AN OPTIMAL SLIDING MODE CONGESTION CONTROLLER FOR CONNECTION–ORIENTED COMMUNICATION NETWORKS WITH LOSSY LINKS

ANDRZEJ BARTOSZEWICZ, PIOTR LEŚNIEWSKI

Institute of Automatic Control
Technical University of Łódż, Bohdana Stefanowskiego 18/22, 90-924 Łódż, Poland
e-mail: andrzej.bartoszewicz@p.lodz.pl, piotr.lesniewski2@gmail.com

A new discrete-time sliding-mode congestion controller for connection-oriented networks is proposed. Packet losses which may occur during the transmission process are explicitly taken into account. Two control laws are presented, each obtained by minimizing a different cost functional. The first one concentrates on the output variable, whereas in the second one the whole state vector is considered. Weighting factors for adjusting the influence of the control signal and appropriate (state or output) errors are incorporated in both the functionals. The asymptotic stability of the closed-loop system is proved, and the conditions for 100% bottleneck node bandwidth utilization are derived. The performance of the proposed algorithm is verified by computer simulations.

Keywords: optimal control, sliding-mode control, flow control, discrete-time systems.

1. Introduction

In connection-oriented communication networks, data units sent by sources pass through a series of intermediate nodes before reaching their destinations. If an intermediate node due to a limited data flow rate of its outgoing link cannot pass on all the data it receives, then congestion occurs. Consequently, in order to maximize throughput as well as minimize queuing delays and jitter in modern communication networks, congestion control algorithms are applied. The main difficulty in appropriate congestion control algorithm design is caused by large propagation delays in the networks. The delays are inevitable since information about congestion at a specific node must be dispatched to all sources transmitting data through this node, in order to enable adjustment of the source transmission rates, and this action does not affect the congested node immediately, but only with delay usually called the Round Trip Time (RTT). The problem of congestion control in connection-oriented communication networks has been studied in many papers, and an extensive review of the papers can be found in a recent monograph of Ignaciuk and Bartoszewicz (2013).

The main advantage of sliding mode control methods is their strong robustness with regard to a class of disturbance and model uncertainty (Drazenovic, 1969). The robustness is achieved by employing nonlinear control signals to force the system trajectory to attain in finite time a motion along a pre-determined surface in the state space. The sliding mode approach was originally proposed for controller design (Emelyanov, 1967; Gao et al., 1995; Hung et al., 1993; Levant, 1993; Mnasri and Gasmì, 2011; Sira-Ramirez, 1993; Tomera, 2010; Utkin, 1977; Utkin and Shi, 1996). However, sliding mode techniques can also be used in state observers (Edwards and Spurgeon, 1994; Edwards et al., 2012; Floquet et al., 2007; Fridman et al., 2007; Haskara et al., 1998; Veluvolu and Soh, 2009).

Due to the robustness of sliding mode control, various types of sliding mode congestion controllers have been proposed. Jing et al. (2007) used a sliding mode controller with a state predictor, and established the maximum delay necessary for the system stability. A fuzzy controller combining the advantages of linear and terminal sliding modes was also proposed by Jing et al. (2008) for a simplified delay-free network model. For a DiffServ network, an adaptive sliding mode controller (using the back-stepping procedure) for a model which neglects the feedback latency was presented by Zheng et al. (2008). On the other hand, for a DiffServ network with delay a second order sliding mode technique was applied by Zhang...
et al. (2009) in order to reduce the chattering of the control signal. In the work of Jin et al. (2009) the problem of fair (in the max-min sense) data rate distribution among the sources is considered. A binary congestion signal is used to control the data output of sources, and the analysis of this algorithm is performed for a delay-free system.

