Available online at www.sciencedirect.com

ScienceDirect

Procedia Engineering 201 (2017) 567–577

Procedia Engineering

www.elsevier.com/locate/procedia

3rd International Conference "Information Technology and Nanotechnology, ITNT-2017, 25-27 April 2017, Samara, Russia

# Two-phase queueing system optimization in applications to data transmission control

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

[a]*Institute of Radio-engineering and Electronics (IRE) of Russian Academy of Sciences, 11 Mokhovaya str., Moscow, 125009, Russia*
[b]*Moscow Institute of Physics and Technology, 9 Institutskiy per., Dolgoprudny, Moscow Region, 141701, Russia*
[c]*Moscow Aviation Institute, 4 Volokolamskoe shosse, Moscow, 125993, Russia*

**Abstract**

We consider the controlled tandem queuing network with two single-server finite-length queues with non-stationary Poisson input flow of packets. The first station controls the service rate and drops any incoming packet if its buffer is full. The second station controls the packet acceptance probability. The tandem is described by a controlled finite-state Markov process and optimized on a fixed time interval by minimizing average number of dropped packets with constraints on the average sojourn time and energy consumption.


Peer-review under responsibility of the scientific committee of the 3rd International Conference "Information Technology and Nanotechnology".

*Keywords:* two-phase queueing system, controlled Markov process, optimal control

## 1. Introduction

Multi-phase (tandem) queuing systems allow modeling processes where jobs sequentially go through several stages. Sequential job service is widely used in practice, e.g., in help desk centers [1], in multimedia data transmission in wireless networks [2] and in data transmission control between nodes of multi-agent robotic systems [3]. To avoid overloading most important stations, multi-phase queuing systems allow jobs blocking. Two-phase systems with blocking were studied in [4,5] where the steady-state probabilities were derived under various assumptions on input flow and service time distribution.

The feedback control problem for two-phase systems was first studied in [6] where the optimality of a threshold-type policy was established given the criterion of minimum average load. This result was obtained using the dynamic programming method for Markov decision processes. This technique was then extended to a more general class of controlled queuing networks [7,8]. Recent studies on optimization of two-phase queuing systems [9,10] consider the optimal access control problem with explicit threshold levels for the optimized criterion. The analysis of these and similar publications shows that the majority of the research in this field is still limited to studying the steady-state

* Corresponding author.
*E-mail address:* kuznetsov@cplire.ru

1877-7058 
Peer-review under responsibility of the scientific committee of the 3rd International Conference "Information Technology and Nanotechnology".
10.1016/j.proeng.2017.09.630

version of the control problem with a single performance functional. So the tandem queues controlled over a finite horizon are still waiting for the comprehensible analysis on the basis of constrained optimization.

The approach of Markov decision processes optimized with respect to several indicators in steady-state mode was developed in [11]. The methods for constrained optimization of continuous-time Markov chains over a finite horizon were described in [12]. The application of this methodology to a single-server queuing system was presented in [13].

In this article, we consider the controlled two-phase queuing system comprised of two single-server stations with finite buffers. The first station is blocked whenever the second station's buffer is full. The first system receives non-stationary Poisson input flow of packets. The packets are processed at a controlled rate on the first server. Incoming packets are dropped in the case of overload of the first system. The second station tries to prevent overload by controlling the acceptance probability (or the blocking probability). The decrease of the acceptance probability leads to a slower packet forwarding from the first system. The queueing network is described by a controlled finite-state Markov process. Its optimization is performed on a fixed time interval by minimizing the average number of dropped packets with constraints on the average sojourn time and energy consumption in the first system. We present an algorithm to synthesize the optimal control on the class of centralized control policies. To illustrate the form of the optimal control policies we also provide a numerical simulation of the data transmission process in the two-agent robotic system.

## 2. Model description

We consider an open tandem queuing network, consisting of two single-server systems: the transmitter and the base station. The transmitter receives packets and after a random delay forwards them to the base station for further processing.

The number of packets on both stations is limited with $M$ and $N$, respectively. If the first queue is full, then any incoming packet is dropped. Packet loss in the network is treated as the most undesirable event, so the probability of dropping the packet should be made as small as possible. To do that we control transmission rate $\mu \geq 0$ of the first station and non-blocking probability $\vartheta \in [0, 1]$ of the second system. The second station rejects the packet coming from the first one with probability $1 - \vartheta$. This blocking mechanism is used to prevent base station overload. If a packet is rejected by the base station, it remains on the transmitter to be retransmitted later. So product $\mu\vartheta$ is an effective transmission rate from the transmitter to the base station.

The queuing network under consideration provides certain sojourn time — the interval between the moment, when a packet comes to the transmitter, and the moment, when it is processed on the base station and leaves it. The processing rate of the base station is considered to be constant.

To decrease both the number of dropped packets and the sojourn time, the transmission rate $\mu$ should be increased. However, $\mu$ is limited due to energy-consumption constraints.

Non-blocking probability $\vartheta$ allows balancing the network state between two extremes: when all packets are accepted ($\vartheta = 1$) and when all packets are blocked ($\vartheta = 0$). Accepting all packets may lead to base station's buffer overflow and, therefore, to collective packet loss.

