BULLETIN OF THE POLISH ACADEMY OF SCIENCES
TECHNICAL SCIENCES, Vol. 60, No. 2, 2012
DOI: 10.2478/v10175-012-0039-0

DEDICATED PAPERS

Quantity-based flow control strategy for connection-oriented communication networks

M. KARBOWAŃCZYK[1]* and J. ŻUK[2]

[1] Institute of Information Technology, Technical University of Łódź, 215 Wólczańska St., 90-924 Łódź, Poland
[2] Institute of Automatic Control, Technical University of Łódź, 18/22 B.Stefanowskiego St., 90-924 Łódź, Poland

Abstract. In this paper a new flow control strategy for connection-oriented communication networks is presented. It utilises methods of control theory, in particular the Smith predictor and dead-beat control, to achieve desirable dynamics of the considered network. In contrast to a number of earlier proposals in which the controller command is interpreted as the rate transmission, in our solution it is interpreted as the quantity of data that the controlled node is expected to send. This allows us to model a single virtual connection with non-persistent data source as a time-delay system in which the delay may temporarily exceed its assumed boundary. Favourable properties of the proposed control strategy are formulated as mathematical theorems and thoroughly discussed.

Key words: telecommunication networks, congestion control, Smith predictor.

1. Introduction

Various forms of digital data transfer play more and more important role in the global economy as well as our day-to-day life. Business transactions, video streaming, worldwide publishing, WEB 2.0 – all these features of modern Internet causes the rapidly growing demand for high data transfer rates. To fulfil this demand, new physical layer technologies are introduced so that the available throughput of network links is increased. As the data transmission through a network link is subject to the signal propagation delay (which is a physical property of the link), modern communication networks are characterised by large bandwidth-delay product. Moreover, multiple data flows, that impose different quality of service (QoS) requirements (such as minimum delay, maximum throughput or transfer reliability) are passing through the network links in parallel, sharing the bandwidth. Taking this into account we can clearly see that the throughput available for a specific flow may vary with time in an unpredictable way. Thus, to achieve satisfactory throughput utilisation and transfer reliability, a suitable flow control strategy must be applied to such networks.

The problem of data flow control in fast communication networks has been subject of a number of research efforts. A good survey on earlier data flow control schemes is presented in [1]. Afterwards, various control strategies, employing artificial intelligence [2–4], fuzzy logic [5], game theory [6, 7] and other approaches, were proposed. Furthermore, it is worth notifying that the communication network can be modelled as a time-delay control system. This allowed many researchers to apply numerous methods of control theory (such as classic PD [8, 9], PID [10], sliding mode [11–15], stochastic [16] or adaptive [17] controllers) to the problem of data flow control in these networks. As the propagation delays may be significant, the use of the Smith predictor combined with proportional, dead-beat, on-off and other types of controllers has been the subject of many research projects [18–26].

In this paper we propose a flow control strategy for fast, connection-oriented communication networks. The strategy combines the benefits of the Smith predictor and dead-beat control. On the contrary to the most earlier solutions, the strategy described here assumes that the data source in virtual connection interprets the control value (received from the controller placed at the node) as the quantity of data that it is expected to send, instead of the rate at which it should transfer the data [11–15]. The change of the control value interpretation is motivated by the fact that the idea of rate-based control is noticeably inconsistent with the way the real communication networks work. First, it cannot be assumed that the source is able to send data at the rate that is exactly equal to the value established by the controller. In fact, the source always sends data at the rate determined by the physical layer that supports the transmission. Therefore, the term "rate control" can be referred only to an average rate value calculated within some time window. Taking into account the granularity of data units that are transmitted in the network we can obviously state that the data transmission rate can be controlled only with some finite accuracy. Moreover, we cannot guarantee that the propagation delay between the controlling node and the source is constant. Even if we assume that the overall delay is variable but bounded, the upper bound may be exceeded, i.e. due to the congestion of some of the intermediate nodes. Combining both issues mentioned above one should easily conclude that it is impossible to guarantee node congestion avoidance if we apply rate-based flow control scheme to a real communication network.

The remainder of the paper is organised in the following way. The model of the considered network is described in

*e-mail: michal.karbowanczyk@p.lodz.pl

the second section. Then the design of the proposed control strategy is presented, along with theorems and discussions of its important properties, such as congestion avoidance and effective throughput utilisation. Finally, the fourth section concludes the paper.

2. Network model

We consider a packet-switching telecommunication network, consisting of data sources, intermediate nodes and destinations. The data stream partitioned into packets is transmitted from a source to its destination through a number of intermediate nodes. Every intermediate node works in the store&forward mode, that is each packet received by the node is stored in a buffer where it waits until it is sent to another node. As the capacity of the buffer is limited, excessive inflow rate may cause that the received packet need to be dropped. In such a situation we state that the node is congested. An occurrence of the node congestion and packet loss usually impose several negative consequences on the network performance, in particular throughput degradation.

We assume that the network is of connection-oriented type, which means that before any data is transmitted, a virtual connection between the source and the destination has to be established. The virtual connection defines a set of nodes that transmit packets from the source to the destination, called the path. The path for an established virtual connection remains unchanged until the connection is closed.

We take into account a set of J (where $J > 1$) virtual connections that pass through a specified intermediate node. We assume that this node is the bottleneck node for every virtual connection belonging to set J, that is, the available throughput at the node limits the throughput of each connection. Thus the state of the other intermediate nodes is negligible from the perspective of data flow control. This allows us to reduce the considered virtual connections to the configuration consisting of their sources, destinations and the single bottleneck node, as it is shown in Fig. 1. Similar approach is widely used in research works [18–20, 22–26].

