = \{W_z(\cdot)\}_{z\in\mathcal{Z}}$ with values from $\mathbb{R}^{\mathcal{Z}}$ must be determined:

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

(iii) Parametric problem of optimization must be solved:

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

(iv) Solution $\phi(t,\lambda) = \{\phi_z(t,\lambda)\}_{z\in\mathcal{Z}}$ must be obtained for the dynamic programming equation:

$$\begin{aligned} \dot{\phi}_z(t,\lambda) = -\min_{u\in\mathrm{U}} W_z(t, \phi(t,\lambda), u, \lambda), \\ t \in [0,T], \; \phi_z(T,\lambda) = 0. \end{aligned} \tag{18}$$

Then, the desired control and optimal functional are determined using the rule

$$\begin{aligned} \tilde{U}(t,\lambda) = \tilde{u}_{Z(t)}(t, \phi(t,\lambda), \lambda) \\ \text{and} \;\; \min_{U\in\mathcal{U}} \langle\lambda, J[U]\rangle = \langle\phi(0,\lambda), \pi(0)\rangle. \end{aligned} \tag{19}$$

We present expressions for function

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

Figure 1 shows possible transitions from the given state (*x*, *y*). The corresponding expressions for function $W_{x,y}(t,\phi(m,v),\lambda)$ are written as

(a) $\alpha\phi_{1,0} - \alpha\phi_{0,0} - \lambda_1\bar{S}\alpha + \lambda_2(-\bar{E}/T)$, if $x = y = 0$;

(b) $\alpha\phi_{x+1,0} + mv\phi_{x-1,1} - (\alpha + mv)\varphi_{x,0} + \lambda_1(x - \bar{S}\alpha) + \lambda_2(m - \bar{E}/T)$, if $0 < x < M,\ y = 0$;

(c) $mv\phi_{M-1,1} - mv\phi_{M,0} + \lambda_0\alpha + \lambda_1 M + \lambda_2(m - \bar{E}/T)$, if $x = M,\ y = 0$;

(d) $\alpha\phi_{x+1,y} + \nu\phi_{x,y-1} - (\alpha + \nu)\phi_{x,y} + \lambda_1(x + y - \bar{S}\alpha) + \lambda_2(m\mathrm{I}\{x > 0\} - \bar{E}/T)$,
if $x = 0, y > 0$ or $x < M, y = N$;

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

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

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

For brevity, we omit time dependences of $\alpha(t)$.

To determine the optimal strategy, we represent $W_{x,y}(t,\phi(m,v),\lambda)$ as a function of variables *m* and *v* and use notation "..." for the dependence on the remaining variables:

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

Then desired optimal strategy (17) is represented as

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

where $\mathrm{MV}(a;b)$ is the notation for the solution of the minimization problem

$$m(av + b) \to \min_{m,v}: \ \underline{m} \le m \le \bar{m},\ \ \underline{v} \le v \le \bar{v}. \quad (23)$$

It is parametrized using constants *a* and *b* and has the following solution:

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

Then, theorem 3 of [15] can be used to obtain the following result.

Theorem 1. *The Cauchy problem for the system of ordinary differential equations* (18) *has single solution* $\phi(t,\lambda) = \{\phi_z(t,\lambda)\}_{z\in\mathfrak{X}}$ *that coincides with the Bellman function, so that*

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

where g(*t*, *z*, *u*) *is the function from* (12).

In particular, the optimum for problem (11) *is given by*

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

where $\pi(0)$ *is the initial distribution of the process* $Z(t) = (X(t), Y(t))$, $t \in [0, T]$.

The optimal control in problem (11) *is constructed in accordance with expressions* (19) *and* (22):

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

where MV(*a*; *b*) *is given by* (24).

The structure of the above control $\tilde{U}(t,\lambda) = (\mu(t), \vartheta(t))$ can be interpreted as follows.

Difference $a = \phi_{x-1,y+1}(t,\lambda) - \phi_{x,y}(t,\lambda)$ determines costs related to packet sending from the transmitter to the base station. If $a \le 0$, the packet sending must be preferred to packet storage and, in accordance with expression (24), optimal acceptance probability $\vartheta(t)$ is assumed to be equal to the upper bound $\bar{v}$. When $a > 0$, a packet must be stored at the transmitter, so that acceptance probability is minimized $\vartheta(t) = \underline{v}$.

For optimal strategy $\mu(t)$, we choose between alternatives $\underline{m}$ and $\bar{m}$ using comparison of parameter a multiplied by acceptance probability $\vartheta(t)$ and coefficient λ_2 that determines the importance of limitation on energy consumption. When $a\vartheta(t) + \lambda_2 \ge 0$, the energy saving is more important than the costs related to the packet storage at the transmitter. Thus, the service rate is minimized: $\mu(t) = \underline{m}$. If $a\vartheta(t) + \lambda_2 < 0$, the packet sending is more important than the energy consumption and we have $\mu(t) = \bar{m}$.

5. SYNTHESIS OF CONTROL IN THE PROBLEM WITH CONSTRAINTS

We consider problem (8) of the minimization of functional $J_0[U]$ on the control class with complete data $\mathcal{U}$ in the presence of constraints $J_l[U] \le 0$, $l = 1, 2$.

For the construction of the optimal control in this problem, we employ the procedure of [15], which involves the following stages.

(i) The constrained problem must be formulated as the equivalent minimax problem for the augmented criterion

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

where Λ is an appropriate convex compact set that contains coefficient vectors $\lambda = \mathrm{col}[1, \lambda_1, \lambda_2]$ with nonnegative coordinates.

(ii) The following dual problem must be solved:

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

(iii) The desired control is the control that is optimal with respect to the augmented functional given the found coefficient vector, so that $\hat{U}(t) = \tilde{U}(t, \hat{\lambda})$.

Note that minimax formulation (25) is equivalent to the constrained problem if operation "max" is changed by "sup" and set Λ contains arbitrary vectors represented as $\lambda = \mathrm{col}[1, \lambda_1, \lambda_2]$, $\lambda_1 \ge 0$, and $\lambda_2 \ge 0$. Set Λ is bounded if any vector $\hat{\lambda}$ that satisfies the Kuhn–Tucker conditions belongs to the set. The results of [17] show that it is sufficient to determine control $U^o \in \mathcal{U}$ on which the Slater condition is satisfied ($J_l[U^o] < 0$, $l = 1, 2$) and assume that

$$\Lambda = \{\lambda \in \mathbb{R}^3 : \lambda_0 = 1,\ 0 \le \lambda_1 \le c_1,\ 0 \le \lambda_2 \le c_2\}, \tag{27}$$

where constants c_1 and c_2 satisfy the inequality $c_l \ge -J_0[U^o] / J_l[U^o] > 0$.