The papers mentioned in the previous paragraph use the continuous time network model. However, any flow control algorithm for a data transmission network must be implemented as a digital controller. Therefore, in some works a discrete-time approach to the problem of data flow control was used. For example, a sliding mode controller was presented by Yan et al. (2007), but the result of that paper was derived without considering the system delays. In the work of Yang et al. (2007) it is shown that any max-min fair system with a stable symmetric Jacobian matrix maintains asymptotic stability under arbitrary directional delays. This means that, if the controller is designed so that the system has a symmetric Jacobian matrix, its stability can be examined based on the corresponding undelayed system. A dead-beat sliding mode controller for multi-source networks with a priori known round trip times is presented by Bartoszewicz and Zuk (2009), whereas in the work of Ignaciuk and Bartoszewicz (2008) an LQ optimal sliding mode controller for single-source networks is proposed. The same approach is then extended for multi-source networks by Ignaciuk and Bartoszewicz (2009), who also design a similar optimal flow controller for multi-source networks with the round trip times which may change during the control process (Ignaciuk and Bartoszewicz, 2011).

In most papers published up to now only packet losses due to buffer overflows are considered, and the occurrence of lossy links in the network is neglected. As in real networks transmission losses are inevitable, in this paper we present an LQ optimal sliding mode controller for a single connection in which packets may be lost during the transmission process.

2. Network model

In this paper we consider a single virtual circuit in a connection-oriented network. The virtual circuit consists of a single data source, intermediate nodes and a destination. Data sent by the source are passed from node to node, until they reach their destination. It is assumed that one of the intermediate nodes (further in the paper called the “bottleneck” node) cannot pass on all the data it receives, due to the limited bandwidth of its outgoing link. Thus congestion occurs, and the surplus data accumulate in the buffer located at the bottleneck node. The block diagram of the network is shown in Fig. 1.

We assume that the source is persistent, i.e., it always has enough data to transmit at the maximum rate allowed by the network. Therefore, the congestion problem can be solved through an appropriate adjustment of the data rate of the source. This rate is determined by the controller placed at the bottleneck node. The source receives the signal from the controller (denoted by \( u \)) after backward delay \( T_B \). It then sends the specified amount of data, which reaches the bottleneck queue after forward delay \( T_F \). It is assumed that during the transmission some data packets are lost, so that only \( \alpha u \) (\( \alpha \in (0, 1) \)) data packets arrive at the bottleneck node. The round trip time \( RTT \) (the delay between generating the control signal, and the requested data arriving at the bottleneck queue) can be expressed as a sum of backward and forward propagation delays,

\[
RTT = T_B + T_F. \tag{1}
\]

Further in the paper, \( T \) represents the sampling time. The queue length at time instant \( kT \) is denoted by \( y(kT) \), and its demand value by \( y_d \). It is assumed that before the beginning of data transmission the buffer is empty, i.e., \( y(kT < 0) = 0 \). The controller output at time \( kT \) is denoted by \( u(kT) \). We also assume that the round trip time is a multiple of the sampling time, i.e., \( RTT = mRTT \), where \( mRTT \) is a positive integer.

The amount of data which may leave the bottleneck buffer at each time instant is modeled as an \( a\text{-priori} \) unknown bounded function of time \( d(kT) \), and its maximum value is denoted by \( d_{\text{max}} \). Because at some time instants there can be less data in the queue than can be sent, an additional function of time \( h(kT) \) is introduced, which represents the data actually leaving the bottleneck queue. Consequently,

\[
0 \leq h(kT) \leq d(kT) \leq d_{\text{max}}. \tag{2}
\]

The queue length for any \( kT \geq 0 \) can be expressed as follows:

\[
y(kT) = \alpha \sum_{j=0}^{k-1} u(jT - RTT) - \sum_{j=0}^{k-1} h(jT) = \alpha \sum_{j=0}^{k-mRTT-1} u(jT) - \sum_{j=0}^{k-1} h(jT). \tag{3}
\]

The network can also be described in the state space as

\[
x[(k + 1)T] = Ax(kT) + Bu(kT) + Oh(kT), \quad y(kT) = q^T x(kT), \tag{4}
\]

where \( x(kT) = [x_1(kT) \quad x_2(kT) \quad \cdots \quad x_n(kT)]^T \) is the state vector, \( \dim(x) = n = mRTT + 1 \), \( y(kT) = x_1(kT) \) is the queue length, and

\[
x_i(kT) = u[(k - n + i - 1)T] \tag{5}
\]
An optimal sliding mode congestion controller for connection-oriented communication networks...

