DEAD-TIME COMPENSATION FOR ABR TRAFFIC CONTROL OVER ATM NETWORKS

F. Gómez-Stern, J.M. Fornés, and F.R. Rubio
Dept. Ingeniería de Sistemas y Automática. Escuela Superior de Ingenieros. Universidad de Sevilla. Camino de los Descubrimientos s/n. 41092-Sevilla. SPAIN
e-mail: {fgs,fornes,rubio}@cartuja.us.es

Abstract

This paper presents the design of a rate-based controller for throttling Available Bit Rate (ABR) input rates in high speed Asynchronous Transfer Mode (ATM) networks with significant propagation delays. First, a classical control scheme based on the Smith predictor is analyzed in order to point out its drawbacks in terms of performance and stability. Saturation issues are handled with anti-windup techniques, widely used in industrial PID controllers. Moreover, performance is notably improved by means of the feedback of an estimation of the ABR disturbance, which remains unknown to the switch. This reduces the average queue level drastically, thus guaranteeing the shortest delays possible while keeping the channel fully occupied. Finally, sensitivity to delay estimation errors is analysed, and the limitations of the proposed controller are discussed with respect to ABR traffic parameters.

List of Acronyms

ABR	Available Bit Rate
ACR	Allowed Cell Rate
ATM	Asynchronous Transfer Mode
CCR	Current Cell Rate
CDV	Cell Delay Variation
CLR	Cell Loss Rate
CTD	Cell Transfer Delay
ER	Explicit Rate
FIFO	First In First Out queue policy
ICR	Initial Cell Rate
MCR	Minimum Cell Rate
NRM	Network Resource Management
PCR	Peak Cell Rate
PID	Proportional-Integral-Derivative linear controller
QoS	Quality of Service
RM	Resource Management
VBR	Variable Bit Rate
VC	Virtual Channel

1 Introduction

In ATM Networks, ABR service is provided to make use of the available bandwidth left unused by other co-existing types of traffic. Traffic sources expect fair bandwidth assignment among all ABR connections sharing the link. ABR is thus non-deterministic, and end-users are notified of the available capacity by tagging Resource Management (RM) cells at rate 1/NRM cells per second. ABR bandwidth can be modeled as a disturbance to the system, and delays appear due to the time deployed by RM cells traversing the links.

In high speed communication networks, long propagation delays are critical for the stability of closed-loop congestion control algorithms. Smith's principle is a key tool to design a feedback control law for congestion avoidance and high speed ATM networks. Many algorithms dealing with congestion control using Smith's predictor have been proposed (see [1], [2], [3] and [4] for arguments for using the Smith predictor). Due to large delays inside the feedback loop, queue level dynamics might exhibit oscillations. Since the model of the communicaction system is known without parameter uncertainty, a controller can be designed following Smith's principle. It is worth noting that, in wide area networks, roundtrip delays are mostly determined by propagation delay. To include in the analysis the jitter of roundtrip time due to queueing time, a model containing different delays could be considered.

A complete fluid-flow model considering transport delays is proposed in the literature [1], where input-output stability is guaranteed when using a proportional Smith predictor based controller. The controller performance is improved by adding an integral part to it, thus becoming a linear PI that can better reject disturbances [2]. This model is not complete as it lacks saturations in the control action.

Regarding the control strategy, a fundamental aspect must be revised: the queue, modeled as an integrator, is an unstable system in the disturbance-output transfer function and, therefore, it will not be eliminated by a Smith predictor [10]. This implies a direct dependence between the steady-state queue level and the real ABR bandwidth. As a consequence, the higher the occupancy level of the queue, the greater the delays.

This paper is organized as follows: First, a state-of-the art in ATM traffic management is included and ABR service on ATM is defined. Section 3 illustrates the system model and settles the control objective. Section 4 begins with a basic proportional controller inside a Smith predictor designed to handle transmission delays. A discussion on the sensitivity of the output under different ABR conditions leads, through various improvements, to a final PI controller with feedforward and anti-windup mechanisms which appears to be the best choice for this type of systems. Further simulation studies illustrate the efficiency of the proposed controller. The results section analyses the advantages and drawbacks of these structures and points out the most significant parameters to evaluate the output. A final conclusive analysis and further suggestions are included in section 6.