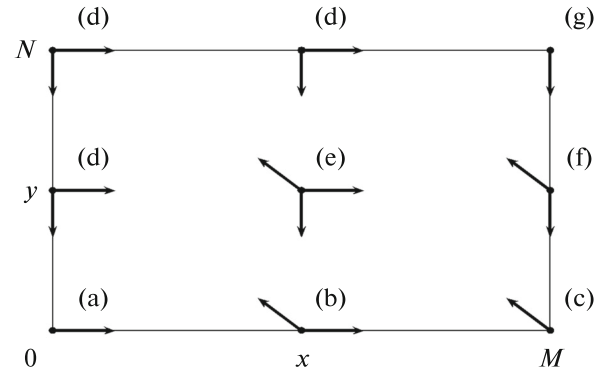

Fig. 1. Transitions between the states of system.

Dual problem (26) is a convex program. Following the approach of [17], we solve the problem using the conditional gradient method [19, 22] or quasi-Newton algorithm [20] adapted for a box-constrained optimization.

For substantiation of the last step, it suffices to show that optimal control $\tilde{U}(t, \lambda)$ continuously depends on coefficient vector λ (see theorem 4 of [15]). However, expressions (22)–(24) cause jumps in optimal service rate $\tilde{m}$ and acceptance probability $\tilde{v}$ when the corresponding coefficients pass through zero. Thus, the continuous dependence of optimal strategy $\tilde{u}_z(t, \phi, \lambda)$ on λ cannot be guaranteed.

Nevertheless, the application of the above procedure (i)–(iii) is possible due to transition to regularized minimax problem [21]:

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

where $\Sigma^{\varepsilon}[U]$ is the stabilizing functional

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

and $\varepsilon_1 > 0$ and $\varepsilon_2 > 0$ are the regularization parameters. We choose integral stabilizer by analogy with functional $\langle \lambda, J[U] \rangle$ and with allowance for the fact that minimum of $\Sigma^{\varepsilon}[U]$ is reached at constant strategy $(\underline{m}, \bar{v})$ providing the minimum energy consumption and the maximum acceptance probability.

The optimal control in the regularized unconstrained problem

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

is constructed using theorem 1 with allowance for the fact that function $\{W_z(\cdot)\}$ that determines the righthand side of dynamic programming equation (18) must be changed by the regularized version

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

and the optimal strategy

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

is uniquely determined from the solution to the problem that is similar to (23)

$$\varepsilon_1(m-\underline{m})^2/2 + m\left(av + b + \varepsilon_2(\overline{v}-v)^2/2\right) \to \min_{m,v}:\ \underline{m}\le m\le\overline{m},\ \ \underline{v}\le v\le\overline{v}.$$

We can easily demonstrate that the solution to such a problem is pair $(\tilde{m},\tilde{v})$, such that

$$\tilde{m} = \begin{cases}\underline{m}, & \tilde{c}\ge 0;\\ \underline{m}-\tilde{c}/\varepsilon_1, & -\varepsilon_1(\overline{m}-\underline{m})\le\tilde{c}\le 0;\\ \overline{m}, & \tilde{c}\le-\varepsilon_1(\overline{m}-\underline{m}),\end{cases}$$
$$\tilde{v} = \begin{cases}\underline{v}, & a\ge\varepsilon_2(\overline{v}-\underline{v});\\ \overline{v}-a/\varepsilon_2, & 0\le a\le\varepsilon_2(\overline{v}-\underline{v});\\ \overline{v}, & a\le 0,\end{cases} \tag{33}$$

where $\tilde{c} = a\tilde{v} + b + \varepsilon_2(\overline{v}-\tilde{v})^2/2$. Using notation $\mathrm{MV}^\varepsilon(a;b)$ for pair $(\tilde{m},\tilde{v})$, we represent strategy (32) as

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

Owing to the regularization, optimal strategy (34) continuously depends on λ and, hence, scheme (i)–(iii) really leads to the solution of minimax problem (28). Moreover, dual problem

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

is a smooth convex program owing to uniqueness of optimal control (30).

Below, we formulate the theorem that determines the properties of regularized problem and its relation to the original problem of the optimal control with constraints.

Theorem 2. *We assume that control $U^o\in\mathcal{U}$ satisfies the Slater condition $J_l\left[U^o\right]<0$ $(l=1,2)$ and set Λ is determined with the aid of rule* (27) *using constants c_1 and c_2 such that $c_l \ge -\left(J_0\left[U^o\right]+\Sigma^\varepsilon\left[U^o\right]\right)/J_l\left[U^o\right]$.*

Then, the following statements are valid.

(i) The optimal control in the regularized unconstrained problem (30) *is calculated as $\tilde{U}^\varepsilon(t,\lambda) = \tilde{u}^\varepsilon_{X(t),Y(t)}\left(t,\phi^\varepsilon(t,\lambda),\lambda\right)$ using strategy* (34) *and function $\phi^\varepsilon(t,\lambda)$ determined from the solution to the Cauchy problem*

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

where $W^\varepsilon(\cdot)$ is the function given by expression (31).

(ii) Objective function in dual problem (35) *given by $\underline{L}^\varepsilon(\lambda) = \left\langle\phi^\varepsilon(0,\lambda),\pi(0)\right\rangle$ is the convex differentiable function with gradient*

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

(iii) Control $\tilde{U}^\varepsilon\left(t,\hat{\lambda}^\varepsilon\right)$ that corresponds to solution $\hat{\lambda}^\varepsilon$ of regularized dual problem (35) *obeys constraints* (8) *and satisfies inequalities*

$$J_0\left[\hat{U}\right] \le J_0\left[\tilde{U}^\varepsilon\left(\cdot,\hat{\lambda}^\varepsilon\right)\right] \le J_0\left[\hat{U}\right] + \Sigma^\varepsilon\left[\hat{U}\right], \tag{38}$$

where $\hat{U}(t)$ is the optimal control for the original problem with constraints (8).

6. OPTIMIZATION ON THE CLASS OF DECENTRALIZED STRATEGIES

We consider the control problem in the absence of complete information on the state of the network. For this purpose, we represent optimization expression (7) for class $\mathcal{U}_d$ of decentralized controls (3) as

$$J_0[U] \to \min_{U\in\mathcal{U}}:\ \ J_1[U]\le 0,\ \ J_2[U]\le 0, \tag{39}$$

where functionals $J_1[U]$ and $J_2[U]$ are given by expressions (9) and (10).