3. Proposed control strategy

In this section a control-theoretic approach is applied to design a discrete-time sliding-mode controller for the system considered. We begin with deriving the general form of a sliding-mode controller. The parameters of the sliding hyperplane are then chosen minimizing two different quality criteria. One involves the output of the system, and the other takes into account the whole state vector. Finally, the stability of the closed-loop system and its other important properties are proved.

3.1. Sliding-mode controller design. We introduce a sliding hyperplane described by the following equation:

$$ s(kT) = c^T e(kT) = 0, $$

where $c^T = [c_1 \ c_2 \ \ldots \ c_n]$ is a vector satisfying $c^T b \neq 0$, and $e(kT) = x_d - x(kT)$ denotes the closed-loop system error. Substituting (4) into $c^T e((k + 1)T) = 0$, the following feedback control law can be derived:

$$ u(kT) = (c^T b)^{-1} c^T [x_d - A x(kT)]. $$

Using (6), (7) and (8), we can present (10) as

$$ u(kT) = e_n^{-1} \left[ c_1 y_d - c_1 x_1(kT) - \sum_{i=2}^{n} c_{i-1} x_i(kT) \right]. $$

With the application of this control strategy, the state matrix of the closed-loop system has the form $A_c = [I_n - b(c^T b)^{-1} c^T] A$. The characteristic polynomial of $A_c$ can be found as follows:

$$ \det(zI_n - A_c) = z^n + \sum_{i=2}^{n} c_{i-1} x_i(kT). $$

(11)

The properties of the closed-loop system will be determined by the choice of the sliding hyperplane parameters $c_1, c_2, \ldots, c_n$. In the remaining part of this subsection, two performance indices will be considered. For both of them appropriate selection of elements of vector $c$, minimizing the different cost functionals will be obtained.

Case 1. In optimization problems we often consider a performance index involving values of the control signal and the output error. In this case we seek for a sliding-mode control $u_{opt}(kT)$ that minimizes the following cost functional:

$$ J_1(u) = \sum_{k=0}^{\infty} \left\{ u^2(kT) + w[y_d - y(kT)]^2 \right\}, $$

where $w$ is a positive weighting factor adjusting the influence of the control signal and the output variable on the functional. According to Kwakernaak and Sivan (1972), for the time-invariant discrete-time system (4) the optimal control $u_{opt}(kT)$ that minimizes the cost functional (13) can be presented as

$$ u_{opt}(kT) = -g x(kT) + r. $$

(14)
where
\[
g = b^* K (I_n + bb^* K)^{-1} A,
\]
(15)
\[
r = G_e(1)^{-1} y_d = \left[ q^r (I_n - A + bg)^{-1} b \right]^{-1} y_d.
\]
(16)
where \((\cdot)^*\) denotes the complex conjugate matrix transpose, \(G_e(z)\) is the transfer function of the closed-loop system, and semipositive matrix \(K\) satisfies \(K^* = K\) and is determined by the following Riccati equation:
\[
K = A^* K (I_n + b^* K)^{-1} A + wqq^*.
\]
(17)
Because all elements of \(A, b\) and \(q\) are real numbers, the complex conjugate matrix transpose \((\cdot)^*\) is equivalent to the matrix transpose \((-\cdot)^T\), and the elements of \(K\) are also real numbers. Therefore, condition \(K^* = K\) implies that \(K\) is symmetric.

In the case of the network system considered in this paper, the Ricatti equation needs to be solved analytically for the system of an arbitrary order \(n\). The method proposed here is similar to the one used by Ignaciuk and Bartoszewicz (2008). It involves iterative substitution of \(K\) into the right hand side of Eqn. (17) and comparing the result with its left-hand side, so that at each iteration the number of independent elements of \(K\) is reduced.

