JOURNAL OF APPLIED
COMPUTER SCIENCE
Vol. xx. No xx (200x), pp. xx-xx

Sliding mode approach to congestion control in connection-oriented communication networks

Andrzej Bartoszewicz, Justyna Żuk
Technical University of Łódź
Institute of Automatic Control
18/22 Stefanowskiego St., 90-924 Łódź, Poland
e-mail: andrzej.bartoszewicz@p.lodz.pl, justynazuk@gmail.com

Abstract. *In this paper, a novel sliding mode flow controller design for the connection-oriented communication networks is proposed. The networks are modeled as discrete time systems with the available bandwidth acting as disturbance. The proposed controller is designed in such a way that the closed-loop system stability and fast, finite time error convergence are ensured. In order to avoid the problem of excessive control signal magnitude, a sliding mode controller with saturation is proposed. When this controller is applied no bottleneck link buffer overflow and full utilization of its available bandwidth are guaranteed. Furthermore, transmission rates generated by the controller are always upper bounded and nonnegative.*

1. Introduction

High-speed connection-oriented communication networks may allow various kinds of applications to run under a uniform infrastructure. In these networks the sequence of application data units are transmitted by a source and reach their destination via a path of intermediate switches. On each switch a server schedules and forwards data units along the path from their source to their destination in the network. The difficulty of the flow control is mainly caused by long propagation delays in the network. If congestion occurs at a specific switch, information about these circumstances must be conveyed to all the sources transmitting data units through the switch. This information is used to adjust source rates and may affect the congested switch after the round trip propagation delay.

Flow control in connection-oriented communication networks has recently become an exciting research field and valuable results have been reported in many papers [3, 6, 10-15, 17]. Their authors proposed ‘on-off’ [3, 6], classical proportional-derivative (PD) [14], fuzzy proportional-integral-derivative (fuzzy PID) [17], stochastic [11], adaptive [13] and neural network based [12] controllers. Due to the significant propagation delays several researchers also applied the Smith predictors [3, 10, 15] for the flow control in such networks.

On the other hand, it is well known that sliding mode control is an attractive and efficient strategy which offers robustness and good dynamic performance of the controlled systems [2, 5, 7, 16, 18]. Therefore, in this paper we attempt to apply discrete time sliding mode approach [1, 4, 8, 9] to the flow control in a connection-oriented communication network. We consider a model of the networks which provide feedback mechanism. An example of such networks is Available Bit Rate (ABR) service in Asynchronous Transfer Mode (ATM) standard. The proposed control algorithms employ an appropriately defined sliding hyperplane, which ensures the closed-loop system stability and finite time error convergence to zero. In order to avoid the problem of excessive control signal magnitude, we propose a sliding mode controller with saturation. When this controller is applied no data loss and full link bandwidth utilization are ensured. These desirable properties are explicitly proved. Moreover, the relation between the data flow rate and the consumed bandwidth is derived.

The remainder of this paper is organized as follows. In Section 2, detailed description of the network model is given. Afterwards, in Section 3, the proposed sliding mode flow controller design and the system performance when the controller is applied are presented. In this section the important properties of the controlled network are also stated (in a lemma and three theorems) and explicitly proved. A simulation example, illustrating the discussed properties, is presented in Section 4. Finally, Section 5 concludes the paper.

2. Network model

In this paper a single virtual circuit in a connection-oriented communication network is considered. Furthermore, it is assumed that there is only one bottleneck node in the network. The source sends data (as determined by the controller at the bottleneck node) and special control units. The control units carry information about the network state. After reaching their destination, they are immediately sent back to the source, along the same path they arrived. The information carried by the control units is used to adjust the amount of data transmitted by the source at each control period. The control units are processed by the intermediate nodes on a priority basis, i.e. they are not queued but sent to the next node without delay. Consequently, the round trip time *RTT* of the control units in the virtual circuit is constant. Moreover, this time can be expressed as the sum of forward and backward propagation delays denoted as T_F and T_B, respectively

$$RTT = T_F + T_B \tag{1}$$

The block diagram of the flow control system considered in this paper is shown in Figure.

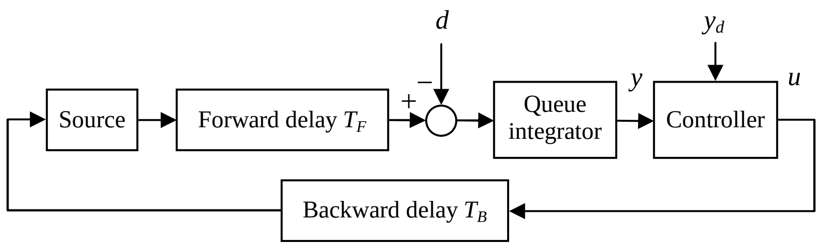

Fig. 1. Network model

Further in the paper, T represents the discretisation period, $y(kT)$ denotes the bottleneck queue length at time instants kT, $k = 0, 1, 2, \ldots$, and $y_d > 0$ is the demand value of $y(kT)$. It is assumed that before setting up the connection, the bottleneck buffer is empty, i.e. $y(kT < 0) = 0$. Moreover, in this paper we assume that the round trip time is a multiple of the discretisation period, i.e. $RTT = m_{RTT} T$, where m_{RTT} is a positive integer.

