Optimization of data transmission over a stochastic channel with partially observed state

D V Myasnikov and K V Semenikhin

1Kotelnikov Institute of Radio Engineering and Electronics (RAS), Moscow, Russia, 125009
2Moscow Aviation Institute (National Research University), Moscow, Russia, 125993
3Moscow Institute of Physics and Technology, Dolgoprudnyi, Moscow region, Russia, 141700

e-mail: dmitry.myasnikov@phystech.edu, siemenkv@gmail.com

Abstract. The optimization problem for data transmission over a stochastic channel is considered. The data transmission model is described by the triple “queue–channel–observations”. The queueing system is fed by a non-stationary Poisson stream of data packets for further transmission over a communication channel governed by a non-homogeneous Markov chain. The system is considered with a single server, finite buffer, and service rate proportional to the transmission rate with channel-dependent factor. The observations are formed by round-trip times of sent packets and described by the Markov counting process with channel-dependent intensity. The transmission rate is to be optimized within the class of feedback controls given two performance characteristics: the average number of lost packets and the mean level of energy consumption. The approach proposed for control optimization is based on the optimal filter equations, the complete-information control algorithm, and Monte Carlo techniques. The paper presents both theoretical and simulation results.

1. Introduction

For many of data transmission networks there is a problem of incomplete information about the current state of communication links, queuing systems, servers’ load, etc. Basically, the actual information available for network monitors consists of round-trip times and loss messages in the packet traffic. Despite the fact that these data cannot fully characterize the network state, one can elaborate efficient methods for congestion control in noisy communication lines and partially observed queuing systems [1–3]. The problem of incomplete information is especially important for the communication networks with onboard units acting autonomously in the presence of randomly changing environment [4–6]. Such an equipment aims at achieving several conflicting objectives: minimum losses in incoming useful traffic; decrease of delays in data processing; maximum battery life. These applications motivate further development of optimal control methods for info-telecommunication systems given noisy observations.

In this paper we consider a single-server finite-buffer queuing system on a fixed time interval. Incoming packets enter the system in a non-stationary Poisson stream for further transmission over a loss-free communication channel. Randomly changing conditions of communication affect delays in transmission of packets: the worse the channel state, the longer the time for sending a packet. Losses of incoming packets occur only if the queue is overflowed. Similarly to the Gilbert model [7], the communication channel state is of few variants such as “good” and “bad”. Its dynamics is modelled by a continuous-time Markov chain. Its transition rates are supposed to be time-dependent in order to take into account changing environment conditions caused by
motion of the mobile transmitter according to the known itinerary. Indirect information about the unknown channel state is available from the Markov arrival process used to describe round-trip times of sent packets. Both the service rate in the queuing system and the intensity of the observation stream are proportional to the controlled transmission rate with channel-dependent factors.

Our goal is to propose an approach for constructing feedback control policies and analyzing their performance for the queuing system with the partially observed Markov state. The synthesis of the transmission rate control is based on the optimal filter equations for the unobserved channel state and the aggregated optimization of two performance indices: the average number of lost packets and the mean level of energy consumption. The performance analysis of the constructed policies is performed using the techniques of Monte Carlo simulation.

2. Model formulation and problem statement
Consider two continuous-time Markov chains $X_t$ and $\theta_t$ defined on a fixed time interval $[0, T]$. The controlled state of the queuing system is described by $X_t$ while the hidden state of the communication channel is governed by $\theta_t$. Let $N$ be the largest number of packets that can be placed in the queue and $K$ be the number of aggregated states for the channel. Following the martingale approach in description of stochastic control systems [8], we assume that the state spaces for $X_t$ and $\theta_t$ are comprised of unit column-vectors:

$$\mathbb{S}^X = \{e_0, e_1, \ldots, e_N\} \subset \mathbb{R}^{N+1} \text{ and } \mathbb{S}^\theta = \{f_1, \ldots, f_K\} \subset \mathbb{R}^K.$$  

(1)

Since $X_t$ is a birth-death process, its model is completely specified by two transition rates:

- the arrival rate $a_{n,n+1} = \alpha(t)$ for transition $e_n \rightarrow e_{n+1}$, $n = 0, 1, \ldots, N - 1$,
- the service rate $a_{n,n-1} = m(d, v)$ for transition $e_n \rightarrow e_{n-1}$, $n = 1, \ldots, N$, where $m = \mu_t$ is the controlled transmission rate and $d, v$ is a factor dependent of the channel state $v = \theta_t$ ($\langle \cdot, \cdot \rangle$ denotes the inner product).