We begin with
\[
K_0 = \begin{bmatrix}
k_{11} & k_{12} & k_{13} & \cdots & k_{1n} \\
k_{12} & k_{22} & k_{23} & \cdots & k_{2n} \\
k_{13} & k_{23} & k_{33} & \cdots & k_{3n} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
k_{1n} & k_{2n} & k_{3n} & \cdots & k_{nn}
\end{bmatrix}.
\]
(18)
Because matrix \(K\) is symmetric, in the following equations, in order to save space, we will represent the elements positioned below the diagonal by \(^*\). After the first analytical iteration, we obtain the following form:
\[
K_1 = \begin{bmatrix}
k_{11} & \alpha (k_{11} - w) & k_{13} & \cdots & k_{1n} \\
* & \alpha^2 (k_{11} - w) & k_{23} & \cdots & k_{2n} \\
* & * & k_{33} & \cdots & k_{3n} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
* & * & * & \cdots & k_{nn}
\end{bmatrix}.
\]
(19)
The next step is substituting \(K_1\) given by (19) into (17) and comparing its left and right-hand sides. We arrive at the next form of \(K\), which expresses the values of more elements in terms of \(k_{11}\),

\[
K_2 = \begin{bmatrix}
k_{11} & \alpha (k_{11} - w) & \alpha (k_{11} - 2w) & \cdots & k_{1n} \\
* & \alpha^2 (k_{11} - w) & \alpha^2 (k_{11} - 2w) & \cdots & k_{2n} \\
* & * & \alpha^2 (k_{11} - 2w) & \cdots & k_{3n} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
* & * & * & \cdots & k_{nn}
\end{bmatrix}.
\]
(20)
We repeat this procedure until all elements of \(K\) are expressed as functions of \(k_{11}\), weighting factor \(w\) and system order \(n\).

\[
K = A^* K (I_n + b^* K)^{-1} A + wqq^*.
\]
(17)
In order to determine \(k_{11}\), we substitute (21) into the right-hand side of (17) and compare the first elements of the obtained matrices. This results in
\[
\alpha^2 k_{11}^2 - \alpha^2 w (2n - 1) k_{11} + \alpha^2 w^2 (n^2 - n) - w = 0.
\]
(22)
Equation (22) has the following roots:
\[
k'_{11} = \frac{1}{2\alpha} \left[ \alpha (2n - 1) w + \sqrt{w (\alpha^2 w + 4)} \right]
\]
\[
k''_{11} = \frac{1}{2\alpha} \left[ \alpha (2n - 1) w - \sqrt{w (\alpha^2 w + 4)} \right]
\]
(23)
The determinant of every principal minor of \(K\) is given by \(\det(K_r) = \alpha^2 w^{r-1} k_{11} - (r - 1) w\), where \(r\) is the dimension of the minor considered. Therefore, since \(w > 0\), \(K\) will be semipositive definite if and only if the condition \(k_{11} \geq (n - 1) w\) is satisfied. Therefore, only \(k'_{11}\) guarantees that matrix \(K\) is semipositive definite. Having found \(K\), we derive vector \(g\) by substituting \(k'_{11}\) into (18) and obtain
\[
g = \gamma_1 [1/\alpha 1 \cdots 1],
\]
(24)
where
\[
\gamma_1 = \frac{1}{2} \left( \alpha \sqrt{w (\alpha^2 w + 4)} - \alpha^2 w \right)
\]
(25)
From (16) we obtain
\[
r = \frac{\gamma_1 y_d}{\alpha}.
\]
(26)
Finally, the optimal control law can be written as
\[
u_{opt}(kT) = \gamma_1 \left[ \frac{y_d - y(kT)}{\alpha} - \sum_{i=2}^{n} \xi_i(kT) \right].
\]
(27)

**Case 2.** We can also analyze the situation where instead of the output error the whole state error vector is taken into account. In this case we seek for \(u_{opt}(kT)\) that minimizes the following cost functional:
\[
J_2(u) = \sum_{k=0}^{\infty} \left[ u^2(kT) + e^T(kT) W e(kT) \right],
\]
(28)
where
\[
W = \begin{bmatrix}
w_1 & 0 & \cdots & 0 \\
0 & w_2 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & w_2
\end{bmatrix}.
\]
(29)
with \(w_1\) and \(w_2\) being positive constants that adjust the influence of the queue length error and the amount of data under way error, respectively. The optimal control \(u_{\text{opt}}(kT)\) can be found using (14), (15) and (16). Equation (17) needs to be modified to

\[
K = A^T K \left( I_n + bb^T K \right)^{-1} A + W.
\] (30)