The amount of data to be sent is generated by the controller placed at the bottleneck node. The controller output at time kT is denoted as $u(kT)$. This amount of data will be sent by the source after backward delay T_B and will arrive at the bottleneck node T_F later. Consequently, the bottleneck buffer for any time $kT \le RTT$ remains empty. It is assumed that before setting up the connection

$$u(kT < 0) = 0 \tag{2}$$

The amount of data which may leave the bottleneck buffer at time kT is modelled as an *a priori* unknown bounded function of time $d(kT)$, for $k = 0, 1, 2, \ldots$. The maximum value of $d(kT)$ is denoted by $d_{\max}$ and $h(kT)$ represents the amount of data actually leaving the bottleneck node at time kT. Consequently

$$\forall_{k \ge 0} \quad 0 \le h(kT) \le d(kT) \le d_{\max} \tag{3}$$

Initially, the bottleneck buffer is empty

$$y(kT = 0) = 0 \tag{4}$$

Then, for any $k \ge 1$, the queue length can be expressed as follows

$$y(kT)=\sum_{j=0}^{k-1}u(jT-RTT)-\sum_{j=0}^{k-1}h(jT)=\sum_{j=0}^{k-m_{RTT}-1}u(jT)-\sum_{j=0}^{k-1}h(jT) \tag{5}$$

3. Sliding mode controllers

In this section, the flow control problem for the described network is considered. First, a chattering free discrete time sliding mode controller is designed so that fast and finite time error convergence to zero is achieved. Then, we propose a modified sliding mode control strategy which takes into account physical constraints and ensures that the maximum admissible flow rate is never exceeded.

3.1. Proposed control strategy

This subsection focuses on the design of a discrete time sliding mode flow controller for the communication network, whose model was introduced in Section 2. For this purpose, first a discrete time state space model of the controlled system is formulated. Then, an appropriate sliding plane is introduced and its parameters are determined in such a way that the closed-loop system is stable and the error converges to zero in finite time.

Let us consider the following discrete time model of the network

$$\begin{aligned}&\boldsymbol{x}\left[(k+1)T\right]=\boldsymbol{A}\boldsymbol{x}(kT)+\boldsymbol{b}u(kT)+\boldsymbol{p}h(kT)\\&y(kT)=\boldsymbol{q}^{\mathrm{T}}\boldsymbol{x}(kT)\end{aligned} \tag{6}$$

where $\boldsymbol{x}(kT) = [x_1(kT)\ \ x_2(kT)\ \ \dots\ x_n(kT)]^{\mathrm{T}}$ is the state vector with $x_1(kT) = y(kT)$, $\boldsymbol{A}$ is $n \times n$ state matrix, $\boldsymbol{b}$, $\boldsymbol{p}$ and $\boldsymbol{q}$ are $n \times 1$ vectors

$$\boldsymbol{A}=\begin{bmatrix}1&1&0&\dots&0\\0&0&1&\dots&0\\\vdots&\vdots&\vdots&\ddots&\vdots\\0&0&0&\dots&1\\0&0&0&\dots&0\end{bmatrix}\quad \boldsymbol{b}=\begin{bmatrix}0\\0\\\vdots\\0\\1\end{bmatrix}\quad \boldsymbol{p}=\begin{bmatrix}-1\\0\\\vdots\\0\\0\end{bmatrix}\quad \boldsymbol{q}=\begin{bmatrix}1\\0\\\vdots\\0\\0\end{bmatrix} \tag{7}$$

and $n = m_{RTT} + 1$. Alternatively, the state space equation can be written as follows

$$\begin{cases} x_1[(k+1)T] = x_1(kT) + x_2(kT) - h(kT) \\ x_2[(k+1)T] = x_3(kT) \\ x_3[(k+1)T] = x_4(kT) \\ \vdots \\ x_{n-1}[(k+1)T] = x_n(kT) \\ x_n[(k+1)T] = u(kT) \end{cases} \tag{8}$$

In this model the available bandwidth $h(kT)$ is represented as unmatched disturbance. The desired state of the system is denoted by $\boldsymbol{x}_d = [x_{d1}\ \ x_{d2}\ \dots\ \ x_{dn}]^{\mathrm{T}}$. It can be noticed from (8) that all components x_{di} of vector $\boldsymbol{x}_d$ for $i = 2, \dots, n$ are equal to zero when $h(kT) = 0$. Let us denote the first state variable x_{d1} representing the demand queue length by y_d.