In accordance with expression (3), any control $U(t)$ from $\mathcal{U}_d$ is determined by strategy

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

where $m_x(t)$ and $v_y(t)$ are the service rate at the transmitter and packet acceptance probability at the base station for the system in state (x, y) at moment t. We assume that strategy $\mathbf{u}(t)$ forms a piecewise-continuous function determined on interval $[0, T]$ with values from set $\mathbf{U} = [\underline{m},\overline{m}]^{\mathcal{X}}\times[\underline{v},\overline{v}]^{\mathcal{Y}}$. We use notation $\mathcal{K}$ for such a class of functions.

By analogy with expression (13), each functional $J_l[U]$ is represented as

$$J_1(\pi,\mathbf{u}) = \int_0^T \sum_{(x,y)\in\mathcal{X}} f_{l,(x,y)}\left(t, m_x(t), v_y(t)\right)\pi_{x,y}(t)\,dt, \quad l=0,1,2. \tag{41}$$

Here, distribution of the states of network $\pi(t) = \{\pi_{x,y}(t)\}$ and strategy (40) serve as arguments.

Then, we may proceed from expression (39) to the problem of optimal control

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

for deterministic system

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

that is considered on [0, T] with fixed initial condition π^o and piecewise-continuous controls (40).

The shortened representation of system of differential equations (43) is written as

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

using notation $\bar{A}(t, \mathbf{v})$ that determines a linear operator in space $\mathbb{R}^{\mathcal{X}}$ at fixed $t \in [0, T]$ and $\mathbf{v} \in \mathbf{U}$. Evidently, $\bar{A}(t, \mathbf{u}(t))$ coincides with the generator of the controlled Markov process that corresponds to strategy $\mathbf{u}(\cdot)$.

A similar notation must be used for short-form representation of the vector consisting of functionals (41)

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

At fixed arguments $t \in [0, T]$ and $\mathbf{v} \in \mathbf{U}$, $\bar{F}(t, \mathbf{v})$ is a linear operator from $\mathbb{R}^{\mathcal{X}}$ to $\mathbb{R}^3$.

The maximum principle for deterministic systems makes it possible to formulate the necessary condition for optimality for the class of controlled queuing networks under study.

Theorem 3. *If piecewise-continuous strategy*

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

determines an optimal solution $\hat{U}(t) = (\hat{m}_{X(t)}(t), \hat{v}_{Y(t)}(t))$ *to the control problem with incomplete information* (39), *there exist vector and function*

$$\hat{\lambda} = \operatorname{col}[\hat{\lambda}_0, \hat{\lambda}_1, \hat{\lambda}_2], \; \hat{\lambda}_l \ge 0, \; l = 0, 1, 2,$$
$$\phi(t) = \{\varphi_{x,y}(t) : x \in \mathcal{X}, y \in \mathcal{Y}\},$$

that are not simultaneously zeros

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

such that

(i) at each moment t

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

where function $\bar{W}(\cdot)$ *is represented as*

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

and pair $\pi(t)$, $\phi(t)$ *forms a solution to two-point boundary-value problem*

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

(ii) the complementary slackness condition is satisfied

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

We consider the problem of minimization (46). For brevity, we use variables π, $\varphi \in \mathbb{R}^{\mathcal{X}}$ instead of functions $\pi(t)$ and $\varphi(t)$. Given $\mathbf{v} = \{m_x, v_y : x \in \mathcal{X}, y \in \mathcal{Y}\}$ and notation (16), the function that is minimized in [46] is written as

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

Using expression (21), we represent this relationship as

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

Unfortunately, explicit minimization of this expression is impossible. Its bilinear structure shows that the desired minimum with respect to $\mathbf{v}$ is reached at the vertexes of parallelepiped $\mathbf{U}$. Therefore, the direct minimization can be implemented only at a relatively small number of states. In addition, probabilities $\pi_{x,y}(t)$ and dual variables $\phi_{x,y}(t)$ are needed to determine $\mathbf{u}(t)$. Therefore, two-point boundary-value problem (48) cannot be reduced to the Cauchy problems. An additional obstacle for the application of Theorem 3 is related to unguaranteed positivity of coefficient λ_0. As distinct from the dynamic programming equation (Theorem 1), conditions (46)–(49) that are determined by the maximum principle do not allow formulation of a practical algorithm for solving the optimal control problem with constraints.

For the synthesis of a numerical method for optimization, we consider a class of stationary controls

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

where strategy **u** belongs to set **U**. Notation S is used for the class of stationary controls. With allowance for the fact that **u** represents a point in a finite-dimensional space, we use notation $J_l(\mathbf{u})$ for functional in expression (41) and omit the dependence on distribution $\pi(t)$ that is unambiguously determined using the system of differential equations

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

For the numerical solution of the problem

$$J_0(\mathbf{u}) \to \min_{\mathbf{u} \in \mathbf{U}}, \quad J_1(\mathbf{u}) \le 0, \quad J_2(\mathbf{u}) \le 0, \quad (52)$$

we use the augmented Lagrangian method [19, 22]. The augmented Lagrangian for problem (52) is written as

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

where $\lambda = \mathrm{col}[1, \lambda_1, \lambda_2]$ is the vector of Lagrange multipliers, ε_1 and ε_2 are the regularization parameters, and $(\cdot)_+$ denotes positive part.

The numerical method involves sequential application of the gradient descent (ascent) with respect to variables **u** and λ, respectively. However, a step of the gradient method with respect to strategy **u** is changed in the below algorithm by the numerical solution of the minimization problem. The description of a detailed implementation of such an approach follows the formulation of the algorithm.

Algorithm 1. Error level $\delta_\lambda > 0$ and a rate of a decrease in the norm of gradient $\delta_N \in (0, 1)$ must be determined. The initial conditions must also be determined for regularization parameters $\varepsilon_1^{(0)}, \varepsilon_2^{(0)} > 0$, norm of gradient $N_\lambda^{(0)} = +\infty$, and strategy $\mathbf{u}^{(0)} = \{m_k, v_q \colon k \in \mathcal{X}, q \in \mathcal{Y}\}$ (where $m_k = \underline{m}$ and $v_q = \bar{v}$), vector of multipliers $\lambda^{(0)} = \mathrm{col}[1, 0, 0]$, and iteration number $s = 0$.