Since the input flow is non-stationary, both controlled parameters $\mu$ and $\vartheta$ depend on the time instant and system state. So any control policy is supposed to be of the feedback form based on *complete information* on the current state of the network. We call such a control policy *centralized* since the transmitter and base station are working in coordination with each other.

The described model can be applied to a problem of video and sensor data transmission from a UAV to the base station. The transmitter on the UAV converts the original stream of various information into a sequence of unified packets. In this case the transmission rate $\mu$ is defined by the time needed to create and send the packet. The non-blocking probability $\vartheta$ depends on how frequently the base station confirms the successful packet delivery.

So the problem described above requires optimizing the tandem infocommunication network on a fixed time interval taking into account the non-stationary arrivals and constraints on energy consumption.

## 3. Problem statement

In this section we formulate the optimization problem for the queuing network in question.

Let $X(t)$ and $Y(t)$ be a number of packets at the moment $t$ on the transmitter and the base station correspondingly. Then stochastic process $Z(t) = (X(t), Y(t))$, describing the current network state, takes values from the set $\mathcal{Z} = \mathcal{X} \times \mathcal{Y}$, where $\mathcal{X} = \{0, 1, \dots, M\}$, $\mathcal{Y} = \{0, 1, \dots, N\}$.

We assume that packets coming on the transmitter form non-stationary Poisson arrivals with continuous rate $\alpha(t)$. Let us first consider a constant control policy $u = (m, v) \in \mathsf{U}$, where $\mathsf{U} = [\underline{m}, \overline{m}] \times [\underline{v}, \overline{v}]$ and $0 \le \underline{m} \le \overline{m} < \infty, 0 < \underline{v} \le \overline{v} \le 1$ are given bounds for transmission rate and non-blocking probability. Then $Z(t)$ is a non-homogeneous Markov process with the generator $A(t, u)$ described by the transition rate matrix $\{a_{z,z'}(t, u)\}_{z,z' \in \mathcal{Z}}$.

There are only three transitions: packet acceptance on the transmitter; packet forwarding from the transmitter to the base station; packet leaving the base station after processing:

$$a_{(x,y),(x+1,y)}(t, u) = \alpha(t), \quad x < M,$$

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

$$a_{(x,y),(x,y-1)}(t, u) = v, \qquad y > 0.$$

Now let us consider the control policy $U(t)$ as a stochastic process

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

where the transmission rate $\mu(t)$ and non-blocking probability $\vartheta(t)$ are defined as functions of time and current state:

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

The functions $m_z(t)$ and $v_z(t)$ are Borelean and take their values in $[\underline{m}, \overline{m}]$ and $[\underline{v}, \overline{v}]$ correspondingly. They are referred to as *control policies*. Since these control policies depend on network state $z \in \mathcal{Z}$ they are *centralized*. Let us denote by $\mathcal{U}$ the class of controls (2).

The controls $U$ belong to the class of *Markov* controls, because $U(t)$ is defined only by the current state of the controlled Markov process $Z(t)$ no matter what previous evolution was.

Now let us come to the optimization problem statement. As mentioned in the previous section, the main criterion for optimization of the queuing network is minimum of the average number of dropped packets on the finite time interval $[0, T]$. Since packet loss happens only when the transmitter's buffer overflows, the performance functional is defined as follows

$$J_0[U] = \int_0^T \mathsf{P}\{X(t) = M\}\alpha(t)\,dt. \tag{3}$$

The functional characterizing the sojourn time can be defined analogously to Little's formula [14, § 5.8]:

$$S[U] = \frac{1}{T}\int_0^T \mathsf{E}\{X(t) + Y(t)\}\,dt \Bigg/ \frac{1}{T}\int_0^T \mathsf{P}\{X(t) < M\}\alpha(t)\,dt. \tag{4}$$

To define the functional related to transmitter energy consumption, we assume that the power consumed by the transmitter server is directly proportional to the transmission rate $\mu$:

$$E[U] = \int_0^T \mathsf{E}\{\mu(t)\mathrm{I}\{X(t) > 0\}\}\,dt. \tag{5}$$

Thus, the optimal control problem can be stated in the following form:

$$J_0[U] \to \min_{U \in \mathcal{U}} \quad \text{under constraints} \quad S[U] \le \overline{S}, \quad E[U] \le \overline{E}, \tag{6}$$

where $\overline{S}$ and $\overline{E}$ are the upper bounds on sojourn time and energy consumption.

## 4. Optimal control for extended functional

We transform optimal control problem (6) to the following equivalent form:

$$J_0[U] \to \min_{U \in \mathcal{U}}: \quad J_1[U] \le 0, \quad J_2[U] \le 0, \tag{7}$$

where

$$J_1[U] = \int_0^T \mathsf{E}\{X(t) + Y(t) - \overline{S}\, \mathrm{I}\{X(t) < M\}\alpha(t)\}\, dt, \tag{8}$$

$$J_2[U] = \int_0^T \mathsf{E}\{\mu(t)\, \mathrm{I}\{X(t) > 0\} - \overline{E}/T\}\, dt. \tag{9}$$

In this section we consider the unconstrained optimization problem for a linear combination of the functionals in (7)

$$\langle \lambda, J[U] \rangle = \lambda_0 J_0[U] + \lambda_1 J_1[U] + \lambda_2 J_2[U] \to \min_{U \in \mathcal{U}}, \tag{10}$$

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 a vector of non-negative coefficients and $U$ belongs to the class $\mathcal{U}$.