The components of $d \in \mathbb{R}^K$ are known positive quantities, which define the following empirical rule: the better the state of the communication channel, the less the average time spent for sending a packet. So we suppose the states $f_1, \ldots, f_K$ are arranged in accordance with

$$d_1 > \ldots > d_K.$$

Thus, the generator of the non-homogeneous Markov process $X_t$ is defined by the matrix $A(t, v, m) = \{a_{n,p}(t, v, m)\}_{n,p=0,1,\ldots,N}$ parameterized by channel state $v$ and transmission rate $m$.

Let $B(t) = \{b_{k,l}(t)\}_{k,l=1,\ldots,K}$ be the generator of the Markov process $\theta_t$. Its transition rates are also assumed to be time-dependent, but the reason of this assumption is to take into account mobility of the transmitter which leads to time-varying behavior of the communication channel.

Indirect information about the hidden state $\theta_t$ is contained in the observation counting process $R_t$, where $R_t$ is equal to the number of packets whose successful delivery is confirmed up to instant $t$. The intensity of $R_t$ depends on the channel state and reflects the following natural assumption: the better the state of communication, the higher the reliability of transmission. So if $\langle c, \theta_t \rangle$ stands for the current intensity of the observation process, then the components of $c \in \mathbb{R}^K$ satisfy inequalities $c_1 > \ldots > c_K > 0$.

The dependence between all the three processes has a statistical sense only: any changes in the one process never lead to an immediate response in the other. So jumps $\Delta X_t$, $\Delta \theta_t$, or $\Delta R_t$ cannot occur at the same time.

Thus, the stochastic control system can be represented in the form

$$dX_t = A^*(t, \theta_t, \mu_t)X_t \, dt + dM^X_t,$$

$$d\theta_t = B^*(t)\theta_t \, dt + dM^\theta_t,$$

$$dR_t = \langle c, \theta_t \rangle \, dt + dM^R_t,$$  

(2)
where $A^*$ and $B^*$ are the transposed matrices and $M^X_t$, $M^θ_t$, $M^R_t$ are pairwise orthogonal martingales [8].

A random process $µ_t$ is called a control with complete information if it:

(a) takes values from a prespecified segment $[m, M]$, where $m > 0$;
(b) has piecewise continuous paths;
(c) is predictable with respect to filtration generated by $X_t$ and $θ_t$.

Analogously, $µ_t$ is called a control with incomplete information if conditions (a) and (b) are fulfilled but assumption (c) is replaced with the following:

(d) $µ_t$ is predictable with respect to filtration generated by $X_t$ and $R_t$.

The classes of controls with complete and incomplete information are denoted by $C$ and $I$, respectively. The assumption $µ_t ∈ C$ means that the current value of $µ_t$ is fully determined by the prehistory $\{X_s, θ_s: s ∈ (0, t)\}$ of the queuing system and the communication channel. But for the case $µ_t ∈ I$ we do not have direct information about the channel state, so the control $µ_t$ depends functionally on the known dynamics of the queue $\{X_s, s ∈ (0, t)\}$ and the observations of the counting process $\{R_s, s ∈ (0, t)\}$.

Consider two performance indices:

$$J_0[µ] = \int_0^T P \{X_t = e_N\} α(t) dt$$

and

$$J_1[µ] = \int_0^T E_{µ_t} dt,$$  \hspace{1cm} (3)

where $J_0$ and $J_1$ are equal to the average number of packets that are lost and sent, respectively. For broadband channels, the transmission rate $µ_t$ is proportional to the signal power. So the functional $J_1$ is treated as a mean level of power consumption.

Now we can formulate the optimal control problem:

$$J_0[µ] → \min_{µ_t ∈ C} \text{ subject to } J_1[µ] ≤ \overline{J}_1,$$  \hspace{1cm} (4)

where $\overline{J}_1$ is a given bound such that $m < \overline{J}_1/T < M$. The goal of (4) is to find a control $\hat{µ}$ that gives the minimum level of losses among all controls with complete information and limited energy consumption over the time horizon $T$. Analogously, the optimal control with incomplete information is defined as a solution to (4), where $C$ is replaced with the class $I$.

