Combined Nonlinear Control of Flow Rate and Bandwidth for Virtual Paths in ATM based Networks

Andreas Pitsillides* and Petros Ioannou**

***Department of Computer Science**
e-mail:cspitsil@zeus.ucy.ac.cy

****Department of Mathematics and Statistics**
e-mail: mspetros@zeus.ucy.ac.cy

University of Cyprus, P.O. Box 537, Nicosia, Cyprus

Abstract: Multimedia and multiservice broadband networks (e.g. Broadband-ISDN and ATM based LANs) will be implemented using Asynchronous Transfer Mode (ATM). In this paper we consider the combined control problem of flow rate and bandwidth (capacity, service rate) at the virtual path level in ATM based networks under nonstationary conditions. A fluid flow model in the state variable form is used to describe the time varying mean behaviour of a virtual path. A nonlinear control approach is adopted to derive the control strategy for the single node, finite buffer and server, case. The performance of the proposed scheme is evaluated using analysis and simulations. Simulation results show that the proposed scheme achieves effective server and buffer utilisation (as predicted by analysis) and by appropriate choice of the control design variables it achieves zero cell loss (even for traffic demands, which if not controlled would suffer cell losses). Hence the proposed scheme avoids retransmissions which as a result would reduce network utilisation.

Keywords: Nonlinear Control, Asynchronous Transfer Mode (ATM) based Communication Networks, Flow Control and Bandwidth Control.

1. Introduction

Asynchronous Transfer Mode (ATM) is expected to be the transfer mode for implementing multimedia and multiservice broadband networks [1] (e.g. Broadband-ISDN and ATM based LANs). This is prompted by the need to handle a variety of types of services, with diverse demands on the network in terms of the required bit rate. Continuous as well as variable bit rates will be serviced, e.g. data, voice, still and moving pictures, and multimedia. The bandwidth required by a connection may vary over the lifetime of the connection, hence multiplexing and buffering within the network should be provided in order to allow more effective use of resources. Thus the need to allocate bandwidth efficiently, within the constraints set by the underlying facility network [2]. Additionally, we expect that by combining the control of bandwidth with the control of flow rate, we can achieve a better network utilisation. Better network utilisation can be achieved by controlling temporary variations in the traffic flow rate which would in certain cases (e.g. a full buffer and heavy traffic flow), if not stored in a temporary local buffer, result in cell loss (and subsequent retransmissions, which as a result would reduce the network utilisation).

Whilst it is possible to provide controls at different levels in the network, such as cell level, burst level and the connection (i.e. call) level [3], we will only consider the control of traffic flow and bandwidth allocation to groups of connections sharing a common route. The virtual path (VP) [4, 5] is an important aspect of current ITU (International Telecommunications Union) recommendations on Broadband ISDN and can be viewed as a pre-established route through the network with dedicated bandwidth onto which virtual channels (VCs) are grouped. VPs considerably simplify network management and call admission by semipermanently allocating resources to the path [4]. Each network link will have a finite capacity (bandwidth or service rate) which must be allocated among the VPs and VCs in a fair and efficient manner while meeting the specific requirements of each type of traffic. The problem of bandwidth allocation to VPs has been studied in, e.g. [6, 7, 8, 9]. Flow rate control has also been extensively studied, by among others [10, 11, 12, 13].

In this paper the combined problem of flow rate and bandwidth control in communication networks under nonstationary conditions is considered. Communication networks not only must have good steady-state performance but also must deliver acceptable performance under nonstationary and transient conditions [14], [15]. Nonstationary conditions occur in communication networks when the statistics of the arrival or service processes to the network queues vary with time. Nonstationary conditions can, for example, be caused by nonstationary input loads, topological changes to the network, and failures of network resources. This nonstationary behaviour is particularly significant in the context of multimedia and multiservice networks because of the mix of traffic types and the nature of resource sharing. Our proposed nonlinear control solution is in the form of a dynamic feedback controller for bandwidth and flow rate.

In Section 2 we adopt a nonlinear state variable model to describe the behaviour of a virtual path, for different classes of traffic, in terms of time-varying mean quantities. We formulate the nonlinear control problem in section 3 for the case of a single node VP, and obtain the combined flow control and bandwidth allocation control strategy in closed form. Simulation results are presented in Section 4. In section 5 we present our concluding comments.

2. A State Model for Virtual Paths

Figure 1 shows an N×N nonblocking output buffered switch which can route cells from a set of incoming links to a set of outgoing links. The VP of Figure 2 focuses on the series of M switches along one particular VP, showing the buffers and servers of the links used by the VP traffic. The VP is modeled as a one-way connection between an Origin-Destination (OD) pair spanning several ATM switching nodes (see Figures 1 and 2). The nodes are connected by links $(1,...,M)$, associated with deterministic propagation delays τ_i $(i = 1,...,M)$. Several VCs $(\lambda_1^1,...,\lambda_1^S)$ combine to form the VP traffic, and local VP and/or VC traffic $(\lambda_1^b,...,\lambda_M^b)$ appears as background traffic, requiring some access to the shared resources of each link. The queue at each link provides for statistical multiplexing of the incoming traffic streams.