We introduce a sliding hyperplane described by the following equation

$$s(kT) = \boldsymbol{c}^{\mathrm{T}}\boldsymbol{e}(kT) = 0 \tag{9}$$

where $\boldsymbol{c}^{\mathrm{T}} = [c_1\ \ c_2\ \dots\ \ c_n]$ is such a vector that $\boldsymbol{c}^{\mathrm{T}}\boldsymbol{b} \neq 0$. Similarly as it is usually done when designing the control systems, now we neglect the effect of disturbance $h(kT)$ in the controller design process. However, this does not imply that the disturbance is disregarded in the paper, it will be given full consideration when analysing the system performance. The closed-loop system error is denoted as $\boldsymbol{e}(kT) = \boldsymbol{x}_d - \boldsymbol{x}(kT)$. Hence, substituting (6) into equation $\boldsymbol{c}^{\mathrm{T}}\boldsymbol{e}[(k + 1)T] = 0$ the following feedback control law can be derived

$$u(kT) = \left(\boldsymbol{c}^{\mathrm{T}}\boldsymbol{b}\right)^{-1}\boldsymbol{c}^{\mathrm{T}}\left[\boldsymbol{x}_d - \boldsymbol{A}\boldsymbol{x}(kT)\right] \tag{10}$$

When this control signal is applied, the closed-loop system state matrix has the following form $\boldsymbol{A}_c = \left[\boldsymbol{I}_n - \boldsymbol{b}\left(\boldsymbol{c}^{\mathrm{T}}\boldsymbol{b}\right)^{-1}\boldsymbol{c}^{\mathrm{T}}\right]\boldsymbol{A}$. Then the characteristic polynomial of $\boldsymbol{A}_c$ can be found as follows

$$\det\left(z\boldsymbol{I}_n - \boldsymbol{A}_c\right) = z^n + \frac{c_{n-1} - c_n}{c_n} z^{n-1} + \frac{c_{n-2} - c_{n-1}}{c_n} z^{n-2} + \dots + \frac{c_1 - c_2}{c_n} z \tag{11}$$

which leads to the condition $c_n \neq 0$. Asymptotic stability of the discrete time system is ensured if and only if all its eigenvalues are located inside the unit circle. Moreover, in order to ensure the closed-loop system error convergence to zero in finite time, the characteristic polynomial should have all roots equal to zero. Therefore, (11) has to satisfy

$$\det\left(z\boldsymbol{I}_n - \boldsymbol{A}_c\right) = z^n \tag{12}$$

Comparing coefficients on the right-hand sides of (11) and (12), the following form of vector $\boldsymbol{c}$ is obtained

$$\boldsymbol{c}^{\mathrm{T}} = \begin{bmatrix} 1 & 1 & 1 & \ldots & 1 \end{bmatrix} c_n \tag{13}$$

Substituting (7) and (13) into (10) the following state feedback control can be derived

$$u(kT) = y_d - \sum_{i=1}^{n} x_i(kT) \tag{14}$$

Alternatively, from (8), one can get the state variables x_i ($i = 2, 3, \ldots, n$) expressed in terms of the control signal generated by the controller at the previous $n - 1$ samples

$$x_i(kT) = u\left[(k-n+i-1)T\right] \quad \text{for } i = 2, 3, \ldots, n \tag{15}$$

Substituting these expressions into (14) and putting $x_1(kT) = y(kT)$, we obtain

$$\begin{aligned} u(kT) &= y_d - \left\{ y(kT) + u\left[(k-(n-1))T\right] + u\left[(k-(n-2))T\right] + \ldots + u\left[(k-1)T\right] \right\} = \\ &= y_d - y(kT) - \sum_{j=1}^{n-1} u\left[(k-j)T\right] = y_d - y(kT) - \sum_{j=1}^{m_{RTT}} u\left[(k-j)T\right] = \\ &= y_d - y(kT) - \sum_{j=k-m_{RTT}}^{k-1} u(jT) \end{aligned} \tag{16}$$

which actually represents a dynamic sliding mode flow controller. This completes the design of the flow control algorithm which guarantees the closed-loop system stability and fast, finite time error convergence to zero in the considered network.

Unfortunately, the strategy proposed in this section may generate initial flow rate of unacceptable magnitude. Therefore, in order to avoid this undesirable effect, further in the paper a modified sliding mode control strategy will be proposed.

3.2. Modified control strategy

In this section we introduce a new flow control strategy. The amount of data to be sent by the source at time kT is now determined by the controller according to the following formula

$$u(kT) = \min\left\{ y_d - y(kT) - \sum_{j=k-m_{RTT}}^{k-1} u(jT),\quad u_{\max} \right\} \tag{17}$$

where $u_{\max} > d_{\max}$ is the maximum admissible value of the flow rate. One can easily notice that this strategy determines data transmission rates which never exceed the predetermined value $u_{\max}$.

Let us denote the first argument of the min{·,·} function in (17) as $w(kT)$, i.e.

$$w(kT) = y_d - y(kT) - \sum_{j=k-m_{RTT}}^{k-1} u(jT) \tag{18}$$

It directly follows from (17) that at any time instant $kT \geq 0$ inequality

$$u(kT) \leq w(kT) \tag{19}$$

is satisfied. Moreover, we will prove that the flow rate generated according to strategy (17) is always nonnegative. This is shown in the following lemma.