3. Control design

3.1. Unconstrained problem

Following [9], the solution $\hat{µ}$ to the optimal control problem (4) can be constructed in the form $\hat{µ} = µ^ο(λ)$, where $µ^ο(λ)$ is the optimal control with respect to the augmented functional:

$$L[µ, λ] = J_0[µ] + λJ_1[µ] → \min_{µ_t ∈ C},$$  \hspace{1cm} (5)

and $\hat{λ}$ is the Lagrange multiplier found from the dual optimization problem

$$L[µ^ο(λ), λ] → \max_{λ ≥ 0}.$$

The unconstrained control problem (5) can be solved using the dynamic programming approach. To do this, we first determine the compound process as a tensor product $X_t ⊗ θ_t$. For two vectors $x ∈ \mathbb{R}^{N+1}$ and $v ∈ \mathbb{R}^K$, we put $x ⊗ v = xv^* ∈ \mathbb{R}^{(N+1)×K}$. This space of matrices is equipped with the inner product $⟨π, φ⟩ = tr[πφ^*]$.
Then we introduce the generator of the compound process. Given fixed time instant \( t \) and control value \( m \), it can be defined as a linear operator acting in \( \mathbb{R}^{(N+1) \times K} \) by the rule:

\[
(C(t,m)[\varphi])_{n,k} = (A(t,f_k,m)\varphi + \varphi B^*(t))_{n,k} \quad \forall \varphi \in \mathbb{R}^{(N+1) \times K},
\]

where \( n = 0, 1, \ldots, N \) and \( k = 1, \ldots, K \), and \((\ldots)_{n,k}\) stands for \((n,k)\)-entry of any matrix expression given in the brackets.

Now we follow the next scheme:

(i) find a matrix-valued function \( F(\cdot) = \{F_{n,k}(\cdot)\} \) such that

\[
L[\mu, \lambda] = \int_0^T \mathbb{E} \langle F(t,\mu_t,\lambda), X_t \otimes \theta_t \rangle \, dt; \quad \text{for} \quad (8)
\]

(ii) determine a matrix-valued function \( W(\cdot) = \{W_{n,k}(\cdot)\} \)

\[
W(t,\varphi,m,\lambda) = C(t,m)[\varphi] + F(t,m,\lambda); \quad \text{for} \quad (9)
\]

(iii) solve a parametric optimization problem

\[
m^o(\cdot) = \{m^o_{n,k}(\cdot)\} : \quad m^o_{n,k}(t,\varphi,\lambda) = \arg \min_{m \in [\underline{m}, \overline{m}]} W_{n,k}(t,\varphi,m,\lambda); \quad \text{for} \quad (10)
\]

(iv) obtain a solution \( \varphi(t,\lambda) = \{\varphi_{n,k}(t,\lambda)\} \) to the system of ordinary differential equations

\[
\dot{\varphi}_{n,k} = -W_{n,k}(t,\varphi,m^o_{n,k}(t,\varphi,\lambda)), \quad t \in [0,T], \quad \varphi|_{t=T} = 0; \quad \text{for} \quad (11)
\]

(v) construct the optimal control and calculate the criterion value

\[
\mu^o_t(\lambda) = \langle m^o(t,\varphi(t,\lambda),\lambda), X_t \otimes \theta_t \rangle, \quad L[\mu^o(\lambda), \lambda] = \langle \varphi(0,\lambda), \pi(0) \rangle, \quad \text{for} \quad (12)
\]

where \( \pi(0) = \mathbb{E}\{X_0 \otimes \theta_0\} \) is the known initial distribution matrix.

All functions \( F(\cdot), W(\cdot), \varphi(\cdot), m^o(\cdot) \), and \( \pi(t) = \mathbb{E}\{X_t \otimes \theta_t\} \) take values in \( \mathbb{R}^{(N+1) \times K} \).

From (3) it follows that

\[
F_{n,k}(t,m,\lambda) = I\{n = N\} \alpha(t) + \lambda(m - \overline{J}_1/T), \quad \text{for} \quad (13)
\]

where \( I\{\ldots\} \) is the indicator of condition \{\ldots\}. Hence, the function \( W(\cdot) \) is completely specified.