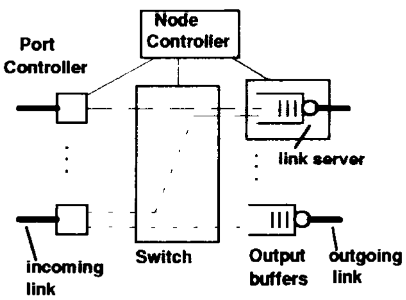


Figure 1. ATM switching node

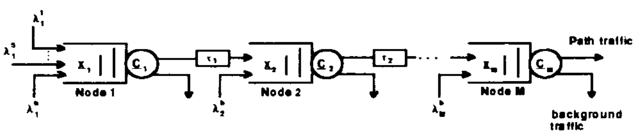


Figure 2. Virtual path spanning M ATM switching nodes

A dynamic model is required for the VP, in a form suitable for a distributed control solution. Our objective is to find a model which captures the essential dynamic behaviour, but has low order complexity relative to detailed probabilistic models such as the Chapman-Kolmogorov equation for determining the time-dependent state probability distribution for a Markovian queue, [16]. Hence, the approximate fluid flow modeling approach described in [16, 17] is adopted.

Using the flow conservation principle, for a single queue, assuming no losses, the rate of change of the average number of cells queued at the link buffer can be related to the rate of cell arrivals and departures by a differential equation of the form:

$$\dot{x}(t) = -f_{out}(t) + f_{in}(t) \qquad (1)$$

where:

$x(t)$ - state of the queue, given by the ensemble average of the number of cells $N(t)$ in the system (i.e. queue+server) at time t, i.e. $x(t)=E\{N(t)\}$

$f_{out}(t)$ - ensemble average of cell flow out of the queue at time t

$f_{in}(t)$ - ensemble average of cell flow into the queue at time t

The fluid flow equation is quite general and can model a wide range of queueing and contention systems as shown in the literature [16, 17, 18].

Assuming that the queue storage capacity is unlimited and the customers arrive at the queue with rate $\lambda(t)$, then $f_{in}(t)$ is just the offered load rate $\lambda(t)$ since no packets are dropped. The flow out of the system, $f_{out}(t)$, can be related to the ensemble average utilisation of the link $\rho(t)$ by $f_{out}(t)=C(t)\rho(t)$, where $C(t)$ is defined as the capacity of queue server. We assume that $\rho(t)$ can be approximated by a function $G(x(t))$ which represents the ensemble average utilisation of the queue at time t as a function of the state variable. Thus, the dynamics of the single queue can be represented by a nonlinear differential equation:

$$\dot{x}(t) = -G(x(t))C(t) + \lambda(t), \quad x(0) = x_o \qquad (2)$$

A commonly used approach to determine $G(x)$ is to match the steady-state equilibrium point of (2) with that of an equivalent queueing theory model where the meaning of "equivalent" depends on the queueing discipline assumed. This method has been validated with simulation by a number of researchers, for different queuing models [16, 17, 18]. System identification techniques can also be used to identify the parameters of the fluid flow equation.

We illustrate the derivation of the state equation for an M/M/1 queue following [16]. The following standard assumptions are made: the packets arrive according to a Poisson process; packet transmission time is proportional to the packet length; and that the packets are exponentially distributed with mean length $1/\mu$.

Then, from the M/M/1 queueing formulas, for a constant arrival rate to the queue the average number in the system at steady state is $\lambda/(\mu C-\lambda)$. Thus requiring that $x(t)=\lambda/(\mu C-\lambda)$ when $\dot{x} = 0$, the state model becomes

$$\dot{x}(t) = -\frac{x(t)}{1+x(t)}\mu C(t) + \lambda(t), \quad x(0) = x_o \qquad (3)$$

Using fluid flow arguments, we can represent a VP as a series of M/M/1 queues. The validity of using an M/M/1 queue to approximately describe the queueing delays with fixed packet length is discussed by Gerla et al [7]. With reference to Figure 2, the following standard assumptions are made: cells at any node along the VP arrive as independent Poisson processes; cells have exponentially distributed lengths; service time is exponentially distributed with mean $1/C_i$, where C_i is the link server capacity at node i; and buffer length at each node is assumed infinite. We assume that the link has a First-In-First-Out

(FIFO) service discipline and a common (shared) buffer. The VP model, an extension of the M/M/1 queue model (1), is:

$$\dot{x}_1(t) = -C_1(t)\left(\frac{x_1(t)}{1+x_1(t)}\right) + \lambda^{vp}(t) + \lambda_1^{background}(t)$$

$$\dot{x}_i(t) = -C_i(t)\left(\frac{x_i(t)}{1+x_i(t)}\right) + \gamma_i^{v}(t) + \lambda_i^{background}(t)$$

$$i = 2,\ldots,M. \qquad (4)$$

where

$C_i(t)$ - bandwidth (capacity, cell service rate) allocated to the VP at node i,

$x_i(t)$ - state of the queue at node i,

$\lambda^{vp}(t)$ - total arrival rate due to VP traffic

$\lambda_i^{b}(t)$ - total arrival rate at cell-queue i due to background traffic,

$\gamma_i^{v}(t)$ - VP traffic, leaving the previous node $i-1$ and entering node i, delayed by a deterministic amount τ_{i-1} due to the transmission propagation.