(i) Strategy $\mathbf{u}^{(s+1)}$ is obtained as a solution to the problem

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

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

(ii) A gradient step must be performed with respect to λ:

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

(iii) The norm of gradient must be calculated at the given point:

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

where the partial derivatives are represented as

$$\frac{\partial M^{\varepsilon}\left(\mathbf{u}^{(s+1)}, \lambda^{(s+1)}\right)}{\partial \lambda_l} = \begin{cases} J_l\left(\mathbf{u}^{(s+1)}\right), \text{ if } J_l\left(\mathbf{u}^{(s+1)}\right) \ge -\varepsilon_l \lambda_l^{(s+1)}, \\ -\varepsilon_l \lambda_l^{(s+1)}, \text{ otherwise}, \end{cases} \quad l = 1, 2.$$

(iv) The iterations must be terminated if the stopping criterion $N_\lambda^{(s+1)} \le \delta_\lambda$ is satisfied.

(v) The regularization parameters must be corrected:

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

(vi) Number s must be increased by unity and the process must be continued from step (i).

The solution to auxiliary problem (53) is found with the aid of the quasi-Newton algorithm of [20] adapted to the optimization of a smooth convex function on the coordinate parallelepiped. For the implementation of the algorithm, we must calculate gradient $\nabla_{\mathbf{u}} M^{\varepsilon}(\mathbf{u},\lambda)$ the coordinates of which are written as

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

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

The expressions for partial derivatives of functionals $J_l(\mathbf{u})$, ($l = 0, 1, 2$) with respect to strategies m_k and v_q ($k \in \mathcal{X}$ and $q \in \mathcal{Y}$) are obtained from (14):

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

$$\frac{\partial J_l(\mathbf{u})}{\partial v_q} = \int_0^T \sum_{(x,y) \in \mathcal{Z}} f_{l,(x,y)}(m_x, v_y) \frac{\partial \pi_{x,y}(t)}{\partial v_q} dt,$$

where distribution $\pi(t) = \{\pi_{x,y}(t)\}$ is determined from system (51) and corresponds to strategy $\mathbf{u} = \{m_k, v_q, k \in \mathscr{X}, q \in \mathscr{Y}\}$.

The differentiation of the equations of system (51) with respect to control parameters enables us to supplement the differential system with new equations for partial derivatives of state probabilities with respect to service rate m_k

$$\frac{\partial \dot{\pi}_{x,y}(t)}{\partial m_k} = \sum_{(x',y')\in\mathscr{Z}} a_{(x',y'),(x,y)}(t, m_{x'}, v_{y'}) \frac{\partial \pi_{x',y'}(t)}{\partial m_k} + \mathrm{I}\{x = k-1, y > 0\} v_{y-1}\pi_{k,y-1}(t) - \mathrm{I}\{x = k, y < N, k > 0\} v_y \pi_{k,y}(t),$$

$$\frac{\partial \pi_{x,y}(0)}{\partial m_k} = 0,$$

and acceptance probability v_q

$$\frac{\partial \dot{\pi}_{x,y}(t)}{\partial v_q} = \sum_{(x',y')\in\mathscr{Z}} a_{(x',y'),(x,y)}(t, m_{x'}, v_{y'}) \frac{\partial \pi_{x',y'}(t)}{\partial v_q} + \mathrm{I}\{x < M, y = q+1\} m_{x+1}\pi_{x+1,q}(t) - \mathrm{I}\{y = q, q < N, x > 0\} m_x \pi_{x,q}(t),$$

$$\frac{\partial \pi_{x,y}(0)}{\partial v_q} = 0.$$

The resulting expressions are sufficient for the numerical solution of auxiliary problem (53).

7. OPTIMIZATION OF DATA TRANSMISSION IN TWO-AGENT ROBOTIC SYSTEM

We assume that data reception, transmission, and processing are performed over time T in two-agent robotic system consisting of an UAV's transmitter and a receiver of the base station. The input data flow is nonstationary, and its rate is presented in Fig. 2. The parameters of the data-transmission network are as follows:

$$T = 100,\ A = \int_0^T \alpha(t)\,dt \approx 167,\ \max_{t\in[0,T]} \alpha(t) = 2,\ \nu = 2.5,$$

$$\underline{m} = 0.5,\ \overline{m} = 4,\ \underline{v} = 0.05,\ \overline{v} = 1,\ M = 10,\ N = 15,$$

where A is the expected number of packets in the input flow.

The number of states for the network under study is $(M+1)(N+1) = 176$. To avoid localization of the stationary distribution on the states that correspond to the overflow of the base station, we use the service rate that is higher than the input flow intensity: $\nu > \alpha(t)$.

We consider stationary strategy $\mathbf{u}^o = \{m_x^o, v^o\}$:

$$m_x^o = \underline{m} \text{ at } x \le 7 \text{ and } m_x^o = \overline{m} \text{ at } x > 7, \text{ and } v^o = \overline{v} = 1. \quad (54)$$

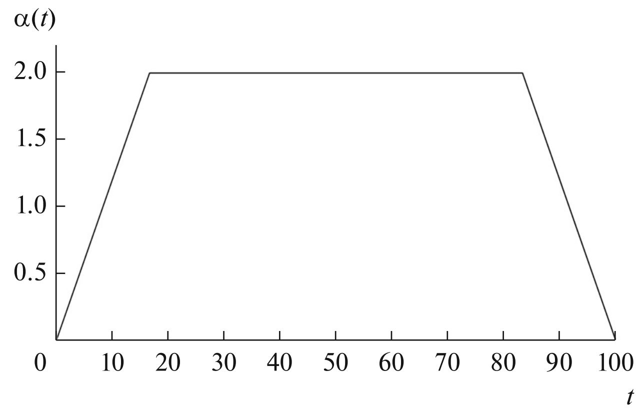

Fig. 2. Arrival rate $\alpha(t)$.

Only two variants are implemented using (54), so that the strategy is called the two-point strategy.

The functionals that describe the total processing time and energy consumption of the UAV's transmitter when the two-point strategy is applied are

$$S(\mathbf{u}^o) = 5.7685,\quad E(\mathbf{u}^o) = 159.3577.$$

We assume that the bounds $\overline{S}$ and $\overline{E}$ are greater about 1%:

$$\overline{S} = 5.8262,\quad \overline{E} = 160.9513.$$

When two-point strategy $\mathbf{u}^o$ is employed, the average number of lost packets and functionals (9) and (10) are

$$J_0(\mathbf{u}^o) = 8.3402,\ J_1(\mathbf{u}^o) = -9.1331,\ J_2(\mathbf{u}^o) = -1.5936.$$

Theorem 2 is used to find strategy $\hat{\mathbf{u}}(t) = \{\hat{m}_{x,y}(t), \hat{v}_{x,y}(t) : x \in \mathscr{X}, y \in \mathscr{Y}\}$ that is optimal on the class of centralized nonstationary controls. In the solution of the dual problem, we use the quasi-Newton algorithm (see [17] for details).