In order to find (10), it suffices to represent \( W(\cdot) \) as a function of control value \( m \):

\[
W_{n,k}(t,\varphi,m,\lambda) = \begin{cases} 
\lambda m + \ldots, & n = 0; \\
-\{d_k(t)(\varphi_{n,k} - \varphi_{n-1,k}) - \lambda\} m + \ldots, & n > 0.
\end{cases}
\]

Putting \( c = d_k(t)(\varphi_{n,k} - \varphi_{n-1,k}) - \lambda \) for brevity, we obtain the optimal policy

\[
m^o_{n,k}(t,\varphi,\lambda) = \begin{cases} 
\text{lower bound} \underline{m} & \text{if} \quad c < 0 \quad \text{or} \quad n = 0; \\
\text{any point in} \quad [\underline{m}, \overline{m}] & \text{if} \quad c = 0 \quad \text{and} \quad n > 0; \\
\text{upper bound} \overline{m} & \text{if} \quad c > 0 \quad \text{and} \quad n > 0.
\end{cases}
\]

\[
\text{for} \quad (14)
\]
The optimal policy describes a choice between two extreme decisions on how to transmit data: either on the minimum rate $m_n$ or on the maximum rate $m$. This choice is made in accordance with comparison of two quantities $d_k(t)(\varphi_{n,k} - \varphi_{n-1,k})$ and $\lambda$. The former stands for the gain of decrease in the queue load and the latter reflects the cost of energy consumed by the system. In case $c < 0$, the energy consumption is a more valuable criterion and the transmission rate $\mu_t$ should be assigned as low as possible: $\mu_t = m_n$. In case $c > 0$, transmitting a packet becomes more important than saving energy; so we make the opposite decision: $\mu_t = m$.

By Theorem 3 from [10] we can claim that (12) is an optimal control in the unconstrained problem (4) whenever the policy $\hat{\mu}$ is a solution of the dual problem (6) and the policy $m^\theta(t, \varphi, \lambda)$ is defined by (14) and the matrix $\varphi(t, \varphi, \lambda)$ forms a solution to (11).

3.2. Optimal control under constraints

Theorem 4 from [10] establishes the following fact: $\mu^\theta(\hat{\lambda})$ is optimal in the constrained control problem (4) if $\hat{\lambda}$ is a solution of the dual problem (6) and the policy $m^\theta(t, \varphi, \lambda)$ can be taken as continuous in $\lambda$.

Finding $\hat{\lambda}$ is a one-dimensional convex program (it is easily solved, say, using the golden-section method). But the continuity condition for (14) is violated. So, in this situation, the method of dual optimization cannot be used directly.

Following [11] we adopt the method of Tikhonov regularization. In addition to the augmented functional $L[\mu, \lambda]$ we introduce its regularized counterpart:

$$L^\varepsilon[\mu, \lambda] = L[\mu, \lambda] + \varepsilon \Sigma[\mu], \quad \Sigma[\mu] = \frac{1}{2(m - m)} \int_0^T E(\mu_t - m)^2 \, dt, \quad \varepsilon > 0. \quad (15)$$

After applying the scheme (i)–(v) to (15) we can obtain the regularized version $m^\varepsilon(t, \varphi, \lambda)$ of the optimal policy $\hat{m}$. It is easy to check that $m^\varepsilon(t, \varphi, \lambda)$ depends continuously on $\lambda$.

Now the method of dual optimization leads to the solution $\hat{\mu}^\varepsilon$ of the constrained problem:

$$J_0[\mu] + \varepsilon \Sigma[\mu] \to \min_{\mu \in \mathcal{C}} \quad \text{subject to} \quad J_1[\mu] \leq \mathcal{J}_1. \quad (16)$$

The theory of Tikhonov regularization [12] yields the following relation:

$$J_0[\hat{\mu}] \leq J_0[\hat{\mu}^\varepsilon] + \varepsilon \Sigma[\hat{\mu}^\varepsilon] \leq J_0[\hat{\mu}] + \varepsilon \Sigma[\hat{\mu}].$$

It can be used to evaluate how far $\hat{\mu}^\varepsilon$ is from the original optimal control $\hat{\mu}$:

$$J_0[\hat{\mu}] \leq J_0[\hat{\mu}^\varepsilon] \leq J_0[\hat{\mu}] + \varepsilon (m - m)T.$$

3.3. Control with incomplete information

In order to construct a control with incomplete information, we have to solve the optimal filtering equation for the hidden state $\theta_t$ given the observation process $R_t$. It can be done on-line together with applying the control policies. Let $\psi_t = \mathbb{E}\{\theta_t | R_t\}$ be the conditional expectation of the hidden state with respect to sigma-algebra $\mathcal{R}_t$ generated by the counting observations $\{R_s, s < t\}$. Since $\theta_t \in \mathcal{S}$, the vector $\psi_t$ consists of conditional probabilities

$$\langle \psi_t, f_k \rangle = \mathbb{P}\{\theta_t = f_k | R_t\}, \quad k = 1, \ldots, K.$$