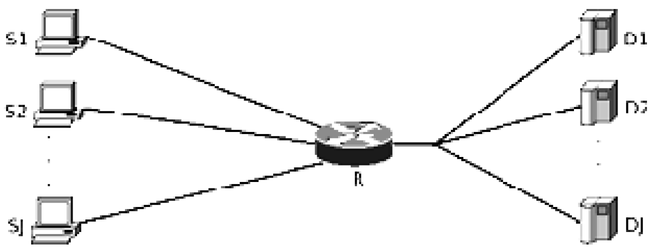

Fig. 1. Model of multi source connection-oriented communication network with bottleneck node R

The queue of packets stored in memory buffer of the bottleneck node can be then modelled as dynamic system depicted in Fig. 2.

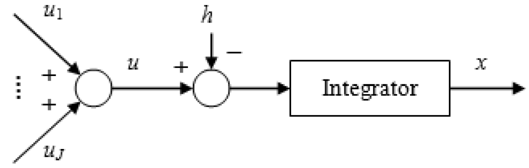

Fig. 2. Model of packets queue in the bottleneck node buffer

Let $j = 1, 2, ..., J$ and t, $t \geq 0$, denote time. The symbol $u_j(t)$ represents the rate at which the data sent by j-th source is being received by the node. Thus, the overall rate at which data reaches the buffer, denoted as $u(t)$, can be calculated as

$$\underset{t\geq 0}{\forall} \; u(t) = \sum_{j=1}^{J} u_j(t). \tag{1}$$

The throughput available for the connections is denoted as $d(t)$. This function cannot be determined a priori, although it is known that its values are nonnegative and bounded by positive constant $d_{\max}$. Function $h(t)$ represents the rate at which the data stored in the buffer is sent by the node. This value is also nonnegative and not greater than available throughput, so we have

$$\underset{t\geq 0}{\forall} \; 0 \leq h(t) \leq d(t) \leq d_{\max}. \tag{2}$$

Then, assuming that the buffer is initially empty, the queue length, denoted by $x(t)$, can be calculated as

$$\underset{t\geq 0}{\forall} \; x(t) = \int_0^t u(\tau)d\tau - \int_0^t h(\tau)d\tau. \tag{3}$$

We expect that the considered network is able to provide explicit feedback to the sources. It is accomplished by using special units called control units, which are sent periodically by the source to the destination and immediately sent back to the source. Every intermediate node is allowed to put the control value into control unit, and the source utilises obtained control value to adjust its operation to the state of the network. A good example of such a solution is ATM network with ABR service category [27]. We assume that the control units are numbered starting from 0, and for every virtual connection it is assured that k-th control unit ($k = 0, 1, 2, \ldots$) arrives at a bottleneck node at time instant $t(k) = kT$, where T denotes the discretisation period. Thus control values calculated by the node form a sequence denoted by $\{a(k)\}$. Further in the paper we use notation $x(k)$, $u(k)$ and for other functions instead of $x(kT)$, $u(kT)$ etc. This applies to every time function except $h(k)$ defined below.

For any positive integer k we also define $h(k)$ (i.e. the amount of data sent by bottleneck node) within time period $[(k-1)T; kT]$

$$\underset{k>0}{\forall} \; h(t) = \int_{(k-1)T}^{kT} h(\tau)d\tau. \tag{4}$$

3. Proposed control strategy

We propose a control strategy that combines the Smith predictor and a dead-beat control. For every nonnegative k the control value is obtained from the following formula

$$\underset{k\geq 0}{\forall}\ a(k) = X^D - x(k) - w(k), \tag{5}$$

where positive reference value X^D is a parameter of the control strategy and $w(k)$ denotes the amount of 'in flight' data. Such a data exists due to the delay in the considered system incurred by the propagation latency on the network links. The k-th control packet sent by the bottleneck node is received by the source of the j-th virtual connection with backward propagation delay denoted as $T_{B:j}$, at the moment $t_{R:j}(k)$. Let $w_{B:j}(t)$ represent the sum of the control values carried by the control packets not yet received by the source. Furthermore, additional delay may be introduced by the source, if it does not have enough data to be sent according to the control value. $t_{S:j}(k)$ denotes the moment when the source of the j-th virtual connection starts sending data packets according to the control value obtained from the k-th control packet. Let $w_{Q:j}(t)$ denote the sum of the control values queued by the source. The data packets sent by the source of the j-th virtual connection are received by the node with forward propagation delay denoted as $T_{F:j}$, at the moment $t_{N:j}(k)$. Let $w_{F:j}(t)$ represent the quantity of data carried by the data packets not yet received by the node. The sum of forward and backward propagation delays is called round-trip time and represented by the symbol *RTTC*$_j$. Consequently, there is always a number of control values that have been previously calculated by control algorithm, but did not yet affect the state of the bottleneck node, and this number can be obtained from the following formula

$$\underset{t\geq 0}{\forall}\ w(t) = \sum_{j=1}^{J} [w_{B:j}(t) + w_{Q:j}(t) + w_{F:j}(t)]. \tag{6}$$

We assume that the time delays are constant, thus the following statements are valid

$$\begin{aligned} \underset{k\geq 0}{\forall}\ & t_{R:j}(k) - t(k) = T_{B:j}, \\ & t_{N:j}(k) - t_{S:j}(k) = T_{F:j}. \end{aligned} \tag{7}$$

We can also distinguish the moment when the first data unit reaches the node

$$t_N(0) = \min\{\, t_{N:j}(0), \quad j = 1, 2, \ldots, J\}. \tag{8}$$

The control value, calculated according to (5), is distributed equally among considered virtual connections. Upon receipt of the control packet the source is obliged to send the amount of data equal to the obtained control value.