2 Available Bit Rate traffic

ABR is an ATM layer service category defined by The ATM Forum [11] that may be used by applications which expect cell loss guarantees but can control their data dynamically as demanded by the network.

In ABR there is a feedback mechanism to control the source rate in response to changing ATM layer transfer characteristics. An end-system that adapts its traffic in accordance with the feedback is expected to experience a low Cell Loss Ratio (CLR), although a quantitative value for CLR is network specific. This feedback is supported by specific control cells called Resource Management Cells, or RM-cells.

The ABR source sends data at a rate less or equal than the Allowed Cell Rate (ACR), which in turn is smaller than a negotiated Peak Cell Rate (PCR) and greater than a negotiated Minimum Cell Rate (MCR). Immediately after establishing a connection and after an idle time-out , ACR is set to an Initial Cell Rate (ICR), which is also negotiated with the network. The source sends an RM-cell after Nrm-1 cells, where Nrm is a parameter. Among the RM cell fields, the Current Cell Rate (CCR) field informs the network about the source's ACR, and the Explicit Rate (ER) field is used by the network to give its rate feedback. The destination returns RM cells back to the source.

The ABR framework is predominantly closed-loop, i.e., sources normally change their rates in response to network feedback. Open-loop control complements closed-loop control when the network delays are long compared to application traffic chunks, or when network feedback is temporarily disrupted.

Although no numeric commitment is made regarding cell transfer delays (see section (2.3.1) of [11]), network providers are expected to advertise delay bounds for the class of ABR as a whole [6].

CTD (Cell Transfer Delay) and CDV (Cell Delay Variation) are agreed on for a VBR source whereas for an ABR source only MCR is agreed on. However, if the agreed CTD and CDV are not demanding, they can be easily satisfied and there may be a possibility to improve the quality of the ABR service by changing the class service priority. If a VBR source is sent, producing a deterministic CTD, the ABR source does not need to wait until the VBR source queue is empty, its CTD and CDV become smaller [7].

The introduction of VBR background traffic makes ABR capacity variable. ABR switch schemes usually use the current demand and capacity to calculate feedback to the sources. The variable demand and variable capacity introduces variance in measurements made by switch schemes, and as a result, in the feedback given.

When the switch is overloaded the source takes feedback delay to respond to new feedback, and it is regarded to minimize its effect in following iterations between the source and the network.

3 System model

The system model considered here for control is the same as used by [5], with a FIFO queue on each switch. We do not need a measure of ABR to design the control law and hence it will be regarded as a disturbance ($d(t)$). We consider a deterministic fluid flow model to obtain the cell rate. The control signal $u(t)$ will be measured in cells/s and it is the ACR offered to ABR sources. The controller output is an indicator of the available capacity conceded to each source. This capacity is expressed as the bitrate allowed to be transmitted by the source.This information is transmitted within an ATM cell and is routed towards the origin of each source. Each virtual channel will receive the same $u(t)$ and hence fairness is guaranteed because the controller fills the cells with the same information for all sources. Nevertheless, each source can receive the cells at different times due to the different paths followed by the cells on their way to the origin. As a consequence not all sources will adapt their traffic simultaneously. This effect is modeled by a different delay for each source. It is not necessary to consider this delay as being random because ATM is connection oriented and the path is fixed for each connection. Each Virtual Channel (VC) will receive the same $u(t)$ and so fairness is assured. The level of the switch queue is represented by $x(t)$ and is measured in cells. This will be the level of the bottleneck queue that constrains the input rate for the VC. The set point $r(t)$ is an occupancy level of the queue and so $x(t) \leq r^o$, where r^o is the queue capacity, is the condition to prevent overflow in the queue. On the other hand, $x(t) > 0$ prevents queue underflow and so full link utilization.