Lemma If the proposed control algorithm is applied, then data transmission rate $u(kT)$ is always nonnegative, i.e.

$$\forall_{k \geq 0} \quad u(kT) \geq 0 \tag{20}$$

Proof: At the initial time function $w(kT = 0) = y_d$. Therefore, the flow rate $u(0)$ either equals y_d or $u_{\max}$. Consequently, inequality (20) is satisfied for $k = 0$. Furthermore, at any time instant $kT > 0$ if the amount of data to be sent by the source is $u_{\max}$, then the flow rate $u(kT)$ is also strictly positive. Hence, in order to complete the proof it is only necessary to show that (20) is satisfied for any $k > 0$ when $u(kT) = w(kT)$. For that purpose, let us notice that the following relation can be derived from (5)

$$y(kT) = y[(k-1)T] + u\left[(k - m_{RTT} - 1)T\right] - h[(k-1)T] \tag{21}$$

which holds for any positive integer k. Since in the analysed case $u(kT) = w(kT)$ and the bottleneck queue length satisfies (21), we obtain

$$
\begin{aligned}
u(kT) &= w(kT) = y_d - y(kT) - \sum_{j=k-m_{RTT}}^{k-1} u(jT) = \\
&= y_d - y[(k-1)T] - u\left[(k - m_{RTT} - 1)T\right] + h[(k-1)T] - \sum_{j=k-m_{RTT}}^{k-1} u(jT) = \\
&= y_d - y[(k-1)T] - \sum_{j=k-m_{RTT}-1}^{k-1} u(jT) + h[(k-1)T] = \\
&= y_d - y[(k-1)T] - \sum_{j=k-m_{RTT}-1}^{k-2} u(jT) - u[(k-1)T] + h[(k-1)T] = \\
&= w[(k-1)T] - u[(k-1)T] + h[(k-1)T]
\end{aligned}
\tag{22}
$$

Taking into account inequality (19) and the fact that the consumed bandwidth is always nonnegative, we obtain

$$u(kT) \geq h[(k-1)T] \geq 0 \tag{23}$$

which shows that inequality (20) indeed holds at any time instant $kT > 0$ when $u(kT) = w(kT)$. This conclusion ends the proof of the lemma.

Thus we conclude that controller (17) generates data transmission rate which is always nonnegative and upper bounded, i.e.

$$\forall_{k \geq 0} \quad 0 \leq u(kT) \leq u_{\max} \tag{24}$$

Clearly, this property is of utmost importance for the practical implementation of the strategy in any real network. In the sequel, three theorems stating further important properties of the proposed flow control scheme are presented. The first one gives the condition which must be satisfied in order to eliminate the risk of data loss as a consequence of exceeding the bottleneck node buffer capacity. Afterwards, the second theorem provides a sufficient condition for the full bottleneck link bandwidth utilization. Finally, a relation between the control signal $u(kT)$ and the consumed bandwidth is formulated in the third theorem.

Theorem 1. If the proposed strategy is applied, then the queue length in the bottleneck buffer is always upper bounded by its demand value, i.e.

$$\forall_{k \geq 0} \quad y(kT) \leq y_d \tag{25}$$

Proof: As it has already been proved, data transmission rate $u(kT)$ is nonnegative at any time instant kT. On the other hand, by definition $u(kT)$ is smaller than or equal to $w(kT)$. Therefore, the following relation holds for any time $kT \geq 0$

$$y_d - y(kT) - \sum_{j=k-m_{RTT}}^{k-1} u(jT) = w(kT) \geq u(kT) \geq 0 \tag{26}$$

Hence, the queue length satisfies

$$y(kT) \leq y_d - \sum_{j=k-m_{RTT}}^{k-1} u(jT) \tag{27}$$

Again taking into account that $u(kT)$ is always nonnegative one concludes that the queue length indeed never exceeds its demand value. This ends the proof of Theorem 1.

Another desirable property of the analyzed system is full bottleneck link bandwidth utilisation. Since the bottleneck link bandwidth $d(kT)$ is fully used if the queue length $y[(k+1)T]$ is strictly greater than zero, then the next theorem specifies a condition which guarantees that the queue length in our scheme is always strictly positive.

Theorem 2. If $u_{\max} > d_{\max}$ and the demand value of the queue length y_d satisfies the following inequality

$$y_d > (m_{RTT} + 1)u_{\max} \tag{28}$$

then for any $k \geq m_{RTT} + 1$ the queue length in the bottleneck buffer is always strictly positive.

Proof: Let us define an auxiliary function

$$\varphi(kT) = y(kT) + \sum_{j=k-m_{RTT}}^{k-1} u(jT) \tag{29}$$

This function represents the amount of data currently waiting in the bottleneck buffer queue, and the amount of ‘in flight’ data, i.e. this data which has already been sent by the source but not yet arrived at the bottleneck node, and that data which will be sent by the source because the controller has already sent out an appropriate command signal to the source.