Consider the j-th virtual connection and let $t \geq t_{N:j}(0)$. The node receives the data that was sent by the source as the realisation of some control value. Let $k_{d:j}(t)$ denote the number of this control value. Obviously $k_{d:j}(t) < [[t/T]]$, where $[[\cdot]]$ denotes the integer part of a number. Furthermore, notice that the control value $a_j(k_{d:j}(t))$ can be divided into two factors: $a_{I:j}(k_{d:j}(t),\ t)$ – amount of data that already reached the node and $a_{O:j}(k_{d:j}(t),\ t)$ – amount of data that is still 'in flight'

$$\begin{aligned} \underset{t\geq t_{N:j}(0)}{\forall}\ & a_j(k_{d:j}(t)) = a_{I:j}(k_{d:j}(t), t) \\ & + a_{O:j}(k_{d:j}(t), t). \end{aligned} \tag{9}$$

On the other hand we define $k_{0:j}$ as the number of the first control value calculated after the first data packets in j-th virtual connection reach the node

$$\begin{aligned} & \underset{j=1,2,\ldots,J}{\forall}\ k_{0:j} \\ & = \min\{\, k = 0, 1, 2 \ldots :\ t_N(0) \leq t(k)\}, \end{aligned} \tag{10}$$

$$k_0 = \min\{\, k_{0:j} = 0, 1, 2 \ldots, J\}. \tag{11}$$

Note also that since before setting up the connections (i.e. for $t < 0$) we assume $u(t) = 0$ and as a consequence there are no data units in the buffer, i.e. $x(k)|_{k<0} = 0$, and even no 'in flight' data, we get from (5) $a(0) = X^D$. Moreover, it can be easily stated that

$$\underset{t\leq t_N(0)}{\forall}\ x(t) = h(t) = 0 \quad \underset{k<k_0}{\forall}\ x(k) = h(k) = 0. \tag{12}$$

Consider again $t \geq 0$. Since time delays may vary among virtual connections, the following three sets can be defined

$$\begin{aligned} A(t) &= \{\, j = 1, 2, \ldots, J :\ t < t_{N:j}(0)\}, \\ B(t) &= \{\, j = 1, 2, \ldots, J :\ t_{N:j}(0) \leq t \leq t_{N:j}(1)\}, \\ C(t) &= \{\, j = 1, 2, \ldots, J :\ t > t_{N:j}(1)\}. \end{aligned} \tag{13}$$

Taking into account notations introduced above we can calculate the length of the queue and amount of 'in flight' data from the formulas

$$\begin{aligned} \underset{t\geq 0}{\forall}\ x(t) &= \sum_{j\in B(t)} a_{I:j}(0,t) \\ &+ \sum_{j\in C(t)} \left[a_j(0) + \sum_{i=1}^{k_{d:j}(t)-1} a_j(i) + a_{I:j}(k_{d:j}(t),t) \right] \\ &- \sum_{i=1}^{[[t/T]]} h(i) - \int_{[[t/T]]T}^{t} h(\tau)d\tau, \end{aligned} \tag{14}$$

$$\begin{aligned} \underset{k\geq 0}{\forall}\ x(k) &= \sum_{j\in B(k)} a_{I:j}(0,t(k)) + \sum_{j\in C(k)} \\ &\cdot \left[a_j(0) + \sum_{i=1}^{k_{d:j}(k)-1} a_j(i) + a_{I:j}(k_{d:j}(k),t(k)) \right] - \sum_{i=1}^{k} h(i), \end{aligned} \tag{15}$$

$$\begin{aligned} \underset{k>0}{\forall}\ w(k) &= \sum_{j\in A(k)} \left[a_j(0) + \sum_{i=1}^{k-1} a_j(i) \right] \\ &+ \sum_{j\in B(k)} \left[a_{O:j}(0,t(k)) + \sum_{i=1}^{k-1} a_j(i) \right] \\ &+ \sum_{j\in C(k)} \left[a_{O:j}(k_{d:j}(k),t(k)) + \sum_{i=k_{d:j}(k)+1}^{k-1} a_j(i) \right]. \end{aligned} \tag{16}$$

The control strategy defined by (5) is characterised by a number of properties, which is presented further in this section in the form of theorems and remarks.

Theorem 1. For any positive integer k the control value is equal to the amount of data sent by bottleneck node within time period $((k-1)T;\ kT]$, that is

$$\underset{k>0}{\forall}\ a(k) = h(k). \tag{17}$$

Proof. We apply the principle of mathematical induction. First, in section (i) we show that Theorem 1 holds for $k = 1$. Then in section (ii) we consider arbitrary $k > 1$ and show that if only $a(i) = h(i)$ for any $1 \le i < k$, then such an equality is also valid for $i = k$.

(i.a) Consider $k_0 > 1$. From the fact that $a(0) = X^D$ we conclude that $w(1) = X^D$, thus from (5) we obtain $a(1) = -x(1)$. Since $k_0 > 1$ we know from (12) that $x(1) = h(1) = 0$, so we have $a(1) = 0 = h(1)$.