The model presented serves as a starting point for the analysis, but it is insufficient for several important reasons. First, it ignores saturations in the control signal and the queue level. The control signal can never exceed the ATM limit of 155 Mbits per second. Moreover, it can never offer negative capacity values to the sources. This limitation becomes apparent

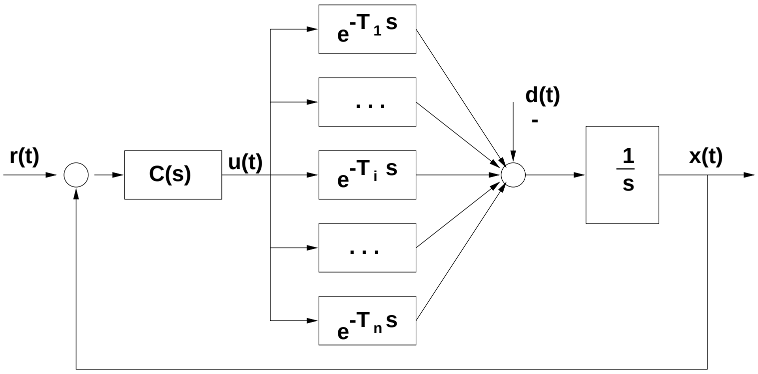

Figure 1: Control scheme for ABR

very easily: whenever the controller tries to compensate for a very high occupancy level in the queue, it will try to lower it by giving negative ABR values to the sources. On the other hand there are queue saturations. As an upper limit there is the available memory of the switch. Over a certain level, cells that cannot be stored are discarded. This could play in favor of the controller, preventing the output from growing to infinity, but cell-loss rate is a QoS parameter that has to be respected. The lower limit is the zero occupancy level, which must be properly modeled for simulations.

4 Control design

The aim of this section is to design a feedback controller that fulfills the following criteria:

- The average occupancy level of the queue must follow the desired set point.
- The set point should be low in order to reduce the size of the output buffer and, therefore, the response time. This parameter is critical in some applications as in voice communications or real-time video. A low set point will also permit oscillations without exceeding the queue capacity.
- As a compromise with the last objective, it is desirable that the queue level does not stay at zero for long intervals, which would mean that the channel is not being fully used. This implies that the set point should be adjusted (raised) taking into account the expected standard deviation of the buffer size.
- The bandwidth must be fairly assigned. This is achieved in the control scheme (1) by generating a unique control signal which is delivered to each traffic source.
- The delays of bandwidth notifications must be treated with dead-time compensation techniques to prevent oscillations.
- The minimum cell rate (MCR) should be considered in the cases where this ATM Forum specification parameter is required.
- The effect of transients in ABR capacity must be studied to view the limitations of the proposed controller.

The presence of time delays in the system can alter the behavior by introducing oscillations or even causing instability. A classic approach to this, which has proved to be robust is the Smith predictor structure [9]. In this type of controller, the delay is separated from the model so that the nominal plant is described as $P_n(s) = G_n(s)e^{-sT}$, $G_n(s)$ being the model without delay. Thus, the controller is designed to compensate $G_n(s)$, and the estimation error is also corrected by a feedforward block $P_n(s)$ with the delay. This scheme can be best understood in Fig.(2).

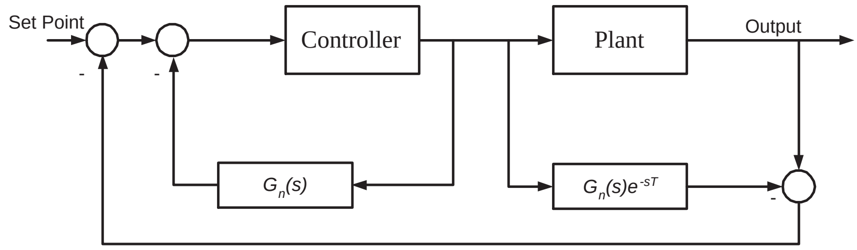

Figure 2: Smith predictor structure

In this figure, $P_n(s) = G_n(s)e^{-T_i s}$ is the nominal plant. The inner controller compensating $G_n(s)$ has to be determined. Re-arranging the structure, we can group this controller into only one block, as seen in Fig.(3).