Substituting formula (5) into (29), one can express function $\varphi(kT)$ as

$$\varphi(kT) = y(kT) + \sum_{j=k-m_{RTT}}^{k-1} u(jT) = \sum_{j=0}^{k-m_{RTT}-1} u(jT) - \sum_{j=0}^{k-1} h(jT) + \sum_{j=k-m_{RTT}}^{k-1} u(jT) =$$
$$= \sum_{j=0}^{k-1} u(jT) - \sum_{j=0}^{k-1} h(jT) \tag{30}$$

Hence, taking into account condition (28) for $k = 0$

$$\varphi(kT) = \varphi(0) = 0 < y_d - (m_{RTT} + 1)u_{\max} < y_d - u_{\max} \tag{31}$$

Furthermore, if for some k the following inequality $\varphi(kT) < y_d - u_{\max}$ is satisfied, then

$$w(kT) = y_d - y(kT) - \sum_{j=k-m_{RTT}}^{k-1} u(jT) = y_d - \varphi(kT) > u_{\max} \tag{32}$$

which implies that $u(kT) = u_{\max}$. Consequently, since $u_{\max} > d_{\max}$, we conclude that if $\varphi(kT) < y_d - u_{\max}$, then function φ increases at least at the rate $u_{\max} - d_{\max}$. Moreover, since for any time $kT < RTT$ the consumed bandwidth $h(kT) = 0$, then if $\varphi(kT) < y_d - u_{\max}$ and condition (28) is satisfied, then $\varphi(kT)$ increases at the rate $u_{\max}$, reaching $m_{RTT} u_{\max}$ at the time $m_{RTT}T$.

On the other hand, the consumed bandwidth for any time $kT \geq RTT$ satisfies inequality $h(kT) \leq d_{\max}$. This implies that $\varphi(kT)$ can decrease at most at the rate $d_{\max}$. Further, we will show that function $\varphi(kT)$ after reaching $m_{RTT} u_{\max}$ never decreases below this value, i.e. we will demonstrate that the following inequality holds for any $kT > RTT$

$$\varphi(kT) > m_{RTT} u_{\max} \tag{33}$$

In order to prove this we will apply the principle of mathematical induction. Let us first check whether (33) holds for $k = m_{RTT} + 1$. If condition (28) is satisfied, then $\varphi(m_{RTT}T) = m_{RTT} u_{\max} < y_d - u_{\max}$. This implies that $u(m_{RTT}T) = u_{\max}$. Consequently

$$\begin{aligned} \varphi\left[(m_{RTT} + 1)T\right] &= \varphi(m_{RTT}T) + u(m_{RTT}T) - h(m_{RTT}T) = \\ &= m_{RTT} u_{\max} + u_{\max} - h(m_{RTT}T) \geq \\ &\geq m_{RTT} u_{\max} + u_{\max} - d_{\max} > m_{RTT} u_{\max} \end{aligned} \tag{34}$$

which shows that (33) is indeed true for $k = m_{RTT} + 1$. Now, let us assume that (33) holds for some $k \geq m_{RTT} + 1$. We will show that this implies that (33) is also satisfied for $k + 1$. For this purpose we will consider the following two cases: the first one when $u(kT) = w(kT)$ and the second one when $u(kT) = u_{\max}$. In the first situation, from (28), (32) and inequality $u_{\max} > d_{\max}$ we obtain

$$\begin{aligned}\varphi\left[(k+1)T\right] &= \varphi(kT) + u(kT) - h(kT) = \varphi(kT) + w(kT) - h(kT) = \\ &= \varphi(kT) + y_d - \varphi(kT) - h(kT) = y_d - h(kT) \geq \\ &\geq y_d - d_{\max} > y_d - u_{\max} > m_{RTT}\, u_{\max}\end{aligned} \tag{35}$$

Then, in the second situation, i.e. when $u(kT) = u_{\max}$, we can write

$$\begin{aligned}\varphi\left[(k+1)T\right] &= \varphi(kT) + u(kT) - h(kT) = \varphi(kT) + u_{\max} - h(kT) \geq \\ &\geq \varphi(kT) + u_{\max} - d_{\max} > \varphi(kT) > m_{RTT}\, u_{\max}\end{aligned} \tag{36}$$

Therefore, we conclude that relation (33) actually holds for any time $kT > RTT$.

Finally, taking into account relations (29), (33) and the fact that the flow rate generated by our controller is always upper bounded by $u_{\max}$, for any time $kT > RTT$, we get

$$y(kT) = \varphi(kT) - \sum_{j=k-m_{RTT}}^{k-1} u(jT) > m_{RTT}u_{\max} - m_{RTT}u_{\max} = 0 \tag{37}$$

which ends the proof of Theorem 2.

The theorem shows that using strategy (17) with condition (28) we ensure full bottleneck link bandwidth utilisation for any time $kT > RTT$. Further, in the next theorem, a relation between the flow rate and the consumed bandwidth is stated and proved.

Theorem 3. If the designed sliding mode flow controller is applied, the demand queue length $y_d > u_{\max}$ and the maximum flow rate $u_{\max} > d_{\max}$, then there exists such a nonnegative integer k_0 satisfying

$$k_0 < \frac{y_d - u_{\max}}{u_{\max} - d_{\max}} + 1 \tag{38}$$

that for any $k > k_0$ the following relation holds

$$u(kT) = h[(k-1)T] \tag{39}$$

Furthermore, when $y_d \leq u_{\max}$, relation (39) is satisfied for any $k \geq 1$.
Proof: First, let us consider the situation when inequality $y_d \leq u_{\max}$ holds. It will be shown that then the following relation is always satisfied

$$\underset{k\geq 0}{\forall} \quad w(kT)\leq u_{\max} \tag{40}$$

which directly implies $u(kT) = w(kT)$.