(i.b) On the other hand, if $k_0 = 1$, then the set $C(k_0)$ is empty. In consequence from (15) and (16) we get respectively

$$x(1) = \sum_{j\in B(1)} a_{I:j}(0, t(1)) - h(1), \tag{18}$$

$$w(1) = \sum_{j\in A(1)} a_j(0) + \sum_{j\in B(1)} a_{O:j}(0, t(1)). \tag{19}$$

Taking into account above relations and the equality $a(0) = X^D$ we notice that

$$\begin{aligned} a(1) &= X^D - x(1) - w(1) \\ &= X^D - \sum_{j\in B(1)} a_{I:j}(0, t(1)) + h(1) \\ &\quad - \sum_{j\in A(1)} a_j(0) - \sum_{j\in B(1)} a_{O:j}(0, t(1)) \\ &= X^D - \sum_{j\in B(1)} [a_{I:j}(0, t(1)) + a_{O:j}(0, t(1))] + h(1) \\ &\quad - \sum_{j\in A(1)} a_j(0) = X^D - \sum_{j\in B(1)} a_j(0) + h(1) \\ &\quad - \sum_{j\in A(1)} a_j(0) = X^D - \sum_{j=1}^{J} a_j(0) + h(1) \\ &= X^D - a(0) + h(1) = X^D - X^D + h(1) = h(1). \end{aligned} \tag{20}$$

(ii) Now let $k > 1$. Suppose that for any $i = 1, 2, ..., k-1$ we have $a(i) = h(i)$.

(ii.a) Assume that $k < k_0$. Since $a(0) = X^D$ and $x(1) = ... = x(k) = h(1) = ... = h(k) = 0$, we easily notice that $a(1) = ... = a(k-1) = 0$, thus $w(k) = X^D$. Taking into account above considerations, from (5) we obtain

$$\begin{aligned} a(k) &= X^D - x(k) - w(k) \\ &= X^D - 0 - X^D = 0 = h(k). \end{aligned} \tag{21}$$

(ii.b) Consider now $k \ge k_0$. Applying (15) and (16) to (5) we get

$$\begin{aligned} a(k) &= X^D - x(k) - w(k) \\ &= X^D - \sum_{j\in B(k)} a_{I:j}(0, t(k)) \\ &- \sum_{j\in C(k)} \left[a_j(0) + \sum_{i=1}^{k_{d:j}(k)-1} a_j(i) + a_{I:j}(k_{d:j}(k), t(k)) \right] + \sum_{i=1}^{k} h(i) \\ &- \sum_{j\in A(k)} \left[a_j(0) + \sum_{i=1}^{k-1} a_j(i) \right] \\ &- \sum_{j\in B(k)} \left[a_{O:j}(0, t(k)) + \sum_{i=1}^{k-1} a_j(i) \right] \\ &- \sum_{j\in C(k)} \left[a_{O:j}(k_{d:j}(k), t(k)) + \sum_{i=k_{d:j}(k)+1}^{k-1} a_j(i) \right] \end{aligned} \tag{22}$$

what can be rearranged as follows

$$\begin{aligned} a(k) &= X^D - \sum_{j\in B(k)} a_{I:j}(0, t(k)) \\ &- \sum_{j\in C(k)} a_j(0) - \sum_{j\in C(k)} \sum_{i=1}^{k_{d:j}(k)-1} a_j(i) \\ &- \sum_{j\in C(k)} a_{I:j}(k_{d:j}(k), t(k)) + \sum_{i=1}^{k} h(i) \\ &- \sum_{j\in A(k)} a_j(0) - \sum_{j\in A(k)} \sum_{i=1}^{k-1} a_j(i) \\ &- \sum_{j\in B(k)} a_{O:j}(0, t(k)) - \sum_{j\in B(k)} \sum_{i=1}^{k-1} a_j(i) \\ &- \sum_{j\in C(k)} a_{O:j}(k_{d:j}(k), t(k)) \\ &- \sum_{j\in C(k)} \sum_{i=k_{d:j}(k)+1}^{k-1} a_j(i) = X^D - \sum_{j\in A(k)} a_j(0) \\ &- \sum_{j\in C(k)} a_j(0) - \sum_{j\in B(k)} a_{I:j}(0, t(k)) \\ &- \sum_{j\in B(k)} a_{O:j}(0, t(k)) - \sum_{j\in C(k)} a_{I:j}(k_{d:j}(k), t(k)) \\ &- \sum_{j\in C(k)} a_{O:j}(k_{d:j}(k), t(k)) - \sum_{j\in A(k)} \sum_{i=1}^{k-1} a_j(i) \\ &- \sum_{j\in B(k)} \sum_{i=1}^{k-1} a_j(i) - \sum_{j\in C(k)} \sum_{i=1}^{k_{d:j}(k)-1} a_j(i) \\ &- \sum_{j\in C(k)} \sum_{i=k_{d:j}(k)+1}^{k-1} a_j(i) + \sum_{i=1}^{k} h(i). \end{aligned} \tag{23}$$