Solution to problem (10) is the first step for optimal control synthesis in the constrained problem (7). Nevertheless, solving unconstrained optimization problem is of interest on its own, because it allows analyzing the sensitivity of the solution with respect to the weight coefficients $\lambda_l$.

Since each of functionals $J_l[\cdot]$ has the form of integrated mathematical expectation, the same form is also valid for their linear combination

$$\langle \lambda, J[U] \rangle = \int_0^T \mathsf{E}\langle \lambda, g(t, Z(t), U(t)) \rangle\, dt \tag{11}$$

given an appropriate choice of the function $g(t, z, u)$. Let us represent this functional in the form

$$\langle \lambda, J[U] \rangle = \int_0^T \sum_{l=0}^{2} \sum_{z \in \mathcal{Z}} \lambda_l f_{l,z}(t, m_z(t), v_z(t)) \pi_z(t)\, dt, \tag{12}$$

where $U(t)$ is the control defined by policies $m_z(t)$ and $v_z(t)$ according to (2), $\pi_z(t) = \mathsf{P}\{Z(t) = z\}$ is the network state probabilities and the functions $f_{l,z}(t, m, v)$ for $z = (x, y)$ are defined as follows:

$$f_{l,(x,y)}(t, m, v) = \begin{cases} \mathrm{I}\{x = M\}\alpha(t), & l = 0, \\ x + y - \overline{S}\, \alpha(t)\, \mathrm{I}\{x < M\}, & l = 1, \\ m\, \mathrm{I}\{x > 0\} - \overline{E}/T, & l = 2. \end{cases} \tag{13}$$

Methods of optimal control synthesis for controlled finite-state Markov processes were developed in [12]. The desired control is shown to be Markov and optimal on the class of all predictable control policies. Therefore, the same control is also optimal among policies (2).

So to determine the control optimal under the extended criterion (10)

$$\widetilde{U}(\cdot, \lambda) \in \arg\min_{U \in \mathcal{U}} \langle \lambda, J[U] \rangle, \tag{14}$$

it is sufficient to do the following:

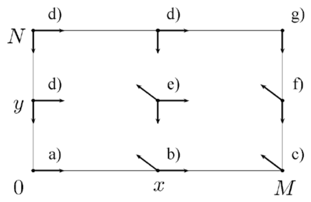


Fig. 1. Transitions between network states.

1. Represent the objective functional in the integral form

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

where $U(t) \equiv u$ is any constant control ($u \in \mathsf{U}$);

2. Define the function $W(\cdot) = \{W_z(\cdot)\}_{z \in \mathcal{Z}}$ with values in $\mathbb{R}^{\mathcal{Z}}$

$$W(t, \phi, u, \lambda) = A(t,u)\phi + F^*(t,u)\lambda, \quad t \in [0,T],\ \phi \in \mathbb{R}^{\mathcal{Z}},\ u \in \mathsf{U}; \tag{15}$$

3. Solve the parametric minimization problem

$$\widetilde{u}_z(t, \phi, \lambda) \in \arg\min_{u \in \mathsf{U}} W_z(t, \phi, u, \lambda); \tag{16}$$

4. Find the solution $\varphi(t, \lambda) = \{\varphi_z(t,\lambda)\}_{z \in \mathcal{Z}}$ of the dynamic programming equations

$$\dot{\varphi}_z(t, \lambda) = -\min_{u \in \mathsf{U}} W_z(t, \varphi(t,\lambda), u, \lambda), \quad t \in [0,T], \quad \varphi_z(T, \lambda) = 0. \tag{17}$$

Then the desired control and the optimal value are given by:

$$\widetilde{U}(t,\lambda) = \widetilde{u}_{Z(t)}(t, \varphi(t,\lambda), \lambda) \quad \text{and} \quad \min_{U \in \mathcal{U}} \langle \lambda, J[U] \rangle = \langle \varphi(0,\lambda), \pi(0) \rangle. \tag{18}$$

Let us represent the function

$$W_z(t, \phi, u, \lambda) = \sum_{z' \in \mathcal{Z}} a_{z,z'}(t,u)\, \phi_{z'} + \sum_{l=0}^{2} \lambda_l f_{l,z}(t,u). \tag{19}$$