In order to prove that relation (40) is indeed satisfied for any time $kT \geq 0$, we apply the principle of mathematical induction. At the initial time $w(0) = y_d \leq u_{\max}$. Therefore, inequality (40) holds for $k = 0$. Now let us assume that (40) is true for some $k \geq 0$ and we will show that it is also satisfied for $k + 1$. Using equations (18) and (21), and taking into account that $u(kT) = w(kT)$ we get

$$\begin{aligned} w\left[(k+1)T\right] &= y_d - y\left[(k+1)T\right] - \sum_{j=k-m_{RTT}+1}^{k} u(jT) = \\ &= y_d - y(kT) - u\left[(k-m_{RTT})T\right] + h(kT) - \sum_{j=k-m_{RTT}+1}^{k} u(jT) = \\ &= y_d - y(kT) - \sum_{j=k-m_{RTT}}^{k-1} u(jT) - u(kT) + h(kT) = \\ &= w(kT) - u(kT) + h(kT) = h(kT) \leq d_{\max} < u_{\max} \end{aligned} \tag{41}$$

This ends the proof of inequality (40).

Since it follows from (40) that at any time instant $kT \geq 0$ the flow rate $u(kT) = w(kT)$, then using expression (22), for any $k \geq 1$, we obtain

$$u(kT) = w[(k-1)T] - u[(k-1)T] + h[(k-1)T] = h[(k-1)T] \tag{42}$$

Equation (42) shows that, if $y_d \leq u_{\max}$, then (39) indeed holds for any positive integer k.

Now let us consider the situation when $y_d > u_{\max}$. If for some k inequality $\varphi(kT) < y_d - u_{\max}$ is satisfied, then it follows from equation (30) and assumption $u_{\max} > d_{\max}$, that function φ increases at least at the rate $u_{\max} - d_{\max}$. Thus, there exists such a finite time instant k_0T, when the following condition

$$\varphi(kT) \geq y_d - u_{\max} \tag{43}$$

becomes satisfied for the first time.

We will determine the latest time instant when inequality (43) can become satisfied for the first time. Since function $\varphi(kT)$ is smaller than the difference $y_d - u_{\max}$ until $k < k_0$, then

$$\varphi\left[(k_0-1)T\right]=\sum_{j=0}^{k_0-2}u(jT)-\sum_{j=0}^{k_0-2}h(jT)<y_d-u_{\max} \tag{44}$$

Moreover, since the flow rate for any $k < k_0$ is equal to $u_{\max}$, then inequality (44) can be rewritten as

$$(k_0-1)u_{\max}-\sum_{j=0}^{k_0-2}h(jT)<y_d-u_{\max} \tag{45}$$

Number k_0 in this equation is the biggest, when for any time from 0 up to $(k_0 - 2)T$, the consumed bandwidth has its greatest possible value $d_{\max}$. Consequently, from relation (45) we get the following inequality

$$(k_0-1)(u_{\max}-d_{\max})<y_d-u_{\max} \tag{46}$$

which gives the estimate of k_0 specified by relation (38).

We will now demonstrate that for any time $kT > k_0T$ condition (43) is indeed satisfied. For that purpose we take into account some $k > k_0$ and we consider the two cases: the first one when $w(kT) \le u_{\max}$, and the second one when $w(kT) > u_{\max}$. In the first case from relations (18) and (29) we obtain

$$w(kT)=y_d-\varphi(kT)\le u_{\max} \tag{47}$$

From this inequality it can be easily noticed that condition (43) actually holds for any $k > k_0$.

Now let us consider the second case, i.e. the situation when $w(kT) > u_{\max}$. In this situation, in order to show that condition (43) holds for any $kT > k_0T$, one can apply the principle of mathematical induction. We have already demonstrated that there exists such a moment k_0T, when inequality (43) is satisfied. Now, let us assume that for some instant $kT > k_0T$ the considered condition holds, and we will show that this implies that the condition is also satisfied at the time instant $kT + T$. Since in the analysed case $w(kT) > u_{\max}$, then $u(kT) = u_{\max}$. Taking into account equation (30) and inequality $u_{\max} > d_{\max}$, we get

$$\begin{aligned}\varphi\left[(k+1)T\right]&=\varphi(kT)+u(kT)-h(kT)=\varphi(kT)+u_{\max}-h(kT)\ge\\&\ge\varphi(kT)+u_{\max}-d_{\max}>\varphi(kT)\ge y_d-u_{\max}\end{aligned} \tag{48}$$

Consequently, we conclude that for any $k > k_0$ inequality (43) is always satisfied.

Condition (43) implies that for any $kT > k_0T$, $w(kT) \le u_{\max}$ and $u(kT) = w(kT)$. Therefore, it immediately follows from equation (22) that relation (39) is indeed satisfied for any $k > k_0$. This ends the proof of Theorem 3.