Simplifying statement (23), using definitions (13), we obtain

$$
\begin{aligned}
a(k) &= X^D + \sum_{i=1}^{k} h(i) - \sum_{j\in A(k)} a_j(0) \\
&- \sum_{j\in C(k)} a_j(0) - \sum_{j\in B(k)} a_j(0) - \sum_{j\in A(k)} \sum_{i=1}^{k-1} a_j(i) \\
&- \sum_{j\in B(k)} \sum_{i=1}^{k-1} a_j(i) - \sum_{j\in C(k)} \sum_{i=1}^{k_{d:j}(k)-1} a_j(i) \\
&- \sum_{j\in C(k)} a_j(k_{d:j}(k)) - \sum_{j\in C(k)} \sum_{i=k_{d:j}(k)+1}^{k-1} a_j(i) \qquad (24)\\
&= X^D + \sum_{i=1}^{k} h(i) - a(0) - \sum_{j\in A(k)} \sum_{i=1}^{k-1} a_j(i) \\
&- \sum_{j\in B(k)} \sum_{i=1}^{k-1} a_j(i) - \sum_{j\in C(k)} \sum_{i=1}^{k-1} a_j(i) \\
&= X^D + \sum_{i=1}^{k-1} h(i) + h(k) - a(0) - \sum_{i=1}^{k-1} a(i).
\end{aligned}
$$

Finally applying $a(0) = X^D$ and the assumption that $a(i) = h(i)$ holds for any $i = 1, 2, ..., k-1$, we get

$$
a(k) = X^D + \sum_{i=1}^{k-1} h(i) + h(k) - X^D - \sum_{i=1}^{k-1} h(i) = h(k). \qquad (25)
$$

As k was chosen arbitrarily, this concludes the proof of Theorem 1.

Remark 1.1. The control values calculated by proposed control strategy are nonnegative and bounded. It is an obvious consequence of Theorem 1. Using definition (4) and assumption (2) set the following

$$
\begin{aligned}
\underset{k>0}{\forall} \quad 0 \le a(k) = h(k) &= \int_{(k-1)T}^{kT} h(\tau) d\tau \\
&\le \int_{(k-1)T}^{kT} d_{\max} d\tau = d_{\max} T.
\end{aligned} \qquad (26)
$$

This remark is very important from the practical perspective. Considering a real telecommunication network it is clearly seen that it is not possible for the data source to execute negative control values. On the other hand, execution of extremely large control value is theoretically possible, but it would impose proportionally large delays due to the limited bandwidth of the network links. In other words, if Remark 1.1 could not be stated, the proposed control strategy would have no practical significance.

Theorem 2. The packet queue length in the bottleneck node buffer never exceeds the reference value X^D.

$$
\underset{t\ge 0}{\forall} \quad x(t) \le X^D. \qquad (27)
$$

Proof. Notice that for $t \le t_N(0)$ relation (27) is obviously valid by virtue of equality (12).

Let $t > t_N(0)$. From relation (14), taking into account Theorem 1, we obtain

$$
\begin{aligned}
x(t) &= \sum_{j\in B(t)} a_{I:j}(0,t) \\
&+ \sum_{j\in C(t)} \left[a_j(0) + \sum_{i=1}^{k_{d:j}(t)-1} a_j(i) + a_{I:j}(k_{d:j}(t), t) \right] \\
&- \sum_{i=1}^{[[t/T]]} h(i) - \int_{[[t/T]]T}^{t} h(\tau) d\tau \le \sum_{j\in B(t)} a_j(0) + \sum_{j\in C(t)} a_j(0) \\
&+ \sum_{j\in C(t)} \left[\sum_{i=1}^{k_{d:j}(t)-1} a_j(i) + a_{I:j}(k_{d:j}(t), t) \right] - \sum_{i=1}^{[[t/T]]} h(i) \\
&\le \sum_{j=1}^{J} a_j(0) + \sum_{j=1}^{J} \sum_{i=1}^{k_{d:j}(t)} a_j(i) \\
&- \sum_{i=1}^{[[t/T]]} h(i) \le a(0) + \sum_{j=1}^{J} \sum_{i=1}^{[[t/T]]} a_j(i) \\
&- \sum_{i=1}^{[[t/T]]} h(i) = X^D + \sum_{i=1}^{[[t/T]]} a(i) - \sum_{i=1}^{[[t/T]]} h(i) \\
&= X^D + \sum_{i=1}^{[[t/T]]} h(i) - \sum_{i=1}^{[[t/T]]} h(i) = X^D
\end{aligned} \qquad (28)
$$

which ends the proof of Theorem 2.

Remark 2.1. If the reference value X^D is smaller than or equal to the capacity of the buffer, Theorem 2 can be restated as: "The memory buffer of the bottleneck node never gets overflown". As a consequence, the problem of node congestion, packet loss and retransmissions is completely eliminated.

Remark 2.2. It is worth mentioning that there is no minimum reference value X^D required for Theorem 2 to be valid. Nevertheless, since the node works in the store&forward mode, it is required that the capacity of the buffer and the reference value are both equal to at least the maximum possible packet size.

Theorem 3. Assume that reference value X^D, and, consequently, the capacity of the buffer, satisfy the following condition

$$
X^D \ge \left(\frac{1}{J} \sum_{j=1}^{J} RTTC_j + 2T \right) d_{\max}. \qquad (29)
$$

Assume also that there exists $k_S > 0$ such that

i) for any $t \ge k_S T$ the throughput available for the bottleneck node is limited to $h_N \le d_{\max}$;

ii) once control packet k_S is sent, every source is able to send data at least at the rate $h_S \ge h_N / J$.