Solving (30) in the same way as before, we arrive at the following form of matrix \(K\):

\[
K = \begin{bmatrix}
  k_{11} & \alpha(k_{11} - w_1) & \cdots & M_1 \\
  \ast & \alpha^2(k_{11} - w_1) + w_2 & \cdots & M_2 \\
  \vdots & \vdots & \ddots & \vdots \\
  \ast & \ast & \cdots & M_3
\end{bmatrix},
\] (31)

where

\[
M_1 = \alpha(k_{11} - (n-1)w_1),
\]
\[
M_2 = \alpha^2(k_{11} - (n-1)w_1),
\]
\[
M_3 = \alpha^2[k_{11} - (n-1)w_1] + (n-1)w_2,
\] (32)

and \(k_{11}\) is determined by

\[
\alpha^2k_{11}^2 - k_{11}w_1\alpha^2(2n-1) + w_1(n-1)(\alpha^2w_1n - w_2) - w_1 = 0.
\] (33)

Equation (33) has two roots,

\[
k'_{11} = \frac{1}{2\alpha} \left[ w_1\alpha(2n-1) \right.
\]
\[
+ \sqrt{w_1^2\alpha^2 + 4w_1w_2(n-1) + 4w_1} \right],
\]
\[
k''_{11} = \frac{1}{2\alpha} \left[ w_1\alpha(2n-1) \right.
\]
\[
- \sqrt{w_1^2\alpha^2 + 4w_1w_2(n-1) + 4w_1} \right].
\] (34)

Only \(k'_{11}\) guarantees that \(K\) is semipositive definite. Now we can obtain \(g\) by substituting \(k'_{11}\) into (15),

\[
g = \gamma_2 \left[ \frac{1}{\alpha} \frac{1}{1} \cdots \frac{1}{1} \right],
\] (35)

where

\[
\gamma_2 = \frac{\alpha^2w_1 + 4w_1w_2(n-1) - \alpha^2w_1}{2 + 2w_2(n-1)}.
\] (36)

Now we can derive \(r\) from (16):

\[
r = \frac{\gamma_2f_{\text{opt}}}{\alpha}.
\] (37)

We conclude that the optimal control law for the second quality criterion is given by

\[
u_{\text{opt}}(kT) = \gamma_2 \left[ \frac{y_d - y(kT)}{\alpha} - \sum_{i=2}^{n} x_i(kT) \right].
\] (38)

Remark 1. As \(w \to 0\) for the first criterion, and as \(w_1 \to 0\) for the second one, the influence of the output error on the value of the functional diminishes, and gains of the controllers decrease to zero. As \(w \to \infty\) for the first criterion, and as \(w_1 \to \infty\) for any finite \(w_2\) in the second case, the impact of the control signal is negligible, and the output error is to be reduced to zero as quickly as possible, regardless of the value of the control signal. The controllers then become dead-beat schemes, and their gains approach unity. State variables \(x_2, x_3, \ldots, x_n\) are the delayed values of the control signal. This means that, as \(w_2 \to \infty\) with finite \(w_1\) for the second cost functional, the value of the control signal dominates the quality criterion and gain \(\gamma_2\) drops to zero. The relation between \(\gamma_1\) and \(w\) for \(\alpha = 0.97\) is shown in Fig. 2. Relations between \(\gamma_2\) and the weighting coefficients \(w_1\) and \(w_2\) for the second controller are depicted in Figs. 3 and 4 respectively.
A discrete-time closed-loop system is asymptotically stable if all the roots of the characteristic polynomial of its system state matrix are located inside the unit circle. The roots of the characteristic polynomial of its discrete-time closed-loop system are described by (39) if \( \gamma \in (0, 2) \). Since both \( \gamma_1 \) and \( \gamma_2 \) satisfy \( \gamma \in (0, 1) \), both of the proposed optimal controllers indeed guarantee the asymptotic stability of the system.

3.3. Properties of the proposed strategy. Properties of both optimal controllers will be formulated and proved simultaneously, again denoting by \( \gamma \) both \( \gamma_1 \) and \( \gamma_2 \).

Theorem 1. If the proposed control strategy is applied, then the queue length never exceeds its demand value \( y_d \).