4. Simulation example

In order to verify the properties of the sliding mode flow control strategy proposed in this paper computer simulations of the network described by equations (6) – (8) have been performed in Matlab-Simulink® environment. First, the model of the network was constructed according to the description given in Section 2. Then the system parameters were chosen as follows: the discretisation period T was selected as 1 ms and the round trip time in the virtual circuit was assumed to be $RTT = m_{RTT}T = 10$ ms ($T_F = 3$ ms, $T_B = 7$ ms). Consequently, the system order $n = m_{RTT} + 1 = 11$. The maximum available bandwidth of the bottleneck link was set as $d_{\max} = 4.8$ Mb per second, and the maximum admissible flow rate as $u_{\max} = 6.1$ Mb per second. The bandwidth actually available for the data transfer is shown in Fig.2. Sudden changes of function d, visible in the figure, reflect the most rigorous networking conditions.

According to Theorem 2, when strategy (17) is applied, the demand value of the queue length required to assure full bottleneck link bandwidth utilization in the analyzed network must be greater than 0.0671 Mb. Consequently, $y_d = 0.0687$ Mb, which is equivalent to 170 ATM cells, is chosen. The transmission rate generated by the controller and the queue length evolution are shown in Figs.3 and 4, respectively.

It can be clearly seen from the figures that the transmission rate is always nonnegative and never exceeds the maximum value $u_{\max}$. Furthermore, the queue length actually never grows beyond its demand value y_d, which ensures no data loss in the network and no need for data retransmission. Moreover, after initial period of 11 ms, i.e. for any time greater than $(m_{RTT} + 1)T = 11$ ms, the queue length is strictly positive, which implies full utilization of the bottleneck link available bandwidth.