The stochastic differential equation for $\psi_t$ is derived in [8, Chap. 7] in a Zakai form for the unnormalized estimate $\tilde{\psi}_t$:

$$d\tilde{\psi}_t = \left\{ B^\gamma(t) \, dt + (\gamma^{-1} \text{diag}[c] - I) \, d(R_t - \gamma t) \right\} \tilde{\psi}_t, \quad \tilde{\psi}_0 = \pi^0(0).$$

$$\text{IOP Publishing}$$
$$\text{doi:10.1088/1742-6596/1096/1/012176}$$
where $\pi^\theta(0) = E \theta_0$ is the initial channel distribution; $\gamma$ is any positive number used for convenience; $\text{diag}[\ldots]$ and $I$ are the diagonal and identity matrices, respectively. Now the desired estimate $\vartheta_t$ is defined by the Bayes rule

$$\vartheta_t = \tilde{\vartheta}_t / \langle \tilde{\vartheta}_t, e \rangle, \quad e = \text{col}[1, \ldots, 1].$$

Suppose that the policy $\{\hat{m}_{n,k}(t), \quad n = 0, 1, \ldots, N, \quad k = 1, \ldots, K\}$ determines the optimal control $\hat{\mu}_t$. We propose a control with incomplete information of the following form:

$$\tilde{\mu}_t = \sum_{n=0}^{N} \sum_{k=1}^{K} \hat{m}_{n,k}(t) I\{X_{t-} = e_n\} \langle \vartheta_t, f_k \rangle. \quad (17)$$
To explain the structure of this control it suffices to note that the policy

\[ \tilde{\mu}_n(t) = \sum_{k=1}^{K} \hat{m}_{n,k}(t) \langle \vartheta_t, f_k \rangle \]  

(18)

coincides with the conditional expectation of the optimal control \( \hat{\mu}(t) \) given the known queue state \( X_{t-} = e_n \) and available observations \( \{ R_s, s < t \} \).

**Figure 3.** Optimal policy \( \hat{m}_{n,1}(t) \) given queue load \( n = 0, 1, \ldots, 8 \) and channel state “good”.

**Figure 4.** Optimal policy \( \hat{m}_{n,2}(t) \) given queue load \( n = 0, 1, \ldots, 8 \) and channel state “bad”.
4. Numerical experiment

The numerical experiment was performed with the following parameters:

\[ N = 8, \quad K = 2, \quad T = 100, \quad J_1 = 200, \quad m = 0.5, \quad \bar{m} = 8, \quad d = \text{col}[1, 0.5], \quad c = \text{col}[1, 0.1]. \]

The arrival rate is shown in Fig. 1. Two channel states \( f_1 \) and \( f_2 \) will be referred to as “good” and “bad”, respectively. The rate of transition “good” \( \rightarrow \) “bad” and the rate of inverse transition are shown in Fig. 2.

The optimal control policies \( \hat{m}_{n,k}(t) \) are shown in Fig. 3 and 4. For the good state, the optimal transmission rate is assigned at the highest level \( \bar{m} \) on the most part of the time segment and, only in the end, it is switched to the lowest level \( m \). If the channel is in state “bad”, the optimal way for control of the queuing system is to decrease the transmission rate to its lower bound \( m \), except for the case of buffer overflow.

Figures 5 and 6 show how the optimal policy is modified when we use the regularized version of the control problem. Introducing a regularized term leads to splitting in the optimal policy with respect to the queue load. At the same time, there is an evident resemblance of the policy \( \hat{m}_{n,1}(t) \) to the transition rate \( b_{1,2}(t) \). So we can see that the more probable the transition to state “bad”, the higher the transmission rate in state “good”. A similar fact holds for \( \hat{m}_{n,2}(t) \) and \( b_{2,3}(t) \).

The behavior of the control with incomplete information (18) is much more complicated because it depends on fluctuations of conditional probabilities \( \langle \vartheta_t, f_k \rangle, k = 1, 2 \) (see Fig. 7). Nevertheless, the policies \( \hat{m}_n(t) \) corresponding different load levels \( n > 0 \) also coincide with each other on the most part of the time segment.

To compare the performance of the controls described above, we present two curves and two clusters of points in Fig. 8. The blue curve stands for the boundary of the attainability set \( \{(J_0[\mu], J_1[\mu]): \mu \in \mathcal{C}\} \). By definition it contains all values of the vector criterion \( \text{col}[J_0[\mu], J_1[\mu]] \) as \( \mu \) runs over the class of controls with complete information. The green curve is used to describe the attainability set corresponding to the regularized vector criterion \( \text{col}[J_0[\mu] + \varepsilon \Sigma[\mu], J_1[\mu]] \).