For $i = 2,\ldots,M-1$

$$\gamma_i^{v}(t) = \left(C_{i-1}(t-\tau_{i-1})\frac{x_{i-1}^{vp}(t-\tau_{i-1})}{1+x_{i-1}(t-\tau_{i-1})}\right).$$

where $x_i^{vp}(t)$ is the ensemble average of the number of cell places in the buffer occupied by VP traffic.

This model has been shown to be reasonably accurate in comparison to discrete event simulation.

This model (4) can be used to represent all possible paths for any origin destination pair.

3. Nonlinear Control of Flow Rate and Bandwidth Allocation

Our control philosophy, shown in Figure 3, is to exert influence on both the flow rate into the server, as well as the allocation of the (cell) server rate, based on feedback from the network state (queue length). Flow control is exercised on the output of the source, which is shown here as a local buffer. Any unmet demanded inflow is stored in this local buffer, and paced into the network, whenever the resources become available.

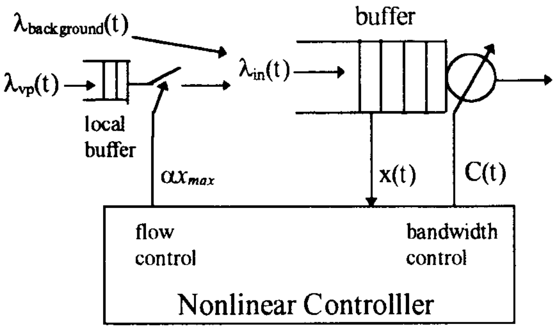


Figure 3. Control concept

To highlight the approach, we only consider a single node with VP traffic and derive the proposed nonlinear control strategy. Extensions to multiple node VPs, with interfering background traffic will be considered in a later study.

We consider two cases:

<u>a) Zero background traffic</u>

In Section II, we derive the dynamic equation for a single node (Equation 3) for the case of infinite buffer places and infinite server capacity. Equation (3) is rewritten here for the case of finite buffer and finite server.

$$\dot{x}(t) = -\frac{x(t)}{1+x(t)}C(t) + \lambda_{in}(t) \quad x(t_o) = x_o$$

$$0 \le x(t) \le x_{buffer\ capacity} \qquad (5)$$

$$0 \le C(t) \le C_{server\ capacity}.$$

Let

$$\bar{x}(t) = x(t) - x_{max}$$

where x_{max} is a design variable, and we select

$x_{max} \le x_{buffer\ capacity}$ and $x_{max} \ge x_o$. (6)

Then from (5) we obtain

$$\dot{\bar{x}}(t) = -\frac{(\bar{x}(t) + x_{max})}{1+\bar{x}(t)+x_{max}}C(t) + \lambda_{in}(t). \qquad (7)$$

We select

$$C(t) = \alpha\,(1+\bar{x}(t)+x_{max}) = \alpha\,(1+x(t)) \qquad (8)$$

where α is a design variable (discussed later), and

$$\lambda_{in}(t) = \alpha\, x_{max}. \qquad (9)$$

where for clarity of exposition we assume that $\lambda_{vp}(t) \ge \alpha x_{max}\ \forall t$, i.e the local buffer is saturated (this condition is relaxed later).

With these choices of $C(t)$ and $\lambda_{in}(t)$ equation (7) becomes

$$\dot{\bar{x}}(t) = -\alpha\,\bar{x}(t) \qquad (10)$$

whose solution is given by

$\bar{x}(t) = e^{-\alpha(t-t_0)}\bar{x}(t_0)$ for all $t \ge t_o$

and

$$x(t) = x_{max} + e^{-\alpha(t-t_0)}(x(t_0) - x_{max}). \qquad (11)$$

Since $0 \le x(t_o) \le x_{max}$, we have $0 \le x(t) \le x_{max}$ for all $t \ge t_o$.

That is $x(t)$ is always upper bounded by x_{max}, and therefore by $x_{buffer\ capacity}$ as long as $x(t_0) \le x_{max}$.

The bandwidth allocation control $C(t)$ given by (8) needs to satisfy the constraints $0 \le C(t) \le C_{server\ capacity}$. Consider $C_{max} \le C_{server\ capacity}$ to be a design parameter to be chosen together with the parameter α so that $C(t)$ satisfies the constraints.

We select α as

$$0 \le \alpha \le \frac{C_{max}}{1+x_{max}} \qquad (12)$$

We now relax the earlier assumption that $\lambda_{vp}(t) \geq \alpha x_{max}$ $\forall t$. In a realistic scenario, $\lambda_{vp}(t)$ can take any value (up to the declared peak rate, which is often greater than the average rate). However the theory presented (see (9)) dictates that $\lambda_{in}(t)$ never exceeds αx_{max}. As already discussed, any unmet demanded inflow is queued at the local buffer, and paced into the network whenever the resources become available. Therefore for efficient operation of the system we must also ensure, if possible, that the demanded cell inflow is met (otherwise the local buffer length will increase toward infinity). In order to achieve this we must increase the lower limit on the control design variable α from 0 to $\frac{C_{aver}}{1+x_{\max}} = \frac{\lambda_{aver}}{x_{\max}}$ where

$C_{aver} = \frac{\lambda_{aver}(1+x_{\max})}{x_{\max}}$. That is

$$\frac{\lambda_{aver}}{x_{\max}} \leq \alpha \leq \frac{C_{\max}}{1+x_{\max}} \tag{13}$$

For setting this limit it is necessary to know the average demanded cell inflow (λ_{aver}). This can either be provided by a suitably designed Connection Admission Control (CAC) strategy, or measured.