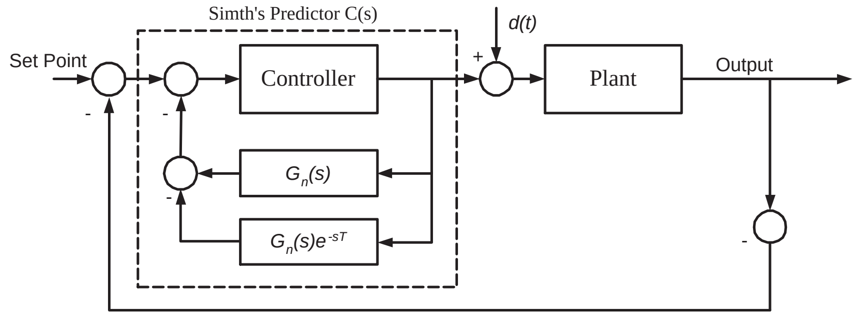

Figure 3: Equivalent Smith predictor structure with disturbance

This structure has the advantage of designing the controller as if there were no delays at all. Once the nominal plant is compensated using classical control techniques, we must observe the closed loop transfer function in order to analyze the behavior of the controller. Assuming a perfect model of the plant:

$$\frac{X(s)}{R(s)} = \frac{C(s)P_n(s)}{1 + C(s)G_n(s)}$$

where $X(s)$ is, again, the output and $R(s)$ is the set point. Another point of interest when using a Smith predictor architecture is the ability of the system to compensate for a nonzero averaged disturbance in the plant. By looking at the disturbance-output transfer function:

$$\frac{X(s)}{D(s)} = P_n(s)\left(1 - \frac{C(s)P_n(s)}{1 + C(s)G_n(s)}\right) \tag{1}$$

The poles of P(s) cannot be eliminated from this relation, and therefore the classic Smith architecture is not suitable to control unstable plants. In the sequel, the different approaches used in order to choose the internal controller of the Smith Predictor will be discussed.

4.1 Smith predictor with proportional controller

The first controller studied is a Smith predictor with a proportional controller inside. It can be easily proved that this controller is sufficient to ensure input-output stability. Here, the transfer function of the controller, including the feedback of the nominal plant is obtained:

$$C(s) = \frac{k/n}{1 + \frac{k/n}{s}(n - \sum_{i=1}^{n} e^{-T_i s})} \tag{2}$$

where n is the number of sources, k is the constant of the internal proportional controller and T_i is the delay of each source i. Including the expression of the controller in the block diagram of figure 1 and the expression of the Laplace transform of the occupancy level of the queue $X(s)$ yields:

$$X(s) = \frac{k/n \sum_1^n e^{-T_i s}}{s+k} R(s) - \frac{1 + k/s - k/(s \cdot n) \sum_1^n e^{-T_i s}}{s+k} D(s) \tag{3}$$

As can be observed in equation 3 the closed loop system is stable for all $k > 0$ and it behaves like a first order system with a time constant $1/k$. In addition to this, oscillations will appear due to the delays. It is also important to have an expression to calculate the steady state queue level. We can see how it is related to the amplitude of the disturbance $d(t)$:

$$x_\infty = r_\infty - \frac{1 + k/n \sum_1^n T_i}{k} d_\infty \tag{4}$$

To view the limitations of the proportional controller the following simulation analysis has been made. The simulated system has six sources with time delays varying between 0.01 and 0.1 seconds and a controller gain of 60.