The blue curve does not pass below the point \( J_0 = 122, J_1 = 141 \) because the lower part of the boundary is flat and it cannot be found directly using the method of dual optimization [11]. In contrast to the blue boundary, the green is constructed completely due to the regularization method.

In the case of filter-based controls such as (17), we do not have a closed-form representation for the expected characteristics \( J_0[\hat{\mu}] \) or \( J_1[\hat{\mu}] \), so we must resort to the Monte Carlo techniques.

The dots depicted in the figure stand for the path estimates of the functionals \( J_0[\hat{\mu}] \) and \( J_1[\hat{\mu}] \)

\[
\mathcal{J}_0^{(i)} = \int_0^T I\{X_t = e_N\} \alpha(t) \, dt, \quad \mathcal{J}_1^{(i)} = \int_0^T \tilde{\mu}_t \, dt, \quad i = 1, \ldots, M,
\]

where \( \tilde{\mu}_t \) is the control with incomplete information (17).

The crosses denote the sample averages based on a bunch of \( M = 50 \) independent path estimates:

\[
\mathcal{J}_0 = \frac{1}{M} \sum_{i=1}^M \mathcal{J}_0^{(i)}, \quad \mathcal{J}_1 = \frac{1}{M} \sum_{i=1}^M \mathcal{J}_1^{(i)}.
\]

The red and violet colors of dots correspond to two values of the Lagrange multiplier: \( \lambda = 0.5 \) and \( \lambda = 0.4 \).

This and other experiments show that the average estimates lie in the attainability set not far from its boundary. In other words, the superiority of complete-information controls is confirmed.
Figure 5. Regularized policy $\hat{m}_{n,1}(t)$ given queue load $n = 0, 1, \ldots, 8$ and channel state “good”.

Figure 6. Regularized policy $\hat{m}_{n,2}(t)$ given queue load $n = 0, 1, \ldots, 8$ and channel state “bad”.

but not overwhelming. This fact is due to fairly accurate a priori information about transitions of the channel state (see Fig. 2). As expected, the controls of the form (18) demonstrate a similar dependence on the Lagrange multiplier: larger values of $\lambda$ correspond to lower energy consumption and more frequent losses.

5. Conclusion

The stochastic control problem is considered for a non-stationary queuing model of data transmission over a Markov channel. The goal of the problem is to choose the transmission rate that minimizes the average amount of lost data while respecting the resource limit. In the case when the channel state is observed directly, a closed-form expression for the optimal policy is presented. To avoid issues caused by the flat boundary of the attainability set, a regularization technique is described. In the case when the channel state is partially observed by measurements of the counting process, a control synthesis method is proposed. It is based on the filtering algorithm and the structure of the optimal policy. To evaluate performance characteristics of the control policies obtained, a numerical experiment is carried out.
Figure 7. Control policy with incomplete information $\tilde{\mu}_n(t)$ given queue states $n = 0, 1, \ldots, 8$.

Figure 8. The boundaries of the attainability set and its regularized version are shown as a blue and green curve, respectively. For the control with incomplete information $\tilde{\mu}_t$, the average values of losses $J_0$ and power consumption $J_1$ are shown as crosses, while their path estimates are shown as dots. The red color is used for the case $\lambda = 0.5$ and the violet for the case $\lambda = 0.4$.

6. References
[1] Altman E, Avrachenkov K and Barakat C 2000 TCP in presence of bursty losses Performance Evaluation 42 129-147
[2] Miller B, Avrachenkov K, Stepanyan K and Miller G 2005 Flow control as stochastic optimal control problem with incomplete information Problems of Information Transmission 41(2) 150-170
[3] Rieder U and Winter J 2009 Optimal control of Markovian jump processes with partial information and applications to a parallel queueing model Math. Meth. Oper. Res. 70 567-596
[4] Choi D H, Kim S H and Sung D K 2014 Energy-efficient maneuvering and communication of a single UAV-based relay IEEE Trans. Aerosp. Electron. Syst. 50(3) 2319-2326
[5] Miller B, Miller G and Semenikhin K 2017 Optimization of the data transmission flow from moving object to nonhomogeneous network of base stations Proc. 20th IFAC World Congress 6334-6339
Acknowledgement
This work is supported by the Russian Science Foundation (Project No. 16-11-00063).