Remark: By recalling Little's law [19], i.e. $E(n)=\lambda E(T)$, where $E(n)$ is the average queue length and $E(T)$ is the average time delay, the lower limit on α ($\alpha_{min} = \frac{C_{aver}}{1+x_{\max}} = \frac{\lambda_{aver}}{x_{\max}}$) can be interpreted as a requirement on the waiting time in the system. ∇∇∇

Since

$$C(t) = \alpha\,(1+x(t)) \leq \alpha\,(1+x_{\max})$$

it follows that

$$C(t) \leq \alpha\,(1+x_{\max}) \leq C_{\max} \leq C_{server\ capacity}.$$

The summarised control strategy is as follows:

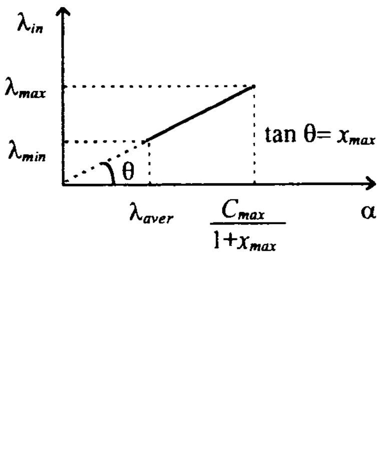
$$C(t) = \alpha\,(1+x(t)), \qquad \lambda_{in}(t) \leq \alpha x_{\max}$$

where α, x_{max}, C_{max} are design parameters chosen as

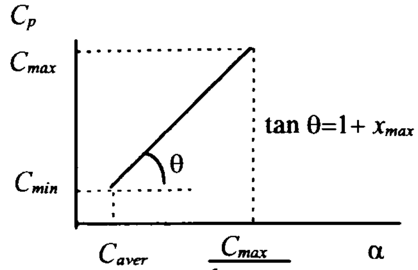
$$\frac{\lambda_{aver}}{x_{\max}} \leq \alpha \leq \frac{C_{\max}}{1+x_{\max}}, \qquad 0 < x_{max} \leq x_{buffer\ capacity}$$

$$x(t_o) \leq x_{\max}, \qquad 0 < C_{max} \leq C_{server\ capacity}.$$

The figures below show (for the case of $\lambda_{vp}(t) \geq \alpha x_{max}$ $\forall t$) the plots of: $\lambda_{in}(t)$ versus α;



and the upper bound C_p for $C(t)$, i.e. $C_p = \alpha\,(1+x_{\max})$ versus α



It is clear that the maximum possible value of $\lambda_{in}(t)$ is $\frac{C_{\max}}{1+x_{\max}} x_{\max}$.

<u>b) Non zero background traffic</u>

In the case of nonzero background traffic instead of equation (5) we have

$$\begin{gathered}\dot{x}(t) = -\frac{x(t)}{1+x(t)}C(t) + \lambda_{vp}(t) + \lambda_{background}(t)\\ x(t_o) = x_o\\ 0 \leq x(t) \leq x_{buffer\ capacity}\\ 0 \leq C(t) \leq C_{server\ capacity}\end{gathered} \tag{14}$$

which may be written in the form

$$\dot{\bar{x}}(t) = -\frac{(\bar{x}(t)+x_{\max})}{1+\bar{x}(t)+x_{\max}}C(t) + \lambda_{vp}(t) + \lambda_{background}(t)$$

where $\bar{x}(t) = x(t) - x_{\max}$. In this case we choose

$$C(t) = \alpha\,(1+\bar{x}(t)+x_{\max}) = \alpha\,(1+x(t))$$

to obtain

$$\dot{\bar{x}}(t) = -\alpha\,\bar{x}(t) - \alpha x_{\max} + \lambda_{vp}(t) + \lambda_{background}(t).$$

Two possibilities exist: measurable or unmeasurable background traffic $\lambda_{background}(t)$.

In the case of measurable $\lambda_{background}(t)$ the strategy remains as presented for case a), but with $\lambda_{vp}(t) = \alpha\,x_{\max} - \lambda_{background}(t)$.

The case of unmeasurable $\lambda_{background}(t)$ traffic is treated next.