It has been stated that when a proportional controller is used, the queue level highly depends on the set point and the available bit rate. Let us suppose that under some particular traffic conditions denoted by ABR_0, the set point r_1 has been established at the minimum required to maintain the queue level always above zero, i.e. the channel being occupied 100% of time. In Fig.(4), R_1 is the actual queue level when the set point is r_1. The ACR for each sequence is shown in Fig.(5); thus, 1 Mbit/s is conceded approximately.

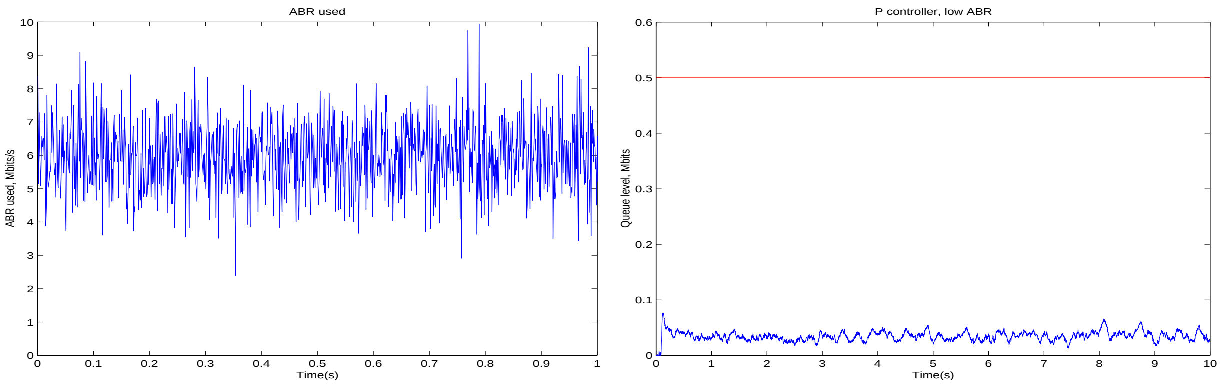

Figure 4: (a) ABR_0[2-10 Mbps], (b)Queue level, P controller

Now suppose that due to external conditions, the ABR traffic level increases to a different situation, namely $ABR_1 > ABR_0$. If the set point is kept at r_1 the queue level will decrease according to (4) and it will be too low to remain above zero, wich means that in some cases, the channel will be underused. Fig.(6) shows the queue level when ABR is ABR_1. By observing the level of the queue we conclude that the channel is left unused used most of the time.

To recover the queue up to a level where the channel is occupied 100% of time it would be necessary to raise the set point. It will be adjusted to $r_2 > r_1$ in order to have an optimal link utilization. Fig.(7) shows the resulting queue level.

It is obvious that an extremely low queue level can be restored by raising the set point. Nonetheless if the external conditions return to ABR_0 the average queue level may be too high and transport delays could exceed their maximum permitted. In Fig.(8)it is shown that keeping the set point at r_2 and changing the ABR to ABR_1 results in an excess of average queue level.

This analysis leads to the conclusion that under bursting ABR conditions, the set point must be readjusted continuously to keep performance optimal (lowest delay and full channel

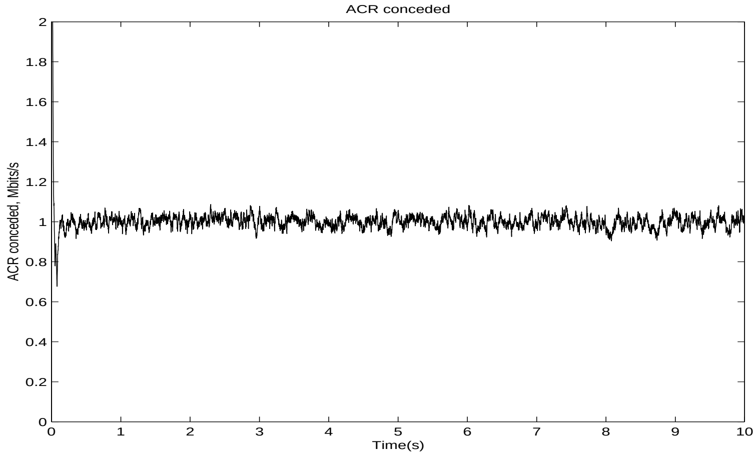

Figure 5: ACR per VC, ABR_0[2-10 Mbps]