The optimal probability of the packet acceptance is unity: $\hat{v}_{x,y} = 1$. Figure 3 presents time-averaged optimal service rate $\hat{m}_{x,y}$. It is seen that an increase in the number of packets x in the queuing system leads to an increase in service rate $\hat{m}_{x,y}$ and service rate decreases to the minimum allowed level in the vicinity of the overflow of the base station.

To illustrate the evolution of strategy $\hat{m}_{x,y}(t)$, we present results of averaging over states of the two stations (Figs. 4 and 5):

$$\hat{m}(t|X = x) = \mathrm{E}\{\hat{m}_{Z(t)}(t)|X(t) = x\},$$

$$\hat{m}(t|Y = y) = \mathrm{E}\{\hat{m}_{Z(t)}(t)|Y(t) = y\}.$$

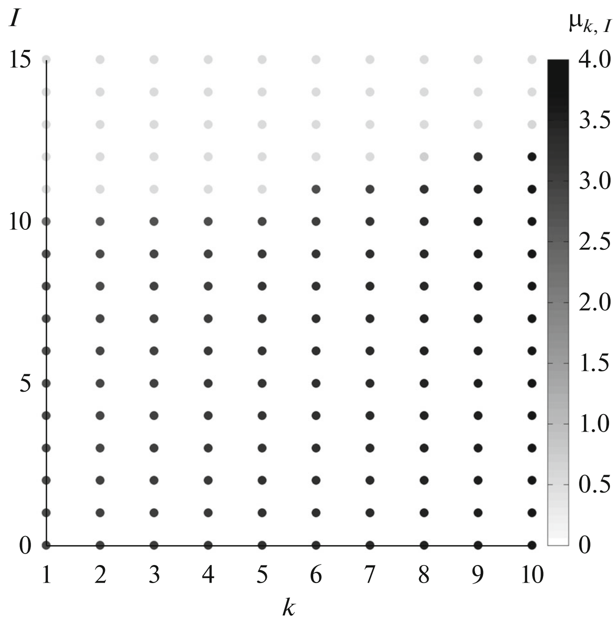

Fig. 3. Time-averaged optimal service rate $\hat{m}_{x,y}$.

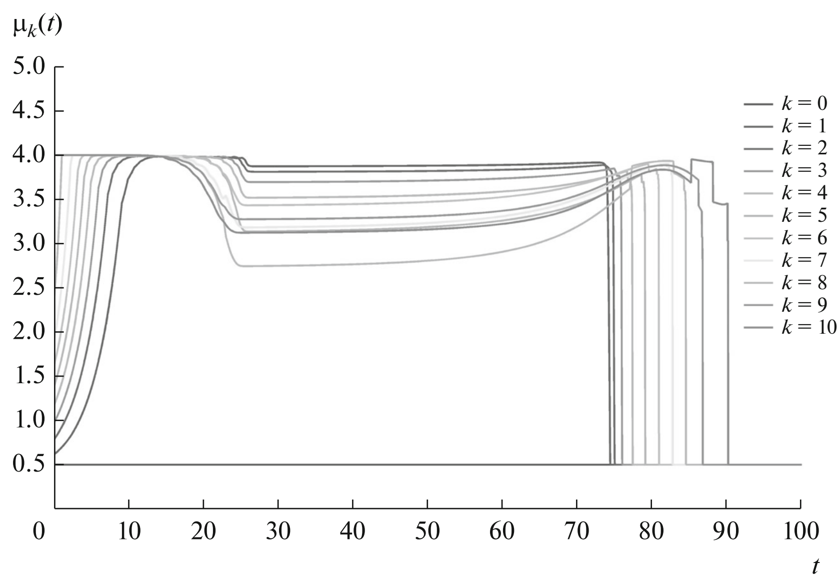

Fig. 4. Optimal service rate averaged over the state of the base station $\hat{m}(t|X = x)$.

Figure 4 shows that changes of strategy $\hat{m}_{x,y}(t)$ follow changes of arrival rate $\alpha(t)$. First, $\hat{m}_{x,y}(t)$ increases to the upper bound $\overline{m}$ and the rate of an increase is higher for more loaded states x. Then, the optimal service rate remains almost constant to the moment when the mode of the input flow is changed. Finally, the rate exhibits a jump in decrease to the lower bound $\underline{m}$. For more loaded states x, the transition to the minimum level is delayed.

Figure 5 proves the dependence of strategy $\hat{m}_{x,y}(t)$ on the evolution of the input flow. Here, we also observe three intervals and the corresponding levels of service rate $\hat{m}_{x,y}(t)$: maximum, intermediate, and minimum. However, a different dependence on the load is obtained for the base station. For the three most loaded states y, the optimal service rate is significantly less than the rate for the remaining states.

Note that state probabilities $\hat{\pi}_{x,y}(t)$ affect the plots of Figs. 4 and 5. Figures 6 and 7 show marginal distri-

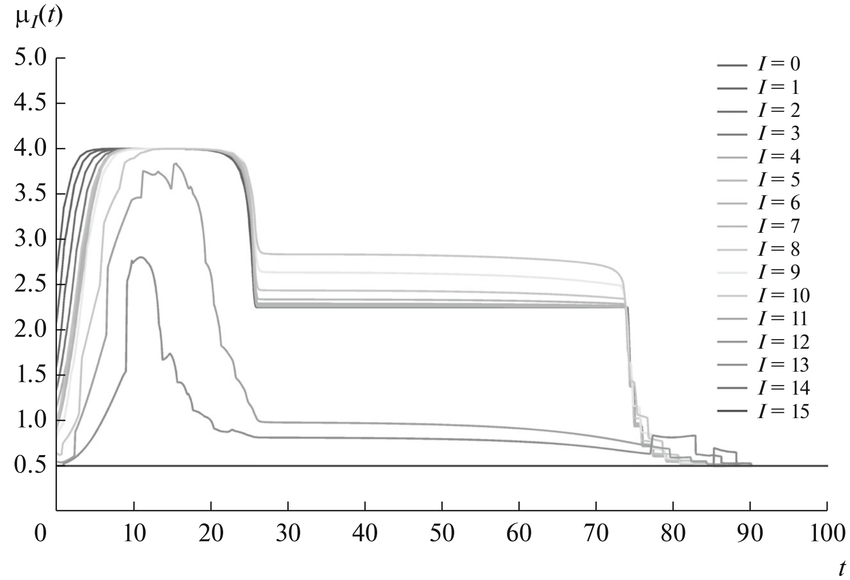

Fig. 5. Optimal service rate averaged over the state of the transmitter $\hat{m}(t|Y = y)$.