We choose (again, as for the case without any background traffic, we assume $\lambda_{vp}(t) \geq \alpha x_{max}$ $\forall t$)

$$\lambda_{vp}(t) = \alpha\,x_{\max} \tag{15}$$

to obtain

$$\dot{\bar{x}}(t) = -\alpha\,\bar{x}(t) + \lambda_{background}(t)$$

or

$$\bar{x}(t) = e^{-\alpha(t-t_0)}\bar{x}(t_0) + \int_{t_o}^{t} e^{-\alpha(t-\tau)}\lambda_{background}(\tau)d\tau$$

If $\lambda_{background}^{max}$ is an upper bound for $\lambda_{background}(t)$ that is known then

$$\bar{x}(t) \le e^{-\alpha(t-t_0)}\bar{x}(t_0) + \frac{\lambda^{max}_{background}}{\alpha}\left(1 - e^{-\alpha(t-t_0)}\right)$$

$$\bar{x}(t) \le \frac{\lambda^{max}_{background}}{\alpha} + e^{-\alpha(t-t_0)}\left(\bar{x}(t_0) - \frac{\lambda^{max}_{background}}{\alpha}\right)$$

or

$$x(t) \le x_{max} + \frac{\lambda^{max}_{background}}{\alpha} + e^{-\alpha(t-t_0)}\left(x(t_0) - \left(x_{max} + \frac{\lambda^{max}_{background}}{\alpha}\right)\right)$$

If we choose $x(t_o)$, x_{max}, α such that

$$x(t_o) \le x_{max} + \frac{\lambda^{max}_{background}}{\alpha} \le x_{buffer\,capacity} \tag{16}$$

then $x(t) \le x_{buffer\,capacity}$ for all $t \ge t_o$.

We consider

$C(t) = \alpha\,(1 + x(t)) \le \alpha\,(1 + x_{buffer\,capacity})$.

If we choose α so that

$$0 < \alpha \le \frac{C_{max}}{1 + x_{buffer\,capacity}} \tag{17}$$

where $0 \le C_{max} \le C_{server\,capacity}$

then $C(t) \le C_{max}$.

The variable α has to satisfy both (16) and (17). From (16) it follows that α should satisfy the inequality

$$\alpha \ge \frac{\lambda^{max}_{background}}{x_{buffer\,capacity} - x_{max}} \tag{18}$$

In this case $x_{max} < x_{buffer\,capacity}$. Combining with (17) we have that α should satisfy

$$\frac{\lambda^{max}_{background}}{x_{buffer\,capacity} - x_{max}} \le \alpha \le \frac{C_{max}}{1 + x_{buffer\,capacity}} \tag{19}$$

Note that, as for the case without any background traffic, for efficient operation of the system we must also ensure, if possible, that the demanded cell inflow is met. In order to achieve this we must set the lower limit α_{min} on the control design variable α to $\alpha_{min} = \max\left(\frac{\lambda_{aver}}{x_{max}}, \frac{\lambda^{max}_{background}}{x_{buffer\,capacity} - x_{max}}\right)$. Therefore α should satisfy

$$\alpha_{min} \le \alpha \le \frac{C_{max}}{1 + x_{buffer\,capacity}} \tag{20}$$

Once again the value of λ_{aver} (now total average, i.e. sum of average of VP and background traffic) can be provided by CAC, or measured.

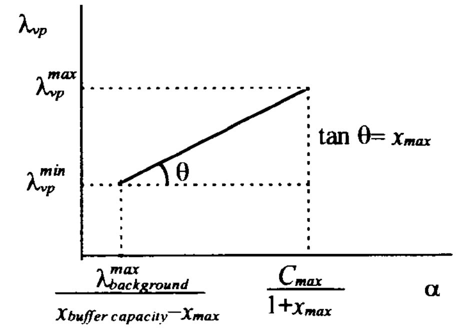
The satisfaction of (20) depends (greatly) on the values of x_{max}, $\lambda^{max}_{background}$ and $C_{max} \le C_{server\,capacity}$.

If $\lambda^{max}_{background} > \frac{C_{server\,capacity}\left(x_{buffer\,capacity} - x_{max}\right)}{1 + x_{buffer\,capacity}}$ it is clear that no α will exist to satisfy (19). In such case the background traffic is so large that no control strategy exists to control the traffic within the desired constraints.

The plot of λ_{vp} versus α when

$$\frac{\lambda^{max}_{background}}{x_{buffer\,capacity} - x_{max}} << \frac{C_{max}}{1 + x_{buffer\,capacity}}$$

is given next.



The plot shows that in the presence of background traffic λ_{vp} has to exceed a certain minimum value λ^{min}_{vp} for the control strategy to be effective. Observe that the maximum value λ^{max}_{vp} of λ_{vp} becomes smaller as $\lambda^{max}_{background}$ increases.

4. Simulation:

We consider a single node with one VP. The allocated service rate cannot exceed 10 cells/msec (i.e. C_{max}=10), as the rest of the server capacity (for 155 Mbit/sec the service rate is $C_{server\ capacity}$=365 cells/msec) has been reserved for the rest of the VPs and VCs. To simulate a small buffer size, we assume that 100 cell places have been reserved for this VP (i.e. x_{max}=100).

Case I) Simulations with zero background traffic

In this set of simulation runs we vary the value of the control design variable α, as shown in the table below:

value (i)	**value (ii)**
$\alpha = \frac{12}{1+x_{\max}} < \frac{C_{\max}}{1+x_{\max}}$	$\alpha = \frac{10}{1+x_{\max}} = \frac{C_{\max}}{1+x_{\max}}$
value (iii)	**value (vi)**
$\alpha = \frac{8.08}{1+x_{\max}} = \frac{C_{aver}}{1+x_{\max}}$	$\alpha = \frac{5}{1+x_{\max}} < \frac{C_{aver}}{1+x_{\max}}$

The values of the control design variable α (see equation 13) have been chosen so that:

value (i) violates the upper limit;

value (ii) equals the upper limit;

value (iii) equals the minimum value which ensures that the demanded inflow is met; and value (iv) is below the minimum value which ensures that the demanded inflow is met.