Fig. 1 shows possible transitions from the given state $(x, y)$. Depending on the case, the expression for $W_{x,y}(t, \phi, (m,\nu), \lambda)$ takes one of the forms:

a) $\alpha\phi_{1,0} - \alpha\phi_{0,0} - \lambda_1 \overline{S}\alpha + \lambda_2(-\overline{E}/T)$, if $x = y = 0$;
b) $\alpha\phi_{x+1,0} + m\nu\phi_{x-1,1} - (\alpha + m\nu)\phi_{x,0} + \lambda_1(x - \overline{S}\alpha) + \lambda_2(m - \overline{E}/T)$, if $0 < x < M$, $y = 0$;
c) $m\nu\phi_{M-1,1} - m\nu\phi_{M,0} + \lambda_0\alpha + \lambda_1 M + \lambda_2(m - \overline{E}/T)$, if $x = M$, $y = 0$;
d) $\alpha\phi_{x+1,y} + \nu\phi_{x,y-1} - (\alpha + \nu)\phi_{x,y} + \lambda_1(x + y - \overline{S}\alpha) + \lambda_2(m\,\mathrm{I}\{x > 0\} - \overline{E}/T)$, if $x = 0$, $y > 0$ or $x < M$, $y = N$;
e) $\alpha\phi_{x+1,y} + m\nu\phi_{x-1,y+1} + \nu\phi_{x,y-1} - (\alpha + m\nu + \nu)\phi_{x,y} + \lambda_1(x + y - \overline{S}\alpha) + \lambda_2(m - \overline{E}/T)$, if $0 < x < M$, $0 < y < N$;
f) $m\nu\phi_{M-1,y+1} + \nu\phi_{M,y-1} - (m\nu + \nu)\phi_{M,y} + \lambda_0\alpha + \lambda_1(M + y) + \lambda_2(m - \overline{E}/T)$, if $x = M$, $0 < y < N$;
g) $\nu\phi_{M,N-1} - \nu\phi_{M,N} + \lambda_0\alpha + \lambda_1(M + N) + \lambda_2(m - \overline{E}/T)$, if $x = M$, $y = N$.

Here the input rate $\alpha(t)$ is denoted by $\alpha$ for brevity.

To determine the optimal control policy, we represent $W_{x,y}(t, \phi, (m,\nu), \lambda)$ as a function of $m, \nu$, hiding dependence on other variables in the notation "...":

$$W_{x,y}(t, \phi, (m,\nu), \lambda) = \begin{cases} m\lambda_2\,\mathrm{I}\{x > 0\} + \dots, & \text{a), d), g);} \\ m(\nu(\phi_{x-1,y+1} - \phi_{x,y}) + \lambda_2) + \dots, & \text{b), c), e), f).} \end{cases} \tag{20}$$

Then the optimal control policy (16) has the form

$$\widetilde{u}_{x,y}(t,\phi,\lambda) = \begin{cases} (\underline{m},\overline{v}), & x = 0 \text{ or } y = N; \\ \mathrm{MV}(\phi_{x-1,y+1} - \phi_{x,y};\, \lambda_2), & x > 0,\ y < N, \end{cases} \tag{21}$$

where $\mathrm{MV}(a;\, b)$ denotes the solution of the minimization problem with parameters $a$, $b$

$$m(av + b) \to \min_{m,v}: \quad \underline{m} \le m \le \overline{m}, \quad \underline{v} \le v \le \overline{v}. \tag{22}$$

It has the explicit form:

$$\mathrm{MV}(a;\, b) = (\widetilde{m},\widetilde{v}): \qquad \widetilde{m} = \begin{cases} \underline{m}, & a\widetilde{v} + b \ge 0, \\ \overline{m}, & a\widetilde{v} + b < 0, \end{cases} \qquad \widetilde{v} = \begin{cases} \underline{v}, & a > 0, \\ \overline{v}, & a \le 0. \end{cases} \tag{23}$$

Therefore, according to Theorem 3 from [12] we can state the following theorem.

Theorem 1. *The Cauchy problem for the system of ordinary differential equations* (17) *has the unique solution* $\varphi(t,\lambda) = \{\varphi_z(t,\lambda)\}_{z\in\mathcal{Z}}$, *which also serves as a Bellman function, i.e.,*

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

*where* $g(t,z,u)$ *is the function from* (11).

*In particular, the optimal value for problem* (10) *equals*

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

*where* $\pi(0)$ *is the initial probability distribution for the stochastic process* $Z(t)$.

*The optimal control in problem* (10) *is described by* (18) *and* (21), *namely,*

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

*where* $\mathrm{MV}(a;\, b)$ *is introduced in* (23).

The structure of the control $\widetilde{U}(t,\lambda) = (\mu(t), \vartheta(t))$ allows the following interpretation.

Difference $a = \varphi_{x-1,y+1}(t,\lambda) - \varphi_{x,y}(t,\lambda)$ defines costs of forwarding the packet from the transmitter to the base station. If $a \le 0$, then forwarding packet proves to be more profitable than delaying it, so according to (23) the optimal non-blocking probability $\vartheta(t)$ is set to the upper bound $\overline{v}$. Conversely, if $a > 0$, then it is preferable to delay the packet on the transmitter, so the non-blocking probability is set to its minimum value $\vartheta(t) = \underline{v}$.

The optimal policy $\mu(t)$ chooses between minimum $\underline{m}$ and maximum $\overline{m}$ values by comparing the product of $a$ and the non-blocking probability $\vartheta(t)$, on the one hand, and coefficient $\lambda_2$, on the other hand. The product $a\vartheta(t)$ represents the effective cost of forwarding the packet, whereas the coefficient $\lambda_2$ defines the weight of the energy-consumption constraint in the objective functional. If $a\vartheta(t) + \lambda_2 \ge 0$, then saving energy is more important than delaying packet on the transmitter. Therefore, in this case the transmitter rate is set to minimum $\mu(t) = \underline{m}$. If $a\vartheta(t) + \lambda_2 < 0$, then forwarding the packet is more important than energy consumption, so $\mu(t) = \overline{m}$.