Proof. Substituting (3) and (5) into (27) or (38), we obtain

\[
u_{\text{opt}}(kT) = \gamma \left[ \frac{y_d}{\alpha} - \frac{k-1}{\alpha} \sum_{i=0}^{k-1} u(iT) + \frac{1}{\alpha} \sum_{i=0}^{k-1} h(iT) \right].
\] (40)

We assume that \( y(mT) \leq y_d \) at some time instant \( m \geq 0 \). We will prove that this theorem is also true for \( m + 1 \). The queue length at time \( m + 1 \) can be expressed as

\[
y[(m + 1)T] = y(mT) + \alpha u[(m - m_{\text{RTT}})T] - h(mT).
\] (41)

Using (3) and (40), we obtain

\[
y[(m + 1)T] = y(mT) - h(mT) + \alpha \left[ \frac{y_d}{\alpha} - \frac{1}{\alpha} \sum_{i=0}^{m-m_{\text{RTT}}-1} u(iT) + \gamma \sum_{i=m-m_{\text{RTT}}}^{m-1} h(iT) \right]
\]

\[
= \gamma y_d + y(mT) - \gamma \sum_{i=m-m_{\text{RTT}}}^{m-1} h(iT) - h(mT)
\]

\[
- \gamma \left[ \sum_{i=m-m_{\text{RTT}}}^{m-1} u(iT) - \sum_{i=0}^{m-1} h(iT) \right]
\]

\[
= y_d - (1 - \gamma)[y_d - y(mT)]
\]

\[
- \gamma \sum_{i=m-m_{\text{RTT}}}^{m-1} h(iT) - h(mT).
\] (42)

Since \( \gamma \in (0, 1) \) and \( h(kT) \) is always nonnegative, \( y[(m + 1)T] \leq y_d \). Because \( y(0) \leq y_d \), we conclude that \( y(kT) \leq y_d \) for any \( k \geq 0 \). This ends the proof.

We notice from (3) that, if \( x_1[(k + 1)T] > 0 \), then the available bandwidth \( d(kT) \) is fully used. The next theorem shows how to choose \( y_d \) in order to ensure that the queue length is strictly positive.

Theorem 2. If the desired queue length

\[
y_d > d_{\text{max}}(m_{\text{RTT}} + 1/\gamma),
\] (43)

then for any \( k \geq m_{\text{RTT}} + 1 \) the queue length is greater than zero.

Proof. From (3) it follows that \( y(kT) = 0 \) for \( k < m_{\text{RTT}} + 1 \). Furthermore, we notice that both (27) and (38) lead to

\[
u[(m + 1)T] = \alpha u(0) - h(m_{\text{RTT}}T)
\]

\[
> d_{\text{max}}(\gamma m_{\text{RTT}} + 1) - d_{\text{max}} > 0.
\] (44)

Now we shall demonstrate that the condition \( y(mT) > 0 \) implies \( y[(m + 1)T] > 0 \). From (42) and (43), we obtain

\[
y[(m + 1)T] = \gamma y_d + (1 - \gamma)y(mT)
\]

\[
- \gamma \sum_{i=m-m_{\text{RTT}}}^{m-1} h(iT) - h(mT)
\]

\[
\geq \gamma y_d - \gamma m_{\text{RTT}} d_{\text{max}} - d_{\text{max}}
\]

\[
= \gamma \left[ y_d - d_{\text{max}} \left( m_{\text{RTT}} + \frac{1}{\gamma} \right) \right] > 0.
\] (45)

It follows from (44) and (45) that, indeed, if (43) holds, then \( y(kT) > 0 \) for \( k \geq m_{\text{RTT}} + 1 \). This completes the induction proof.
In order to be applied in a real network, any flow control algorithm should generate transmission rates that are always nonnegative and limited by some predictable, finite value. This property is demonstrated in the next theorem.