Figure 4 shows the cell inflow into the queue, for the four different values of α as well as the demanded cell inflow (prior to any shaping by the controls). We can see that as α decreases the amount of shaping (or smoothing or control) that is applied to the cell inflow increases. Note that value (iv) also limits the amount of traffic entering the network to below that demanded (that is excessive control is applied to the flow into the system). This can also be seen by the local buffer state $x_{local}(t)$, in Figure 7, where the local buffer queue starts to build up to infinity (i.e. the demanded cell inflow has not been met, and the excess traffic is queued at the local buffer prior to it entering the network). The buffer state is shown in Figure 5, and the allocation of service rate in Figure 6. Again, we can observe the influence of the control design variable α on the controlled system performance. As α decreases we move from not enough control, for value (iv), to excessive control, for value (i). Finally, in Figure 8 the cummulative loss $n(t)$ is shown. Only value (i), as expected, has any loss. This is due to the violation of the upper limit on α.

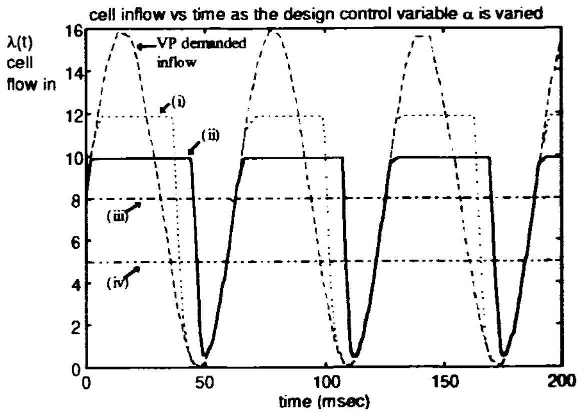


Figure 4. Cell flow into the buffer

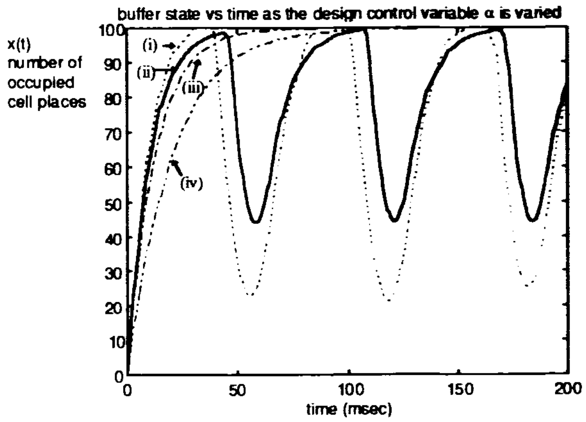


Figure 5. Buffer state

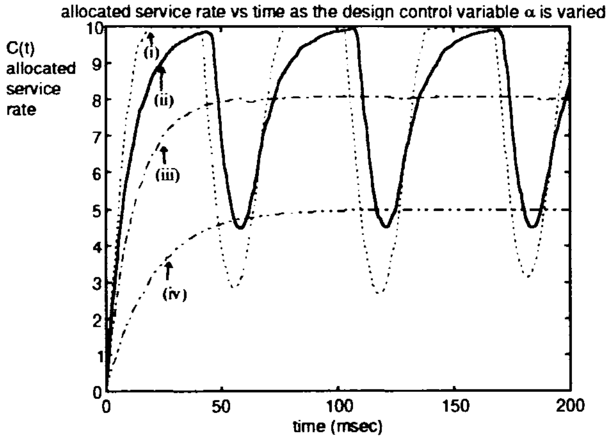


Figure 6. Allocated service rate

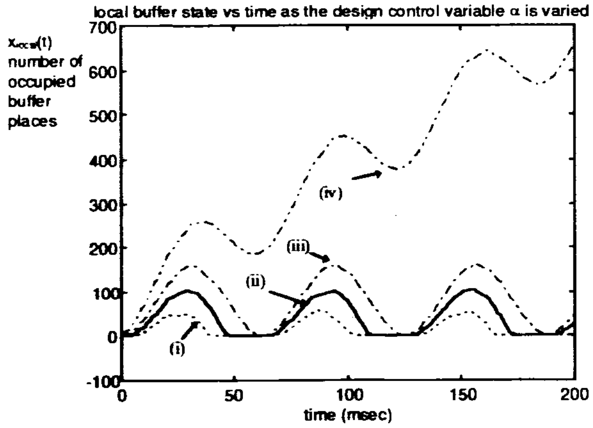


Figure 7. Local buffer state

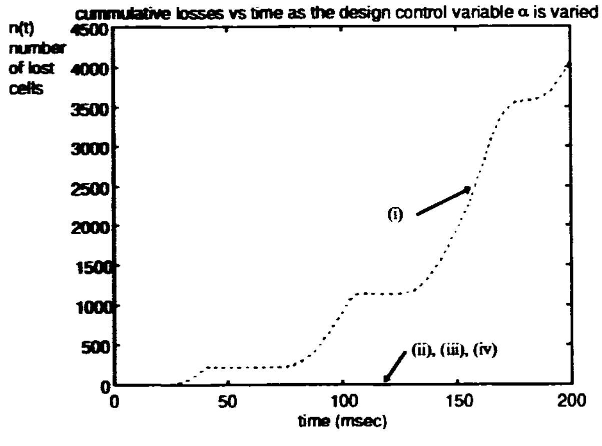


Figure 8. Buffer losses

Case II) Simulations with background traffic

In this set of simulation runs we increase the buffer length to $x_{buffer\ length}$=140 cell places, but keep x_{max} the same as in case I). All other simulation parameters are as in case I). We vary the value of the control design variable α, as shown in the table below:

value (i)	value (ii)	value (iii)
$\alpha=\dfrac{10}{1+x_{\max}}=\dfrac{C_{\max}}{1+x_{\max}}$	$\alpha=\dfrac{8}{x_{max}}=\dfrac{\lambda_{aver}}{x_{max}}=\alpha_{\min}$	$\alpha=\dfrac{2}{x_{buffer\ capacity}-x_{max}}=\dfrac{\lambda^{max}_{background}}{x_{buffer\ capacity}-x_{max}}<\alpha_{\min}$

The values of the control design variable α (see equation 20) have been chosen so that:

value (i) equals the upper limit;

value (ii) equals the minimum value which ensures that the demanded inflow is met; and

value (iii) equals the minimum value which ensures that the control strategy is effective.

The results show similar behaviour as for case I). Figure 9 shows the cell inflow into the queue, for the three different values of α, as well as the demanded total cell inflow (prior to any shaping by the controls), and the VP demanded cell inflow. The demanded background traffic is kept constant at 2 cells/msec. Again, we observe that as α decreases the amount of shaping (or smoothing or control) that is applied to the cell inflow increases. Value (iii) also limits the amount of traffic entering the network to below that demanded (that is excessive control is applied to the flow into the system), which is confirmed by the local buffer state, in Figure 12. For value iii) the local buffer queue starts to build up to infinity (i.e. the demanded cell inflow has not been met, and the excess traffic is queued at the local buffer prior to it entering the network), whereas for values i) and ii) the local queue is used to pace the traffic into the network. The buffer state is shown in Figure 10, and the allocation of service rate in Figure 11, where we can observe the influence of the control design variable α on the controlled system performance. Since the limits on α have not been violated, as expected, no losses are observed for all three values.

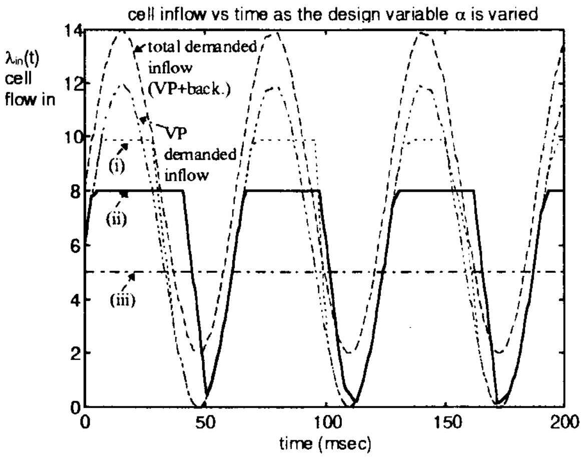


Figure 9. Cell flow into the buffer (with non zero background traffic)

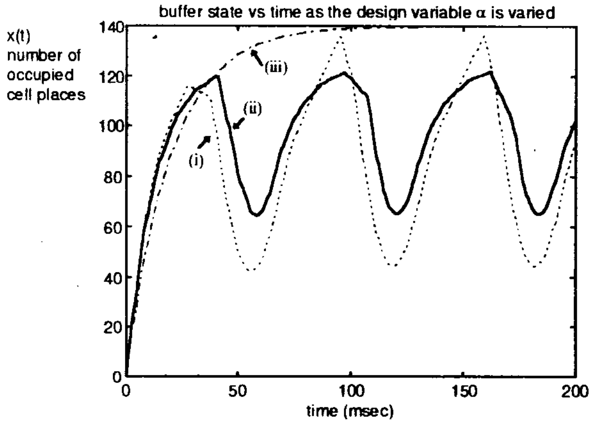


Figure 10. Buffer state (with non zero background traffic)

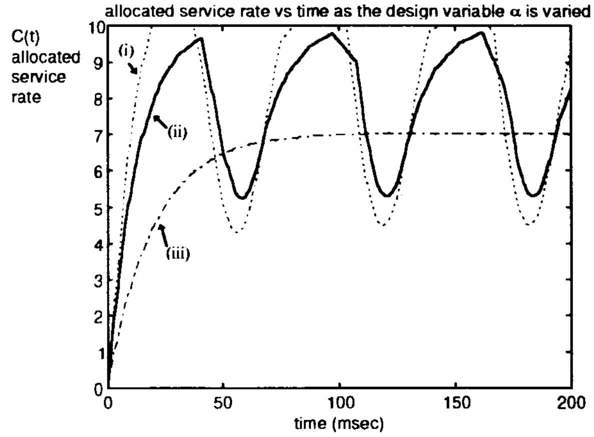


Figure 11. Allocated service rate (with non zero background traffic)