## 5. Optimal control synthesis in constrained problem

In this section we consider problem (7) to minimize the objective functional $J_0[U]$ on the class of centralized controls $\mathcal{U}$ under constraints $J_l[U] \le 0$, $l = 1, 2$.

We use the approach suggested in [12] to synthesize the optimal control in this problem. The following steps are needed to do that:

1. Represent the constrained problem in the equivalent minimax form for the extended criterion

$$\hat{U} \in \arg\min_{U \in \mathcal{U}} \max_{\lambda \in \Lambda} \langle \lambda, J[U] \rangle, \tag{24}$$

where $\Lambda$ is a convex compact set, comprised of vectors $\lambda = \mathrm{col}[1, \lambda_1, \lambda_2]$ with non-negative coordinates;
2. Solve the dual optimization problem

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

3. Define the desired control in the form: $\hat{U}(t) = \widetilde{U}(t, \hat{\lambda})$.

It is worth noting that the minimax problem (24) is equivalent to the constrained problem (10) whenever "sup" operation is used instead of "max" and $\Lambda$ contains arbitrary vectors $\lambda = \mathrm{col}[1, \lambda_1, \lambda_2]$, with $\lambda_1 \geq 0$, $\lambda_2 \geq 0$. When choosing the bounded version of $\Lambda$, we should ensure that any vector $\hat{\lambda}$, satisfying KarushKuhnTucker conditions, belongs to the set. To do that according to [13] it is sufficient to find a control $U^o \in \mathcal{U}$ satisfying the Slater condition: $J_l[U^o] < 0$, $l = 1, 2$. Then the bounded set $\Lambda$ can be defined as follows

$$\Lambda = \{\lambda \in \mathbb{R}^3 : \lambda_0 = 1,\ 0 \leq \lambda_1 \leq c_1,\ 0 \leq \lambda_2 \leq c_2\}, \tag{26}$$

where coefficients $c_1$, $c_2$ satisfy $c_l \geq -J_0[U^o]/J_l[U^o] > 0$.

The dual optimization problem (25) is a convex program. Similar to [13], in order to solve it one can use the conditional gradient method [15,16] or the quasi-Newton algorithm adapted to box-constrained optimization [17].

To justify the last step, it is sufficient to ensure that the optimal control $\widetilde{U}(t, \lambda)$ continuously depends on $\lambda$ (see Theorem 4 from [12]). However, according to (21)–(23) when corresponding coefficients pass through zero the optimal transmission rate $\widetilde{m}$ and non-blocking probability $\widetilde{v}$ change abruptly. Therefore, the optimal policy $\widetilde{u}_z(t, \phi, \lambda)$ may not be continuous in $\lambda$.

Nevertheless, the scheme 1–3 described above can be used for the regularized minimax problem [18]:

$$\hat{U}^{\varepsilon} \in \arg\min_{U \in \mathcal{U}} \max_{\lambda \in \Lambda} \langle \lambda, J[U] \rangle + \Sigma^{\varepsilon}[U], \tag{27}$$

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

$$\Sigma^{\varepsilon}[U] = \frac{1}{2} \int_0^T \mathsf{E}\{\varepsilon_1(\mu(t) - \underline{m})^2 + \varepsilon_2 \mu(t)(\overline{v} - \vartheta(t))^2\}\, dt, \tag{28}$$

where $\varepsilon_1 > 0$ and $\varepsilon_2 > 0$ are the regularization parameters. The stabilizing functional has the integral form similar to $\langle \lambda, J[U] \rangle$ and its minimum is achieved on the constant control policy $(\underline{m}, \overline{v})$, which corresponds to minimal energy consumption and maximum non-blocking probability.

The optimal control in the unconstrained regularized problem

$$\widetilde{U}^{\varepsilon}(\cdot, \lambda) \in \arg\min_{U \in \mathcal{U}} \langle \lambda, J[U] \rangle + \Sigma^{\varepsilon}[U] \tag{29}$$

can be found according to Theorem 1 taking into account that the function $\{W_z(\cdot)\}$ in the right-hand side of the dynamic programming equations (17) should be replaced with its regularized version

$$W_z^{\varepsilon}(t, \phi, u, \lambda) = W_z(t, \phi, u, \lambda) + (\varepsilon_1(m - \underline{m})^2 + \varepsilon_2 m(\overline{v} - v)^2)/2, \quad u = (m, v). \tag{30}$$

Now the optimal control policy

$$\widetilde{u}_z^{\varepsilon}(t, \phi, \lambda) \in \arg\min_{u \in \mathsf{U}} W_z^{\varepsilon}(t, \phi, u, \lambda) \tag{31}$$

is uniquely determined from the problem similar to (22)