**Theorem 3.** With the application of the proposed controller, data transmission rates are always nonnegative and upper bounded, i.e., for any \( k \geq 0 \),

\[
0 \leq u(kT) \leq \max(\gamma y_d / \alpha, d_{\text{max}} / \alpha)
\]

(46)

Proof. Let us assume that (46) holds for some \( m \geq 0 \). We shall prove that the proposition is true also for \( m + 1 \). Using (40), we get

\[
u[(m+1)T] = \gamma \left( y_d / \alpha - \sum_{i=0}^{m} u(iT) + \frac{1}{\alpha} \sum_{i=0}^{m} h(iT) \right) \\
= \frac{\gamma y_d}{\alpha} - \sum_{i=0}^{m-1} u(iT) + \frac{1}{\alpha} \sum_{i=0}^{m-1} h(iT) \\
- \gamma \left( u(mT) - \frac{1}{\alpha} h(mT) \right) \\
= (1 - \gamma)u(mT) + \frac{\gamma}{\alpha} h(mT). \tag{47}
\]

From (40) we have \( u(0) = \gamma y_d / \alpha \). We conclude that (46) indeed holds for any \( k \geq 0 \).

Let us finally notice that, if the source is not persistent or if the round trip time is not known exactly, then a similar approach as the one proposed by Pietrabissa et al. (2006) can be adopted. In that paper an adaptive controller combining the advantages of control-theoretic and fuzzy-logic approaches was proposed to address the issue of source saturation as well as uncertain, possibly time-varying transmission delays. Furthermore, if the round trip time is not known exactly, and in particular when it is not a multiple of the discretisation period, then our optimal controller can be equipped with a saturating element and a similar approach to the one proposed by Bartoszewicz (2006) can also be effectively applied. However, when the round trip time is not known exactly, then conditions stated in Theorems 1 and 2 become more restrictive.

### 4. Simulation results

To verify the properties of the proposed control strategy, computer simulations have been performed. The sampling time \( T \) was selected as 1 ms. The round trip time \( RTT \) in the virtual circuit was assumed to be 8 ms. From this follows \( m_{\text{RTT}} = 8 \) and \( n = 9 \). The maximum available bandwidth \( d_{\text{max}} = 6 \text{ kb} \). The bandwidth actually available for the data transfer is shown in Fig. 5. Sudden changes of large amplitude occur in the function \( d \), which reflects the most difficult possible conditions in the network. It is assumed that 3% of data is lost during the transmission, which corresponds to \( \alpha = 0.97 \).

The derived parameters for the first controller: the gain \( \gamma_1 \) obtained from (25), the minimum demand queue length \( y_d \) calculated from the condition (43), and the queue length actually used in the simulation \( y_d \) are shown in Table 1. The results of the simulations are shown in Figs. 6 and 7. The value of the control signal at the beginning of the transmission process is shown in Fig. 8.

![Fig. 5. Available bandwidth.](image)

![Table 1. Parameters of the first controller.](image)

<table>
<thead>
<tr>
<th>( w )</th>
<th>( \gamma_1 )</th>
<th>( y_d ) [kb]</th>
<th>( y_d ) [kb]</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>0.5</td>
<td>10</td>
<td>0.5</td>
</tr>
<tr>
<td>10</td>
<td>0.5</td>
<td>54.6</td>
<td>56</td>
</tr>
<tr>
<td>0.5</td>
<td>0.4899</td>
<td>60.3</td>
<td>62</td>
</tr>
</tbody>
</table>

![Fig. 6. Buffer occupancy with the first controller for different values of \( w \).](image)

Similar tests were performed with the second controller. The weighting factor \( w_2 \) was selected as 5 and the resulting parameters of the second controller are shown in Table 2. The gain \( \gamma_2 \) was derived from (25), and \( y_d \) from (43). Simulation results are shown in Figs. 9 and 10. Again, the value of the control signal at the beginning of the transmission process is shown more clearly in Fig. 11.

It can be seen from the figures that the transmission rates calculated by both the algorithms are always nonnegative and upper bounded. Furthermore, the queue length
never exceeds its demand value, and for $k \geq m_{RTT} + 1$ it never decreases to zero. This means that there is no risk of data loss resulting from buffer overflow, and that all of the available bandwidth is used. Consequently, maximum throughput possible in the network is achieved.