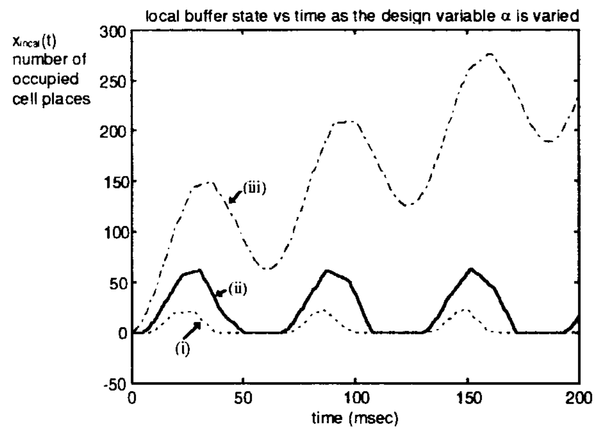


Figure 12. Local buffer state (with non zero background traffic)

5. Conclusions

This paper focuses on the development of a combined dynamic nonlinear control strategy for flow control and bandwidth allocation control for virtual paths under nonstationary network conditions. A state variable model for describing the dynamic behaviour of virtual paths is presented and used to formulate the dynamic nonlinear control strategy within a nonlinear control theoretic framework. Finite constraints on the buffer and server capacity are explicitly handled in the problem solution. The behaviour of the nonlinear control solution is demonstrated using simulation. It is shown that by appropriate choice of the control design variables effective server and buffer utilisation can be achieved (hence avoid cell losses and subsequent retransmissions). This paper suggests that nonlinear control techniques can be usefully employed in the solution of telecommunication problems. Future extension include the case of multiple nodes in a VP. This will necessitate the derivation of a distributed nonlinear control strategy.

References:

[1] CCITT: Recommendation I.121. "Broadband Aspects of ISDN". Geneva, 1991.

[2] R. Warfield, R. Harris, S. Mischnowicz, "Some problems of network management in a fast packet switching network", in Proc. 3rd Fast Packet Workshop, Melbourne, 1988.

[3] J. Filipiak, "M-architecture: A structural model of traffic management and control in broadband ISDNs", IEEE Comm. Mag. Vol. 27, No. 5, pp.25-32, 1989.

[4] J. Burgin, D. Dorman, "Broadband ISDN Resource Management: The Role of Virtual Paths", IEEE ,Comm. Mag. pp44-48, Sept. 1991.

[5] K. Sato, S. Ohta, I. Tokizawa, "Broad-Band ATM Network Architecture Based on Virtual Paths", IEEE Trans. on Comm., Vol. 38, N0. 8, pp1212-1222, Aug. 1990.

[6] J. Burgin, "Dynamic Capacity Management in the BISDN", International Journal of Digital and Analog Communications Systems, Vol. 3, pp. 161-165, 1990.

[7] M. Gerla, J. A. S. Monteiro, R. Pazos, "Topology Design and Bandwidth Allocation in ATM nets", IEEE JSAC, Vol. 7, No. 8, pp. 1253-1262, Oct. 1989.

[8] R. Bolla, et al, "Adaptive bandwidth allocation by hierarchical control of multiple ATM traffic classes", IEEE INFOCOM'92, Florence, May 1992.

[9] A. Pitsillides, J. Lambert, D. Tipper, "A multilevel optimal control approach to dynamic bandwidth allocation in Broadband-ISDN", to appear in the special issue on B-ISDN in the Telecommunication Systems Journal.

[10] A. Pitsillides, J. Lambert, "Adaptive connection admission and flow control: quality of service with high utilisation", IEEE INFOCOM'94, Toronto, Ontario, Canada, June 1994.

[11] P. Newman, "Backward explicit congestion notification for ATM local area networks", Globecom'93, pp.719-723, Houston, December 1993.

[12] A. Pitsillides, A. Sekercioglu, G. Ramamurthy, "Fuzzy backward congestion notification (FBCN) congestion control in asynhronous transfer mode (ATM) networks", submitted to IEEE GLOBECOM 95, Singapore, November 13-17, 1995.

[13] J-C. Bolot, A. Udaya Shankar, "Dynamical behaviour of rate-based flow control mechanisms", Computer Communication Review, pp. 35-49, April 1990.

[14] H. Van As, "Transient analysis of Marcovian queuing systems and its application to congestion control modelling", IEEE JSAC, Vol. 4, No. 6, pp . 891-904, Sept. 1986.

[15] W. Lovegrove, J. L. Hammond, D. Tipper, "Simulation Methods for Studying Nonstationary Behavior of Computer Networks", IEEE JSAC, Vol. 8, No. 9, pp. 1696-1708, Dec. 1990.

[16] D. Tipper, M. K. Sundareshan, "Numerical Methods for modelling Computer Networks Under Nonstationary Conditions", IEEE JSAC, Dec. 1990.

[17] S. Sharma, D. Tipper, "Approximate models for the study of nonstationary queues and their application to communication networks", IEEE ICC'93, May 1993.

[18] J. Filipiak, "Modelling and Control of Dynamic Flows in Communication Networks", Springer-Verlag, 1988.

[19] M. Schwartz, "Telecommunication Networks: Protocols, Modelling and Analysis", Addison Wesley, 1988.