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

One can easily verify that the solution of this problem is a pair $(\widetilde{m}, \widetilde{v})$ such that

$$\widetilde{m} = \begin{cases} \underline{m}, & \widetilde{c} \ge 0, \\ \underline{m} - \widetilde{c}/\varepsilon_1, & -\varepsilon_1(\overline{m}-\underline{m}) \le \widetilde{c} \le 0, \\ \overline{m}, & \widetilde{c} \le -\varepsilon_1(\overline{m}-\underline{m}), \end{cases} \quad \widetilde{v} = \begin{cases} \underline{v}, & a \ge \varepsilon_2(\overline{v}-\underline{v}), \\ \overline{v} - a/\varepsilon_2, & 0 \le a \le \varepsilon_2(\overline{v}-\underline{v}), \\ \overline{v}, & a \le 0, \end{cases} \tag{32}$$

where $\widetilde{c} = a\widetilde{v} + b + \varepsilon_2(\overline{v}-\widetilde{v})^2/2$. If we denote this pair by $\mathrm{MV}^\varepsilon(a; b)$, then the control policy (31) takes the following form

$$\widetilde{u}^\varepsilon_{x,y}(t,\phi,\lambda) = \begin{cases} (\underline{m}, \overline{v}), & x = 0 \text{ or } y = N; \\ \mathrm{MV}^\varepsilon(\phi_{x-1,y+1} - \phi_{x,y}; \lambda_2), & x > 0,\ y < N. \end{cases} \tag{33}$$

Thanks to regularization, the optimal control policy (33) is continuous in $\lambda$ and, therefore, the scheme 1–3 indeed leads to the solution of minimax problem (27). Moreover, due to uniqueness of optimal control (29), the corresponding dual optimization problem

$$\hat{\lambda}^\varepsilon \in \arg\max_{\lambda\in\Lambda} \underline{L}^\varepsilon(\lambda), \qquad \underline{L}^\varepsilon(\lambda) = \min_{U\in\mathcal{U}} \langle \lambda, J[U]\rangle + \Sigma^\varepsilon[U], \tag{34}$$

is a smooth convex program.

The theorem stated below establishes several properties of the regularized problem and its connection to the original constrained optimal control problem.

*Theorem 2. Let $U^o \in \mathcal{U}$ be a control satisfying the Slater condition: $J_l[U^o] < 0$, $l = 1,2$. Let the set $\Lambda$ be defined according to* (26) *with coefficients $c_1$, $c_2$ such that $c_l \ge -(J_0[U^o] + \Sigma^\varepsilon[U^o])/J_l[U^o]$.*
*Then the following statements hold*:

1. *The optimal control in the unconstrained regularized problem* (29) *can be found in the form*

$$\widetilde{U}^\varepsilon(t,\lambda) = \widetilde{u}^\varepsilon_{X(t),Y(t)}(t, \varphi^\varepsilon(t,\lambda), \lambda)$$

*using the policy* (33) *and the function $\varphi^\varepsilon(t,\lambda)$ found from the Cauchy problem*

$$\dot{\varphi}^\varepsilon_z(t,\lambda) = -\min_{u\in\mathsf{U}} W^\varepsilon_z(t, \varphi^\varepsilon(t,\lambda), u, \lambda), \quad t \in [0,T], \quad \varphi^\varepsilon_z(T,\lambda) = 0, \tag{35}$$

*where $W^\varepsilon(\cdot)$ is the function defined in* (30);

2. *The objective function in the dual problem* (34) *is defined by $\underline{L}^\varepsilon(\lambda) = \langle \varphi^\varepsilon(0,\lambda), \pi(0)\rangle$; it is a convex, differentiable function with gradient*

$$\nabla \underline{L}^\varepsilon(\lambda) = J[\widetilde{U}^\varepsilon(\cdot,\lambda)]; \tag{36}$$

3. *The control $\widetilde{U}^\varepsilon(t,\hat{\lambda}^\varepsilon)$ corresponding to the solution $\hat{\lambda}^\varepsilon$ of the regularized dual problem* (34) *meets constraints* (10) *and satisfies*

$$J_0[\hat{U}] \le J_0[\widetilde{U}^\varepsilon(\cdot,\hat{\lambda}^\varepsilon)] \le J_0[\hat{U}] + \Sigma^\varepsilon[\hat{U}], \tag{37}$$

*where $\hat{U}(t)$ is the control, optimal in the original constrained problem* (10).

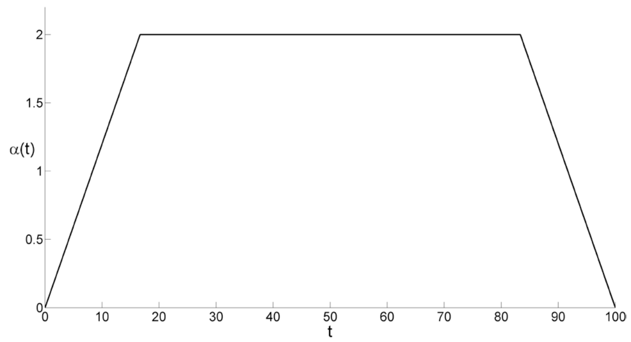


Fig. 2. Input flow rate $\alpha(t)$.