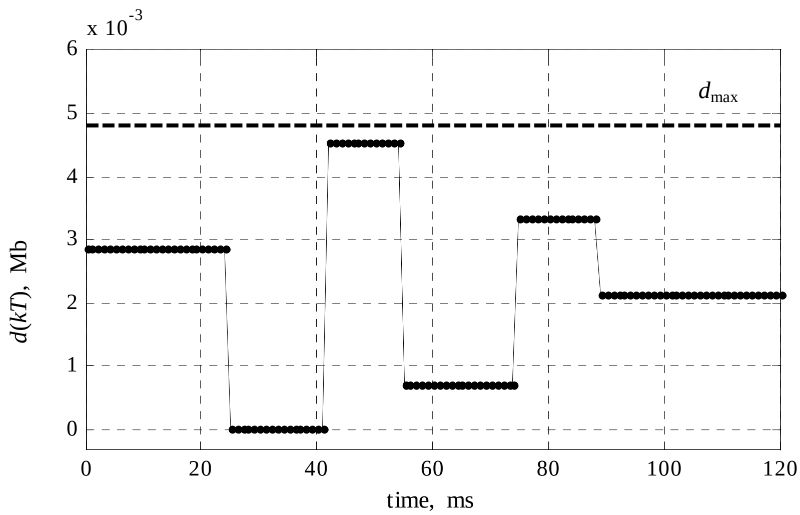

Fig. 2. Available bandwidth at the output link of the bottleneck node

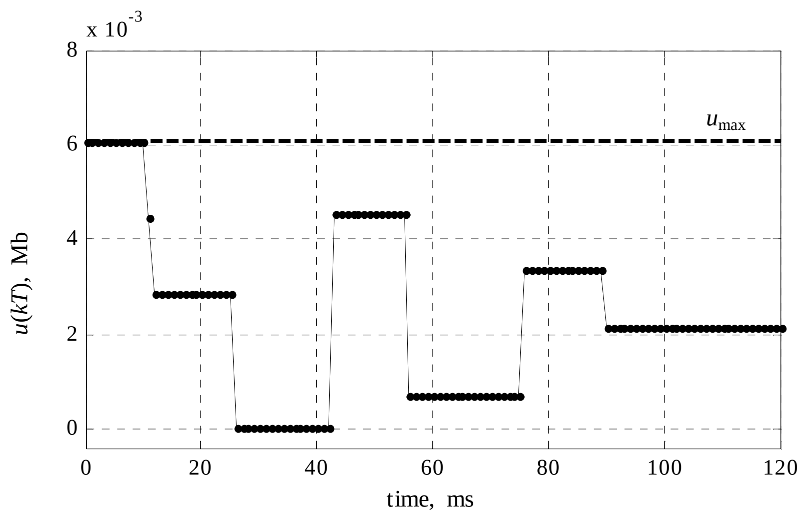

Fig. 3. Transmission rate generated by the controller

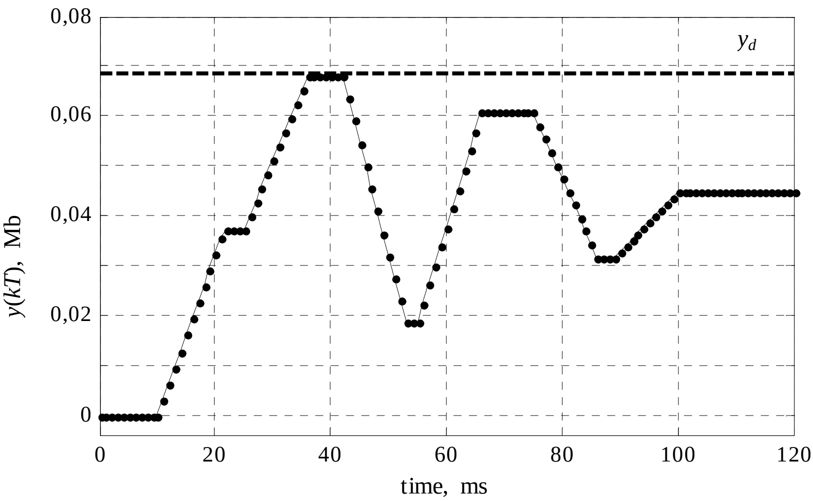

Fig. 4. Queue length

5. Conclusions

A new discrete time sliding mode flow control strategy for a single virtual circuit of a connection-oriented communication network has been presented. The strategy is designed so that the closed-loop system stability and fast, finite time error convergence are ensured. In order to avoid the problem of excessive control signal magnitude, a sliding mode controller with saturation is proposed. When this controller is applied, full bottleneck node link utilization and no data loss in the controlled network are guaranteed. The conditions ensuring these favorable properties are formulated and explicitly proved. Consequently, the need for data retransmission is eliminated and the maximum throughput is achieved. Moreover, as the flow rate generated by the controller is always nonnegative and bounded, the proposed mechanism can be feasibly incorporated in real communication networks. Our further research focuses on adapting the control strategy proposed in this paper for multi-source networks.

Acknowledgement

This work has been financed by the Polish State budget in the years 2008 – 2010 as a research project N N514 300035 “Design of the switching surfaces for the sliding mode control".

References

[1] Bandyopadhyay B., Janardhanan S.: *Discrete-time sliding mode control. A multirate output feedback approach*. Series: Lecture Notes in Control and Information Sciences, Vol. 323, Springer-Verlag Berlin Heidelberg, 2006.

[2] Bartolini G., Fridman L., Pisano A., Usai E.: *Modern Sliding Mode Control Theory. New Perspectives and Applications*. Series: Lecture Notes in Control and Information Sciences, Vol. 375, Springer-Verlag Berlin Heidelberg, 2008.

[3] Bartoszewicz A.: *Nonlinear flow control strategies for connection-oriented communication networks*. IEE Proceedings on Control Theory and Applications, 2006, Vol. 153, No.1, pp. 21-28.

[4] Bartoszewicz A.: *Discrete time quasi-sliding mode control strategies*. IEEE Transactions on Industrial Electronics, 1998, Vol. 45, No.4, pp. 633-637.

[5] Bartoszewicz A., Kaynak O., Utkin V.I. (editors): *Sliding mode control in industrial applications*. Special section: IEEE Transactions on Industrial Electronics, 2008, Vol. 55, No. 11, pp. 3805-4103.

[6] Chong S., Nagarajan R., Wang Y.: *First-order rate-based flow control with dynamic queue threshold for high-speed wide-area ATM networks*. Computer Networks and ISDN Systems, 1998, Vol. 29, pp. 2201-2212.

[7] DeCarlo R.S., Żak S., Mathews G.: *Variable structure control of nonlinear multivariable systems: a tutorial*. Proceedings of IEEE, 1988, Vol. 76, No.3, pp. 212-232.

[8] Furuta K.: *Sliding mode control of a discrete system*. Systems & Control Letters, 1990, Vol. 14, pp. 145-152.

[9] Gao W., Wang Y., Homaifa A.: *Discrete-time variable structure control systems*. IEEE Transactions on Industrial Electronics, 1995, Vol. 42, pp. 117-122.

[10] Gómez-Stern F., Fornés J., Rubio F.: *Dead-time compensation for ABR traffic control over ATM networks*. Control Engineering Practice, 2002, Vol. 10, pp. 481-491.

[11] Imer O., Compans S., Basar T., Srikant R.: *Available bit rate congestion control in ATM networks*. IEEE Control Systems Magazine, 2001, pp. 38-56.

[12] Jagannathan S., Talluri J.: *Predictive congestion control of ATM networks: multiple sources/single buffer scenario*. Automatica, 2002, Vol. 38, pp. 815-820.

[13] Laberteaux K., Rohrs Ch., Antsaklis P.: *A practical controller for explicit rate congestion control*. IEEE Transactions on Automatic Control, 2002, Vol. 47, pp. 960-978.

[14] Lengliz I., Kamoun F.: *A rate-based flow control method for ABR service in ATM networks*. Computer Networks, 2000, Vol. 34, pp. 129-138.

[15] Mascolo S.: *Congestion control in high-speed communication networks using the Smith principle*. Automatica, 1999, Vol. 35, pp. 1921-1935.

[16] Slotine J.J., Li W.: *Applied Nonlinear Control*, Prentice-Hall International Editions, 1991.

[17] Sun D.H., Zhang Q.H., Mu Z.C.: *Single parametric fuzzy adaptive PID control and robustness analysis based on the queue size of network node*. Proceedings of 2004 International Conference on Machine Learning and Cybernetics, 2004, Vol. 1, pp. 397-400.

[18] Utkin V.: *Variable structure systems with sliding modes*. IEEE Transactions on Automatic Control, 1977,Vol. 22, pp. 212-222.