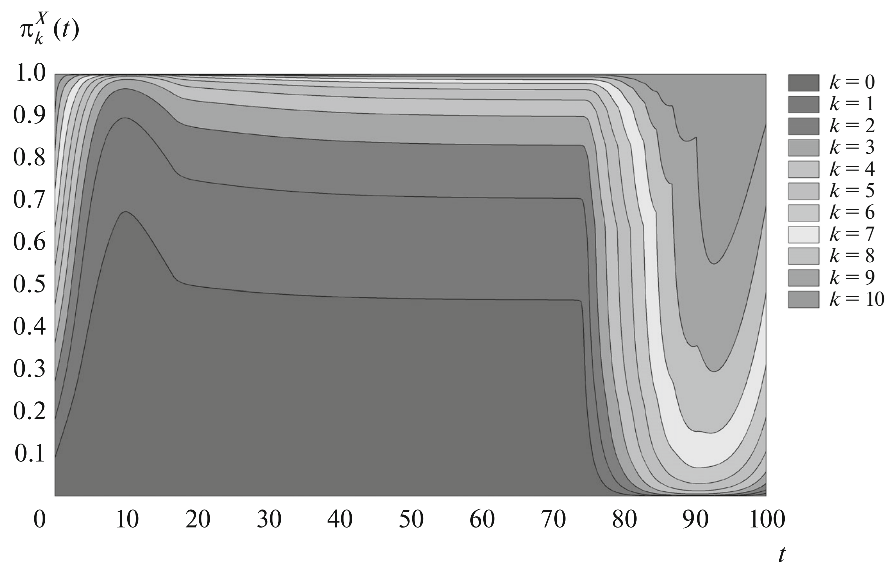

Fig. 6. State probabilities for the transmitter $\hat{\pi}_x^X(t)$ given the optimal centralized control $\hat{\mathbf{u}}$.

butions that correspond to optimal strategy $\hat{\mathbf{u}}(t)$ for the transmitter and base station:

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

Solution $\tilde{\mathbf{u}} = \{\tilde{m}_x, \tilde{v}_y : x \in \mathcal{X}, y \in \mathcal{Y}\}$ of the optimal control problem on the class of stationary decentralized strategies is numerically obtained with the aid of Algorithm 1. As in the case of the centralized control, the optimal strategy of the acceptance probability is $\tilde{v}_y = 1$.

Figure 8 shows the dependence of the optimal service rate $\tilde{m}_x$ on the load of transmitter x. Note similarity to two-point strategy (54) except for difference at $x = 7$. In comparison with the strategy based on the complete information, decentralized strategy $\tilde{m}_x$ exhibits a significantly more active reaction to variations in the load of the transmitter in the absence of data on the state of the base station.

Note that the two-point strategy serves as the initial approximation in Algorithm 1. To verify the insensitivity of the iterative procedure to the starting point, we consider alternative variants of the initial approxima-

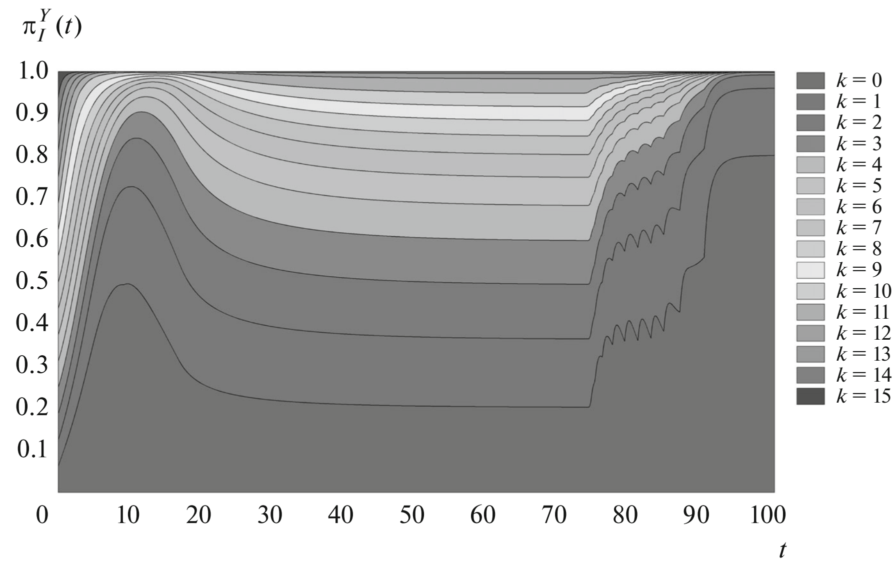

Fig. 7. State probabilities for the base station $\hat{\pi}_y^Y(t)$ given the optimal centralized control $\hat{\mathbf{u}}$.

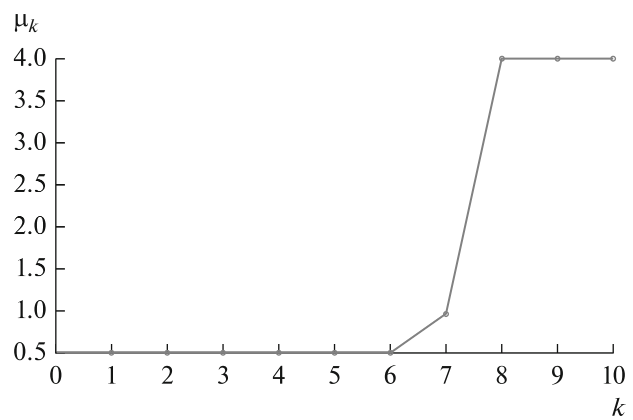

Fig. 8. Optimal decentralized strategy of service rate $\tilde{m}_x$.

tion. The calculations yield almost identical results, so that optimal decentralized strategy $\tilde{m}_x$ is verified.