## 6. Data transmission optimization for two-agent robotic system

We consider a two-agent robotic system, consisting of the UAV transmitter and the base station. The transmitter receives and forwards packets to the base station during connection session $T$. The input data flow is non-stationary with rate $\alpha(t)$ (fig. 2).

The data transmission network has the following parameters:

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

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

where A is the expected number of incoming packets.

The number of states for the queuing network equals $(M + 1)(N + 1) = 176$. The processing rate for the base station is chosen higher than the input rate, i.e. $\nu > \alpha(t)$.

Let us consider the two-point control policy $\mathbf{u}^o = \{m^o_x, v^o\}$:

$$m^o_x = \underline{m} \quad \text{if} \quad x \le 7 \quad \text{and} \quad m^o_x = \overline{m} \quad \text{if} \quad x > 7, \quad v^o = \overline{v} = 1. \tag{38}$$

We choose the upper bounds $\overline{S}, \overline{E}$ for functionals characterizing the sojourn time and energy-consumption for the UAV to be approximately 1 % higher than $S(\mathbf{u}^o)$ and $E(\mathbf{u}^o)$ at the two-point policy.

The optimal control policy $\hat{\mathbf{u}}(t) = \{\hat{m}_{x,y}(t), \hat{v}_{x,y}(t)\colon\ x \in \mathcal{X},\ y \in \mathcal{Y}\}$ was found using Theorem 2. The quasi-Newton algorithm was used to solve the dual optimization problem [13].

Fig. 3 shows time-averaged optimal transmission rate $\hat{m}_{x,y}$. It allows to make two conclusions: increase in the number of packets $x$ in the transmitter system leads to increase in the transmission rate $\hat{m}_{x,y}$; when $y$ approaches the overload state of the base station, the transmission rate $\hat{m}_{x,y}$ decreases to the minimum level.

To illustrate how control policy $\hat{m}_{x,y}(t)$ evolves in time, fig. 4 shows the transmission rate averaged by state of each of two stations $\hat{m}(t|X = x) = \mathsf{E}\{\hat{m}_{Z(t)}(t) \mid X(t) = x\}$, $\quad \hat{m}(t|Y = y) = \mathsf{E}\{\hat{m}_{Z(t)}(t) \mid Y(t) = y\}$.

The left-hand side of fig. 4 demonstrates the fact that the control policy $\hat{m}_{x,y}(t)$ follows the changes in the input flow intensity $\alpha(t)$. First, $\hat{m}_{x,y}(t)$ increases to the upper bound $\overline{m}$, with the growth rate larger for states $x$, corresponding to a higher load. Then the optimal transmission rate remains almost constant until the input flow rate changes. After that it abruptly decreases to the lower bound $\underline{m}$. For states $x$, corresponding to a higher load, transition to the minimal level happens with some delay. The right-hand side of fig. 4 also confirms dependency of the control policy $\hat{m}_{x,y}(t)$ on the input flow evolution. Here we can also distinguish three intervals and corresponding rate levels $\hat{m}_{x,y}(t)$: maximum, intermediate and minimal. However, dependency on the base station state is different. The optimal transmission rate is significantly lower for three states $y$, corresponding to the maximum base station load, than for other states.

Table 1 contains functional values for the two-point control policy $\mathbf{u}^o$ (38) and the optimal control policy $\hat{\mathbf{u}}(t)$ constructed in accordance with Theorem 2.

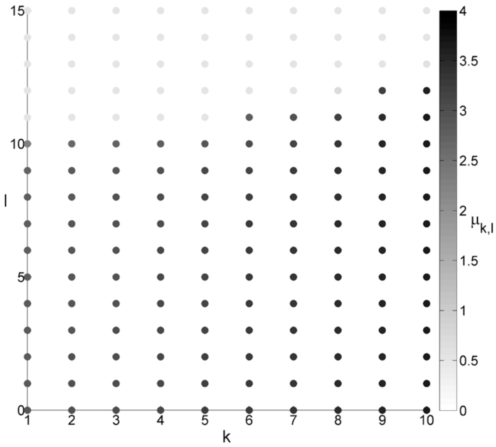


Fig. 3. Time-averaged optimal transmission rate $\hat{m}_{x,y}$.

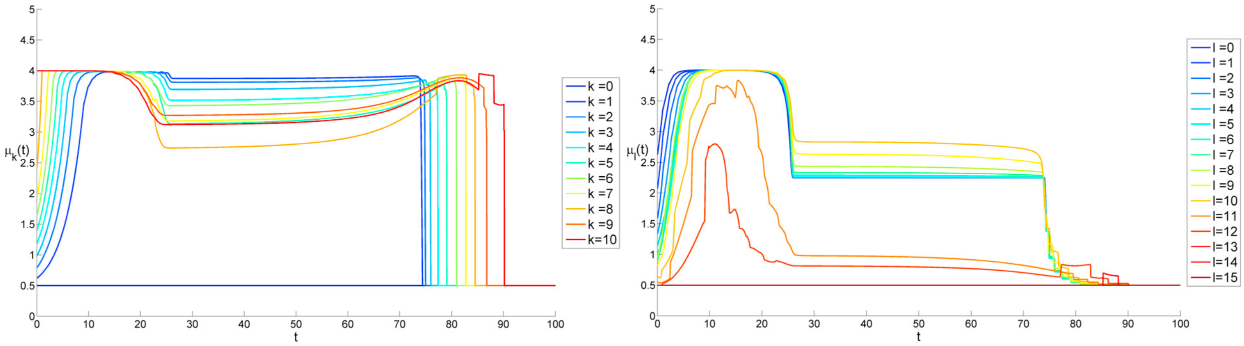


Fig. 4. Optimal transmission rate averaged by state of: $\hat{m}(t|X = x)$ — base station (left); $\hat{m}(t|Y = y)$ — transmitter (right).