<table>
<thead>
<tr>
<th>$w_1$</th>
<th>$\gamma_1$</th>
<th>$y_d$ [kb]</th>
<th>$y'_d$ [kb]</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>0.3779</td>
<td>63.9</td>
<td>66</td>
</tr>
<tr>
<td>0.5</td>
<td>0.1015</td>
<td>107.1</td>
<td>110</td>
</tr>
</tbody>
</table>

Both algorithms can be adjusted to specific requirements using the appropriate weighting factors. As could be expected, changes in $w$ of the first criterion have a similar impact as changes of $w_1$ for the second case. Larger values of $w$ and $w_1$ result in faster tracking of the output flow. This, in turn, allows allocating smaller buffers, while still utilizing the full available bandwidth. On the other hand, smaller $w$ and $w_1$ lead to smaller values of the control signal at the beginning of the transmission process. This also makes the control signal smoother, which is advantageous for transmission consistency. Changes in $w_2$ for the second case have the opposite effect, as has already been stated in the previous section.

In all the above simulations the main goal was to fully utilize the available bandwidth. Therefore, the demand queue length was selected according to Theorem 2. However, the condition (43) may not be satisfied for two reasons. Firstly, there may be not enough physical memory in the congested node. Secondly, we can deliberately lower $y_d$ in order to reduce jitter and latency, and in this way improve the Quality of Service (QoS) in the network. Therefore, for the last simulation scenario we consider the case of $w_1 = 10$, $w_2 = 5$ for the second functional, but now we choose $y_d = 58 < y'_d$. The results are presented in Figs. [12] and [13]. The unused bandwidth, which is equal to $d(kT) - h(kT)$, is presented in Fig. [14]. As can be ob-
served, the proposed control algorithm ensures fairly good bandwidth utilization even with a lowered demand queue length.

![Buffer occupancy for lowered y_d.](image1)

![Transmission rate for lowered y_d.](image2)

![Unused bandwidth for lowered y_d.](image3)

5. Conclusion

In this paper, two LQ optimal sliding-mode flow controllers for a single virtual circuit in data transmission networks were proposed. Possible data losses during the transmission were explicitly taken into account. The design procedure was based on minimization of two different quadratic cost functionals, and solving the resulting matrix Ricatti equation. The closed-loop system stability was demonstrated. The condition for full bandwidth consumption was formulated and proved. It was also proved that the rates generated by the controller are always non-negative and upper bounded. Finally, it is worth pointing out that the results presented by Ignaciuk and Bartoszewicz (2008) may be regarded as a special case of the more general analysis performed in this paper.

Acknowledgment

This work was performed in the framework of the project Optimal sliding mode control of time delay systems financed by the National Science Centre (DEC 2011/01/B/ST7/02582). The kind support provided by the Foundation for Polish Science under a Mistrz grant is also acknowledged.

This work is an extended version of the paper Linear quadratic optimal sliding-mode controller for connection-oriented communication networks with lossy links presented at the 11th International Conference on Diagnostics of Processes and Systems (Łagów Lubuski, Poland, 2013).

References


Pietrabissa, A., Priscoli, F. D., Fiaschetti, A. and Paolo, F. D.


Kwakernaak, H. and Sivan, R. (1972). Linear Optimal Control


A. Bartoszewicz and P. Leśniewski


Andrzej Bartoszewicz received the M.Sc. degree in 1987 and the Ph.D. degree in 1993, both from the Łódź University of Technology, Poland. Then he obtained the postdoctoral degree in control engineering and robotics from the Academy of Mining and Metallurgy in Cracow, Poland. He was a visiting scholar at Purdue University, West Lafayette, USA, and at Strathclyde University, Glasgow, UK. Then for one year he was at the University of Leicester, UK. Currently he is a professor at the Technical University of Łódź, a vice dean of the Faculty of Electrical, Electronic, Computer and Control Engineering, the head of the Electric Drive and Industrial Automation Group and a vice-director of the Institute of Automatic Control. He has published three monographs and over 250 papers, primarily in the fields of sliding mode control and congestion control in data transmission networks.
Piotr Leśniowski received the Master’s degree in control engineering and robotics from the Łódź University of Technology in 2012. Currently he is a Ph.D. student at that university. His main research interests include discrete time sliding mode control and time-delay systems.

Received: 17 April 2013
Revised: 4 August 2013