Provided that the assumptions mentioned above are met we have

$$\underset{k_U > k_S}{\exists} \quad \underset{k > k_U}{\forall} \quad x(k) + \int_{kT}^{(k+1)T} u(\tau)d\tau \geq h_N T. \tag{30}$$

Proof. Consider j-th virtual connection. First notice that for any $t \geq t_{R:j}(k_S)$, within time period $(t;\, t+T]$ exactly one control packet reaches the source, so the value $w_{Q:j}$ can be increased by at most $h_N T$. On the other hand, it is assured by the theorem assumptions that during the same time period the source is able to send $h_S T > h_N T$ data. Thus the value $w_{Q:j}$ is consequently decreased, and beginning from a number denoted as $k_{Z:j}$ there are no queued control values when another control packet reaches the source

$$\underset{j=1,2,\ldots,J}{\forall} \quad \underset{k_{Z:j} > k_S}{\exists} \quad \underset{k > k_{Z:j}}{\forall} \quad w_{Q:j}(t_{R:j}(k)) = 0. \tag{31}$$

It allows us to define $k_{U:j}$ as the number of the first control packet sent by the source after it reaches the state mentioned above

$$\underset{j=1,2,\ldots,J}{\forall} k_{U:j} = \min\{k > k_{Z:j} : t(k) - T_{F:j} \geq t_{R:j}(k_{Z:j})\}, \tag{32}$$

$$k_U = \max\{\, k_{U:j},\ j = 1, 2, \ldots, J\}. \tag{33}$$

Let $k > k_U$. If $x(k) \geq h_N T$, statement (30) is obviously valid. Assume that there exists positive $R \leq h_N T$ such that $x(k) = h_N T - R$.

First notice that

$$w_{Q:j}(t_{R:j}(k)) = w_{Q:j}(k) + w_{B:j}(k) - \int_{t(k)+T_{F:j}}^{t_{R:j}(k)+T_{F:j}} u_j(\tau)d\tau \tag{34}$$

what implies

$$\begin{aligned} &\sum_{j=1}^{J} w_{Q:j}(t_{R:j}(k)) \\ &= \sum_{j=1}^{J}\left[w_{Q:j}(k) + w_{B:j}(k) - \int_{t(k)+T_{F:j}}^{t_{R:j}(k)+T_{F:j}} u_j(\tau)d\tau\right] \\ &= w_Q(k) + w_B(k) - \sum_{j=1}^{J}\int_{t(k)+T_{F:j}}^{t_{R:j}(k)+T_{F:j}} u_j(\tau)d\tau. \end{aligned} \tag{35}$$

Combining this with relation (31) we obtain

$$0 = w_Q(k) + w_B(k) - \sum_{j=1}^{J}\int_{t(k)+T_{F:j}}^{t_{R:j}(k)+T_{F:j}} u_j(\tau)d\tau \tag{36}$$

$$w_Q(k) = \sum_{j=1}^{J}\int_{t(k)+T_{F:j}}^{t_{R:j}(k)+T_{F:j}} u_j(\tau)d\tau - w_B(k). \tag{37}$$

Since $a(k) = X^D - x(k) - w(k) = X^D - x(k) - w_B(k) - w_Q(k) - w_F(k)$ we have

$$\begin{aligned} w_F(k) &= X^D - x(k) - w_B(k) - w_Q(k) - a(k) \\ &= X^D - x(k) - w_B(k) - \sum_{j=1}^{J}\int_{t(k)+T_{F:j}}^{t_{R:j}(k)+T_{F:j}} u_j(\tau)d\tau \\ &\quad + w_B(k) - a(k) = X^D - x(k) \\ &\quad - \sum_{j=1}^{J}\int_{t(k)+T_{F:j}}^{t_{R:j}(k)+T_{F:j}} u_j(\tau)d\tau - a(k). \end{aligned} \tag{38}$$

Consider again j-th virtual connection and time period $[t(k);\, t_{R:j}(k)]$. As a consequence of relation (31), value of w_Q cannot exceed the maximum control value that can reach the source. Particularly

$$w_{Q:j}(t(k)) \leq \frac{1}{J} h_N T. \tag{39}$$

Note that the control values received by the source are also restricted by above relation. Moreover, during the considered time period at most $[[T_{B:j}/T]]$ control packets may reach the source. Summarising these considerations we get

$$\begin{aligned} \int_{t(k)+T_{F:j}}^{t_{R:j}(k)+T_{F:j}} u_j(\tau)d\tau &\leq \left(\left[\left[\frac{T_{B:j}}{T}\right]\right] + 1\right)\frac{1}{J} h_N T \\ &\leq \frac{1}{J} h_N (T_{B:j} + T). \end{aligned} \tag{40}$$

From relation (38), taking into account above estimation and assumptions of the theorem, we obtain

$$\begin{aligned} w_F(k) &= X^D - x(k) - \sum_{j=1}^{J}\int_{t(k)+T_{F:j}}^{t_{R:j}(k)+T_{F:j}} u_j(\tau)d\tau \\ &\quad - a(k) \geq \left(\frac{1}{J}\sum_{j=1}^{J} RTTC_j + 2T\right) d_{\max} \\ &\quad - (h_N T - R) - h_N \sum_{j=1}^{J}\frac{1}{J}(T_{B:j} + T) - h_N T \\ &\geq \left(\frac{1}{J}\sum_{j=1}^{J} RTTC_j + 2T\right) h_N - 2h_N T \\ &\quad + R - h_N \frac{1}{J}\sum_{j=1}^{J}(T_{B:j} + T) \\ &\geq h_N \frac{1}{J}\sum_{j=1}^{J}(RTTC_j - T_{B:j} - T) + 2T h_N \\ &\quad - 2h_N T + R = h_N \frac{1}{J}\sum_{j=1}^{J}(T_{F:j} - T) + R. \end{aligned} \tag{41}$$