Table 1. Functional values and constraints

| $\mathbf{u}$ | $J_0(\mathbf{u})$ | $J_1(\mathbf{u})$ | $J_2(\mathbf{u})$ | $S(\mathbf{u})$ | $E(\mathbf{u})$ |
|---|---|---|---|---|---|
| $\mathbf{u}^o$ | 8.3402 | −9.1331 | −1.5936 | 5.7685 | 159.3577 |
| $\hat{\mathbf{u}}$ | 3.5164 | −424.0999 | −0.0078 | 3.2267 | 160.9435 |
| | Constraints: | 0 | 0 | 5.8262 | 160.9513 |

The numerically determined optimal control policy $\hat{\mathbf{u}}(t)$ meets the specified constraints. At the same time, the energy resources limit is almost completely exhausted: $J_2(\mathbf{u}) \approx 0$ and $E(\mathbf{u}) \approx \overline{E}$. However, the constraint on sojourn time is satisfied with a large margin since $J_2(\mathbf{u}) \ll 0$ or $S(\mathbf{u}) \ll \overline{S}$. The optimal control policy $\hat{\mathbf{u}}(t)$ turned out to be significantly better by the criterion of minimal average number of dropped packets $J_0(\mathbf{u})$ than the two-point policy $\mathbf{u}^o$.

It also should be noticed that the optimal non-blocking probability $\hat{v}$ is always equal to one. This fact means the optimality of the greedy policy in the tandem system under consideration [10].

## Acknowledgments

Research is supported by Russian Science Foundation (project № 16-11-00063)

## References


[1] S. A. Dudin, O. S. Dudina, Help desk center operating model as a two-phase queueing system, Communication Network Theory 49 (2013) 66–82.

[2] V. M. Vishnevsky, A. A. Larionov, O. V. Semenova, Performance evaluation of the highspeed wireless tandem network using centimeter and millimeter-wave channels (in russian), Problemy upravlenija (2013) 50–56.

[3] N. A. Kuznetsov, D. V. Myasnikov, K. V. Semenikhin, Optimal control of data transmission in a mobile two-agent robotic system, Journal of Communications Technology and Electronics 16 (2016) 1456–1465.

[4] A. Gomez-Corral, A tandem queue with blocking and Markovian arrival process, Queueing Systems 41 (2002) 343–370.

[5] D. Breuer, A. N. Dudin, V. I. Klimenok, G. V. Tsarenkov, A two-phase $BMAP/G/1/N \to PH/1/M-1$ system with blocking, Automation and Remote Control 65 (2004) 104–115.

[6] Z. Rosberg, P. P. Varaiya, J. Walrand, Optimal control of service in tandem queues, IEEE Transactions on Automatic Control 27 (1982) 600–610.

[7] S. Stidham, R. Weber, A survey of Markov decision models for control of networks of queues, Queueing Systems 13 (1993) 291–314.

[8] M. Y. Kitaev, V. V. Rykov, Controlled queueing systems, Boca Raton: FL: CRC, 1995.

[9] D. F. Silva, B. Zhang, H. Ayhan, Optimal admission control for tandem loss systems with two stations, Operations Research Letters 41 (2013) 351–356.

[10] B. Zhang, H. Ayhan, Optimal admission control for tandem queues with loss, IEEE Transactions on Automatic Control 58 (2013) 163–167.

[11] E. Altman, Constrained Markov Decision Processes, Boca Raton, FL: Chapman & Hall/CRC, 1999.

[12] B. M. Miller, G. B. Miller, K. V. Semenikhin, Methods to design optimal control of markov process with finite state set in the presence of constraints, Automation and Remote Control 72 (2011) 323–341.

[13] D. V. Myasnikov, K. V. Semenikhin, Control of $m|m|1|n$ queue parameters under constraints, Journal of Computer and Systems Sciences International 55 (2016) 59–78.

[14] B. V. Gnedenko, I. N. Kovalenko, Introduction to Queuing Theory, Birkhauser Boston, 1989.

[15] B. T. Polyak, Introduction to Optimization, Optimization Software, Inc., New York, 1987.

[16] B. D.P., Nonlinear Programming, 2nd ed., Belmont, MA: Athena Scientific, 1999.

[17] D. Kim, S. Sra, I. S. Dhillon, Tackling box-constrained optimization via a new projected quasi-newton approach, SIAM J. Scientific Computing 32 (2010) 3548–3563.

[18] A. N. Tikhonov, V. Y. Arsenin, Solutions of ill-posed problems (in Russian), Moscow: Nauka, 1979.