Figures 9 and 10 present marginal distributions of the queuing system states for the transmitter and base station under optimal decentralized control $\tilde{\mathbf{u}}$. The high-load states are the most probable states for the transmitter, so that the scenario substantially differs from the scenario in which we employ strategy $\hat{m}_{x,y}(t)$ with complete information (cf. with Fig. 6). However, the load distributions for the base station are close to each other (cf. with Fig. 7).

Within class (50), optimal decentralized strategy $\tilde{\mathbf{u}}$ can be compared with the following averaged strategy:

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

where averaging (with respect to time and state of the base system) is performed for optimal service rate $\hat{m}_{x,y}(t)$. Such an approach leads to the decentralized strategy of Fig. 11. Almost constant service rate $\breve{m}_x$ crucially differs from optimal strategy $\tilde{m}_x$ (Fig. 8).

Table 1 presents functionals for four controls: two-point strategy $\mathbf{u}^o$ defined in (54), optimal strategy $\hat{\mathbf{u}}(t)$ constructed in accordance with Theorem 2, optimal decentralized strategy $\tilde{\mathbf{u}}$ resulting from the application of Algorithm 1, and averaged strategy $\breve{\mathbf{u}}$ obtained by the rule (55).

Both numerically synthesized strategies $\hat{\mathbf{u}}(t)$ and $\tilde{\mathbf{u}}$ satisfy the given constraints. Note that the energy

Table 1. Functionals for several controls

$\mathbf{u}$	$J_0(\mathbf{u})$	$J_1(\mathbf{u})$	$J_2(\mathbf{u})$	$S(\mathbf{u})$	$E(\mathbf{u})$
$\mathbf{u}^o$	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
$\breve{\mathbf{u}}$	1.0189	−494.6514	11.6171	2.8400	172.5684
Constraints:		0	0	5.8262	160.9513

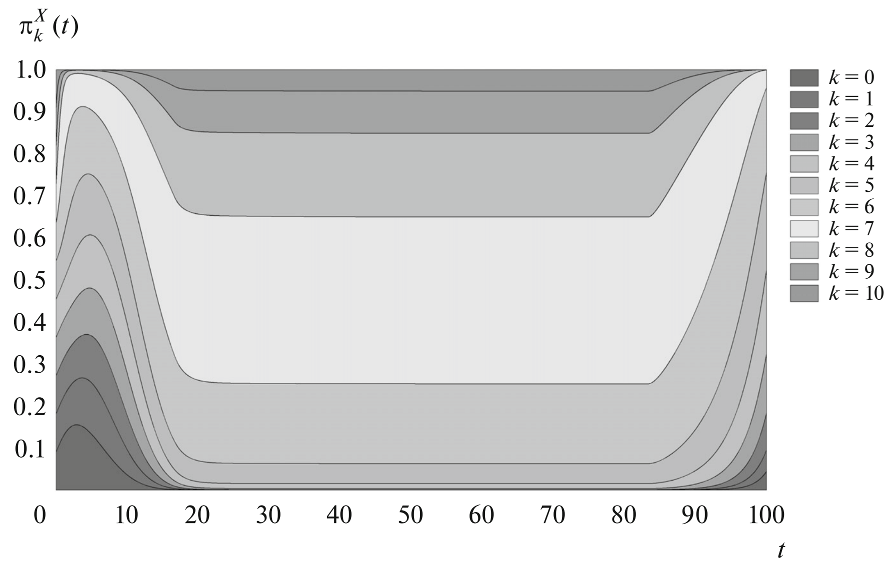

Fig. 9. Probabilities of states of the transmitter $\tilde{\pi}_x^X(t)$ for the optimal decentralized control $\tilde{\mathbf{u}}$.

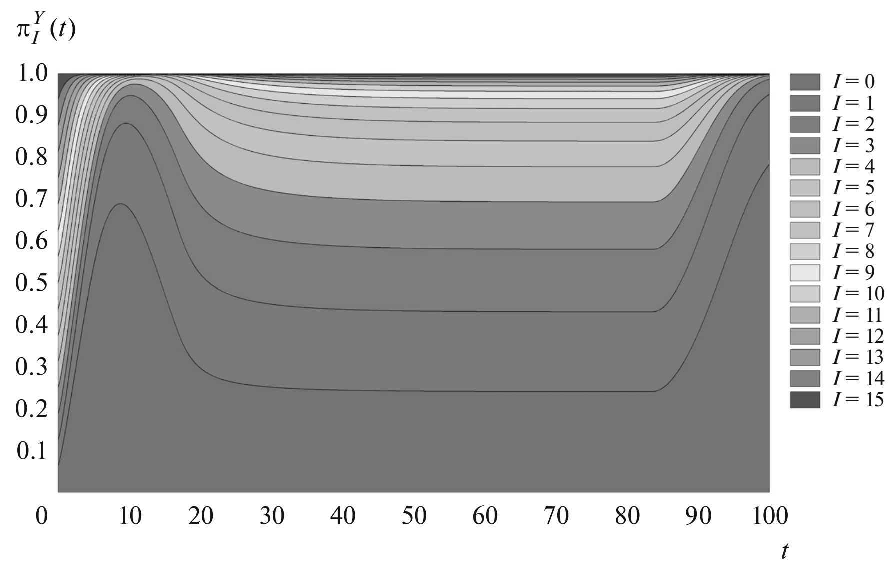

Fig. 10. Probabilities of states of the base station $\tilde{\pi}_y^Y(t)$ for the optimal decentralized control $\tilde{\mathbf{u}}$.

resources are almost exhausted: $J_2(\mathbf{u}) \approx 0$ and $E(\mathbf{u}) \approx \bar{E}$. However, the constraints for the sojourn time of a packet in the network are satisfied with a large margin, since we have $J_2(\mathbf{u}) \ll 0$ and $S(u) \ll \bar{S}$.

With respect to the criterion of the minimum loss number $J_0(\mathbf{u})$, optimal strategy $\hat{\mathbf{u}}(t)$ that employs complete data is significantly better than two-point strategy $\mathbf{u}^o$ and decentralized strategy $\tilde{\mathbf{u}}$ that is optimal for its class. Averaged strategy $\breve{\mathbf{u}}$ makes it possible to minimize the number of lost packets but the constraints on energy consumption are substantially violated: $J_2(\breve{\mathbf{u}}) \gg 0$ and $E(\breve{\mathbf{u}}) \gg \bar{E}$.