Note that w_F denotes amount of data that is already sent by the source but not yet received by the node. The data must reach the node within time period $[t(k);\ t(k)+T_{F:j}]$, so we have

$$\begin{aligned} w_F(k) &= \sum_{j=1}^{J} \int_{t(k)}^{t(k)+T_{F:j}} u_j(\tau)d\tau \\ &= \sum_{j=1}^{J} \int_{t(k)}^{t(k)+T} u_j(\tau)d\tau + \sum_{j=1}^{J} \int_{t(k)+T}^{t(k)+T_{F:j}} u_j(\tau)d\tau \\ &\geq h_N \frac{1}{J} \sum_{j=1}^{J} (T_{F:j} - T) + R. \end{aligned} \tag{42}$$

This can be rewritten as follows

$$\begin{aligned} \sum_{j=1}^{J} \int_{t(k)}^{t(k)+T} u_j(\tau)d\tau \geq h_N \frac{1}{J} \sum_{j=1}^{J} (T_{F:j} - T) \\ + R - \sum_{j=1}^{J} \int_{t(k)+T}^{t(k)+T_{F:j}} u_j(\tau)d\tau. \end{aligned} \tag{43}$$

Now we apply reasoning that led to relation (40) to time period $[t(k)+T;\ t(k)+T_{F:j}]$ and we obtain the following inequality

$$\int_{t(k)+T}^{t(k)+T_{F:j}} u_j(\tau)d\tau \leq \frac{1}{J} h_N (T_{F:j} - T). \tag{44}$$

Combining relations (43) and (44) we get

$$\begin{aligned} &\int_{t(k)}^{t(k)+T} u(\tau)d\tau = \sum_{j=1}^{J} \int_{t(k)}^{t(k)+T} u_j(\tau)d\tau \\ &\geq h_N \frac{1}{J} \sum_{j=1}^{J} (T_{F:j} - T) + R - \sum_{j=1}^{J} \int_{t(k)+T}^{t(k)+T_{F:j}} u_j(\tau)d\tau \\ &\geq h_N \frac{1}{J} \sum_{j=1}^{J} (T_{F:j} - T) + R - \sum_{j=1}^{N} \frac{1}{J} h_N (T_{F:j} - T) = R. \end{aligned} \tag{45}$$

Finally, taking into account inequality (45) and assumption $x(k) = h_N T - R$, we estimate the left-hand side of inequality (30) as follows

$$\begin{aligned} x(k) + \int_{t(k)}^{t(k)+T} u(\tau)d\tau = x(k) + \sum_{j=1}^{J} \int_{t(k)}^{t(k)+T} u_j(\tau)d\tau \\ \geq x(k) + R = h_N T - R + R = h_N T. \end{aligned} \tag{46}$$

Therefore inequality (30) is valid for arbitrarily chosen $k > k_U$. This conclusion ends the proof of Theorem 3.

Remark 3.1. The practical consequence of Theorem 3 is that if its assumptions are satisfied, full utilisation of the available throughput at the bottleneck node is guaranteed by the proposed control strategy. Indeed, since we assumed that the throughput available for bottleneck node is bounded to h_N, maximum amount of data that could be sent within time period $(kT;\ (k+1)T]$ is not greater than $h_N T$. On the other hand, relation (30) ensures that there is at least $h_N T$ data to be sent at the node. Consequently, the available throughput is the only factor that impacts the rate at which the bottleneck node is able to send the data.

Remark 3.2. It is worth emphasizing that the validity of Theorem 3 does not depend on the state of the considered system before k_S control packets are sent by the sources. In other words, it is not required for the sources to be persistent, that is, to be able to send data at maximum rate at any time. Again, this feature is very important from the practical perspective, because in the case of real network the ability of the source to send data is limited not only by the state of the network (i.e. throughput available for the source), but also by the application that provides the source with data to be sent.

4. Summary

In this paper a new flow control strategy for connection-oriented, packet-switching networks has been proposed. It employs the Smith predictor combined with a dead-beat controller. On the contrary to numerous works in this field, control values are interpreted by the data sources as the quantity of data that is to be sent instead of the rate transmission. This is motivated by the fact that in real packet-switching networks it is not possible to precisely control the sending rate, particularly to change the rate at any given time instant. Applying the quantity-based control scheme we assume that the packets are sent at maximum possible rate (typically equal to the physical layer bandwidth), which is coherent with the way the real packet-switching networks operate. The most significant properties of proposed control strategy have been introduced in a form of mathematical theorems, and thoroughly discussed from the perspective of real network applications. First, it is ensured that the control values are nonnegative and bounded, which is necessary for the control strategy to be applicable. Moreover, it is guaranteed that the problem of node congestion and packet loss is completely eliminated, as the length of the packet queue in the buffer of the bottleneck node is bounded by reference value X^D. Furthermore, the proposed control strategy achieves full utilisation of the throughput available to the considered node, provided that the reference queue length is set according to (29). These favourable properties are achieved with no a priori knowledge of the throughput available to considered node, except that it is upper bounded. Finally, it is worth emphasizing that the control strategy described in this paper has been successfully adopted to perform flow control in IP networks.

Acknowledgements. This work has been performed within the framework of a project "Optimal sliding mode control of time delay systems" financed by the National Science Centre of Poland – under the decision number DEC-2011/01/B/ST7/02582.

REFERENCES

[1] R. Jain, "Congestion control and traffic management in ATM networks: recent advances and a survey", *Computer Networks and ISDN Systems* 28 (13), 1723–1738 (1996).