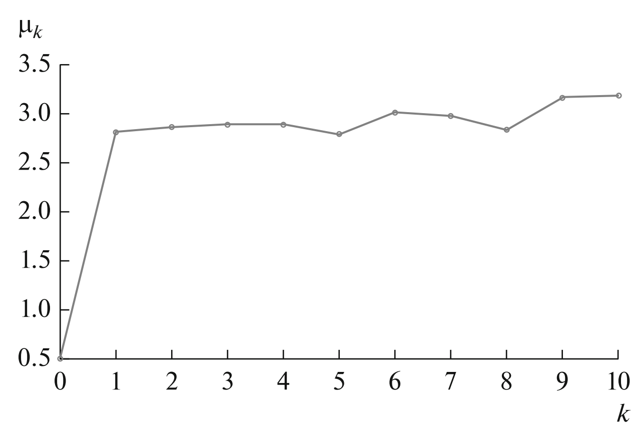

Fig. 11. Service rate $\breve{\mu}_x$ for averaged strategy.

Finally, we compare the results with the conclusions drawn in [3] for a similar model. In accordance with the results of [3], the acceptance probability must not be identically equal to unity (see above): it must decrease to the lower bound when the base station is substantially loaded. Such a difference is caused by a different choice of functionals. Indeed, in [3] two functionals are used to describe the sojourn time of a packet in the transmitter and base station. For the optimal strategies of [3], the constraint on the average total service time in the first station is satisfied with some margin but the time constraint for the second station becomes tight due to active control of the acceptance probability. In the model under study, these two constraints are combined and the network is controlled using only service rate in the first station.

8. CONCLUSIONS

We have considered a controlled two-phase queuing system that contains two single-channel stations with finite buffers and allows blocking of the first server. Such a queuing system is described with the aid of the Markov process that is optimized on a finite time interval using minimization of the average number of lost packets in the presence of constraints on total service time and energy consumption of the first station. We have developed two methods to determine optimal controls on the class of centralized and decentralized strategies. Explicit expressions have been derived for the optimal service rate and acceptance probability using the augmented criterion over the class of centralized strategies. The constrained optimization is performed with the aid of dual optimization method. The necessary conditions for optimality are obtained for the decentralized control. The optimization of the decentralized control is implemented on the class of stationary strategies using an original iterative procedure. The results of numerical experiment show typical optimal strategies for the data transmission control problem in two-agent robotic system.

ACKNOWLEDGEMENTS

The development of models and algorithms for control of nonstationary data flow in two-agent robotic system (Sections 2, 3, and 7) was supported by the Russian Science Foundation (project no. 16-11-00063). The analysis of the problem of constrained optimization of the two-phase queuing system (Sections 4–6) was supported by the Russian Foundation for Basic Research (project nos. 15-37-20611-mol_a_ved and 16-07-00677-a).

REFERENCES

1. S. A. Dudin and O. S. Dudina, "Help desk center operating model as a two-phase queueing system," Commun. Network Theory **49**, 66–82 (2013).
2. V. M. Vishnevskii, A. A. Larionov, and O. V. Semenova, "Performance evaluation of the high-speed wireless tandem network using centimeter and millimeter-wave channels," Probl. Upr., No. 4, 50–56 (2013).
3. N. A. Kuznetsov, D. V. Myasnikov, and K. V. Semenikhin, "Optimal control of data transmission in a mobile two-agent robotic system," J. Commun. Technol. Electron **16**, 1456–1465 (2016).
4. A. Gomez-Corral, "A tandem queue with blocking and Markovian arrival process," Queueing Syst. **41**, 343–370 (2002).
5. L. Broier, A. N. Dudin, V. I. Klimenok, and G. V. Tsarenkov, "A two-phase BM AP/G/1/N → PH/1/M – 1 system with blocking," Autom. Remote Control. **65**, 104–115 (2004).
6. P. P. Bocharov, R. Manzo, and A. V. Pechinkin, "Analysis of a two-phase queueing system with a Markov arrival process and blocking," J. Math. Sci. **132**, 578–589 (2006).
7. Z. Rosberg, P. P. Varaiya, and J. Walrand, "J. Optimal control of service in tandem queues," IEEE Trans. Autom. Control. **27**, 600–610 (1982).
8. B. Hajek, "Optimal control of two interacting service stations," IEEE Trans. Autom. Control. **29**, 491–499 (1984).
9. S. Stidham and R. Weber, "A survey of Markov decision models for control of networks of queues," Queueing Syst. **13**, 291–314 (1993).
10. M. Y. Kitaev and V. V. Rykov, *Controlled Queueing Systems* (CRC, Boca Raton, FL, 1995).
11. G. Koole, "Convexity in tandem queues," Probab. Eng. Inf. Sci. **18** (1), 13–31 (2004).
12. D. F. Silva, B. Zhang, and H. Ayhan, "Optimal admission control for tandem loss systems with two stations," Operat. Res. Lett. **41**, 351–356 (2013).
13. B. Zhang and H. Ayhan, "Optimal admission control for tandem queues with loss," IEEE Trans. Autom. Control. **58**, 163–167 (2013).
14. E. Altman, *Constrained Markov Decision Processes* (Chapman & Hall/CRC, Boca Raton, FL, 1999).
15. B. M. Miller, G. B. Miller, and K. V. Semenikhin "Methods to design optimal control of Markov process with finite state set in the presence of constraints," Autom. Remote Control. **72**, 323–341 (2011).
16. B. M. Miller, G. B. Miller, and K. V. Semenikhin, "Optimal control problem regularization for the Markov process with finite number of states and constraints," Autom. Remote Control. **77**, 1589–1611 (2016).
17. D. V. Myasnikov and K. V. Semenikhin, "Control of M|M|1|N queue parameters under constraints," J. Comput. Syst. Sci. Int. **55**, 59–78 (2016).
18. B. V. Gnedenko, I. N. Kovalenko, *Introduction to Queuing Theory* (Birkhauser, Boston, 1989).
19. B. T. Polyak, *Introduction to Optimization* (Optimization Software, Inc., New York, 1987).
20. D. Kim, S. Sra, and I. S. Dhillon, "Tackling box-constrained optimization via a new projected quasi-Newton approach," SIAM J. Sci. Comput. **32**, 3548–3563 (2010).
21. A. N. Tikhonov and V. Ya. Arsenin, *Solutions of Ill-Posed Problems* (Nauka, Moscow, 1979; Wiley, New York, 1977).
22. D. P. Bertsekas, *Nonlinear Programming*, 2nd Ed. (Athena Scientific, Belmont, MA, 1999).

Translated by A. Chikishev