[2] M. Lee, D.J. Im, Y.K. Lee, J. Lee, S. Lee, K.K. Lee, and H. Kang, "Algorithm for ABR traffic control and formation feedback information", *Proc. Comput. Science and its Applications – ICCSA 2005* 1, CD-ROM (2005).

[3] Y. He, N. Xiong, and Y. Yang, "Data transmission rate control in computer networks using neural predictive networks", *Proc. Parallel and Distributed Processing and Applications Second Int. Symposium, ISPA 2004* 1, CD-ROM (2004).

[4] S. Jagannathan and J. Talluri, "Predictive congestion control of ATM networks: multiple sources/single buffer scenario", *Automatica* 38 (5), 8815–820 (2002).

[5] D.H. Sun, Q.H. Zhang, and Z.C. Mu, "Single parametric fuzzy adaptive PID control and robustness analysis based on the queue size of network node", *Proc. 3rd Int. Conf. on Machine Learning and Cybernetics* 1, 397–400 (2004).

[6] H. Chen and Y. Li, "Intelligent flow control under game theoretic framework" in.: *Telecommunications Optimization*: *Heuristic and Adaptive Techniques*, eds.: D.W. Corne, G.D. Smith, M.J. Oats, Wiley, London, 2000.

[7] I. Sahin and M.A. Simaan, "Competitive flow control in general multi-node multi-link communication networks", *Int. J. Communication Systems* 21 (2), 167–184 (2008).

[8] L. Benmohamed and S.M. Meerkov, "Feedback control of congestion in packet switching networks: the case of a single congested node", *IEEE/ACM Trans. on Networking* 1 (6), 693–708 (1993).

[9] L. Benmohamed and S.M. Meerkov, "Feedback control of congestion in packet switching networks: the case of multiple congested nodes", *Int. J. Communication System* 10 (5), 227–246 (1997).

[10] F. Blanchini, R. Lo Cigno, and R. Tempo, "Robust rate control for integrated services packet networks", *IEEE/ACM Trans. on Networking* 10 (5), 644–652 (2002).

[11] P. Ignaciuk and A. Bartoszewicz, "Linear quadratic optimal discrete-time sliding-mode controller for connection-oriented communication networks", *IEEE Trans. on Industrial Electronics* 55 (11), 4013–4021 (2008).

[12] P. Ignaciuk and A. Bartoszewicz, "Linear quadratic optimal sliding mode flow control for connection-oriented communication networks", *Int. J. Robust and Nonlinear Control* 19 (4), 442–461 (2009).

[13] A. Bartoszewicz and J. Żuk, "Sliding mode approach to congestion control in connection-oriented communication networks", *J. Applied Computer Science* 17 (1), 9–25 (2009).

[14] A. Bartoszewicz and J. Żuk, "Discrete time sliding mode flow controller for multi-source connection-oriented communication networks", *J. Vibration and Control* 15 (11), 1745–1760 (2009).

[15] J. Żuk, "Discrete sliding control of data flow in connection-oriented communication networks", *Ph.D. Thesis*, Technical University of Łódź, Łódź, 2010, (in Polish).

[16] O.C.Imer, S. Compans, T. Basar, and R. Srikant, "Available bit rate congestion control in ATM networks", *IEEE Control Systems Magazine* 21 (1), 38–56 (2001).

[17] K.P. Laberteaux, Ch. Rohrs, and P. Antsaklis, "A practical controller for explicit rate congestion control", *IEEE Trans. on Automatic Control* 47 (6), 960–978 (2002).

[18] S. Mascolo, "Congestion control in high-speed communication networks using the Smith principle", *Automatica* 35 (12), 1921–1935 (1996).

[19] S. Mascolo, "Smith's principle for congestion control in high-speed data networks", *IEEE Trans. on Automatic Control* 45 (2), 358–364 (2000).

[20] S. Mascolo, "Dead-time and feed-forward disturbance compensation for congestion control in data networks", *Int. J. Systems Science* 34 (10–11), 627–639 (2003).

[21] S. Mascolo, "Modeling the Internet congestion control using a Smith controller with input shaping", *Control Eng. Practice* 14 (4), 425–435 (2006).

[22] A. Bartoszewicz and T. Molik, "ABR traffic control over multi-source single-bottleneck ATM networks", *J. Applied Mathematics and Computer Science* 14 (1), 43–51 (2004).

[23] M. Karbowańczyk and A. Bartoszewicz, "Flow control in a single connection ATM network with a limited source capability". *J. Applied Computer Science* 13, 35–46 (2005).

[24] A. Bartoszewicz, "Nonlinear flow control strategies for connection oriented communication networks", *Proc. IET Part D: Control Theory and Applications* 153 (1), 21–28 (2006).

[25] F. Gómez-Stern, J.M. Fornés, and F.R. Rubio, "Dead-time compensation for ABR traffic control over ATM networks", *Control Eng. Practice* 10 (5), 481–491 (2002).

[26] A. Pietrabissa, F. Delli Priscoli, A. Fiaschetti, and F. Di Paolo, "A robust adaptive congestion control for communication networks with time-varying delays", *Proc. IEEE Int. Conf. on Control Applications* 1, 2093–2098 (2006).

[27] ATM Forum Traffic Management Working Group, *Traffic Management Specification Version 4.1*, (1999).

[28] M. Karbowańczyk, "Overloading prevention in IP networks with the use of a feedback and the Smith predictor and the dead-beat controller", *Ph.D. Thesis,* Technical University of Łódź, Łódź, 2011, (in Polish).