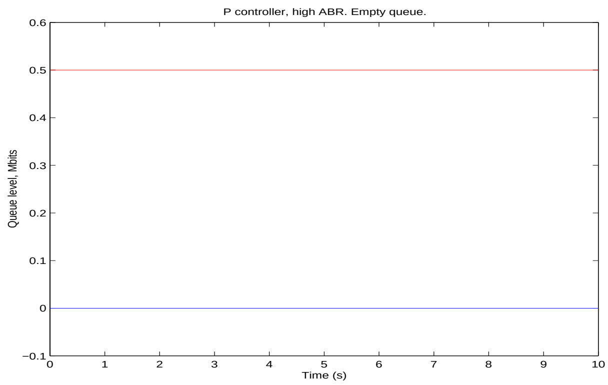

Figure 6: Queue level, P controller, ABR_1[36-44 Mbps]

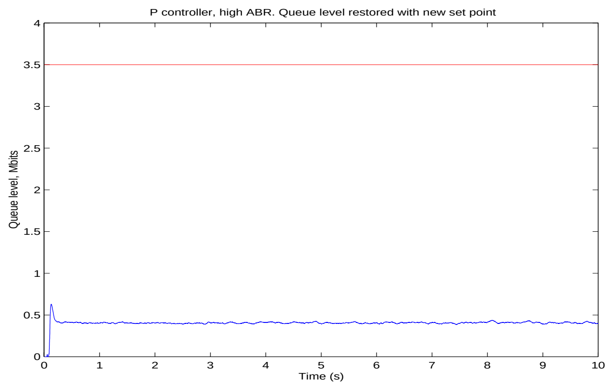

Figure 7: Queue level restored, P controller, ABR_1[36-44 Mbps]

utilization). Some improvements in the controller may reduce overall sensitivity as will be explained in the following sections.

4.2 Smith predictor with PI controller

The addition of an integral term into the inner controller of the Smith Predictor will reduce the undesired sensitivity to ABR conditions. A straightforward calculation leads to the steady-

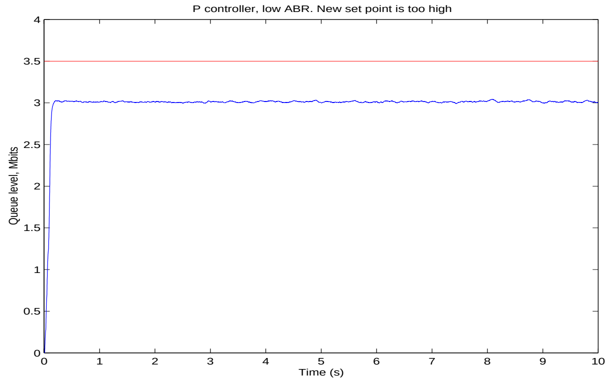

Figure 8: Queue level, P controller, ABR0[2-10 Mbps]

state queue level when using a PI controller:

$$x_{PI\infty} = r_\infty - \frac{\sum T_i}{n} d_\infty \tag{5}$$

The term affecting the disturbance d_∞ can be compared to the one obtained with a proportional controller in equation 4 to observe that the PI is more efficient in rejecting this value:

$$\frac{\sum T_i}{n} < \frac{1 + k/n \sum_1^n T_i}{k} \tag{6}$$

where the term on the left is the expression of the PI case. The improved behavior will be illustrated via simulation. Now let us simulate the system with a PI controller and ABR=ABR_1[36-44 Mbps] as in the last section. The set point r_3 is adjusted in order to have an optimal link utilization, as shown in Fig.(9). Due to the PI action, this reference must not be set as high as before, i.e. $r_3 < r_2$.

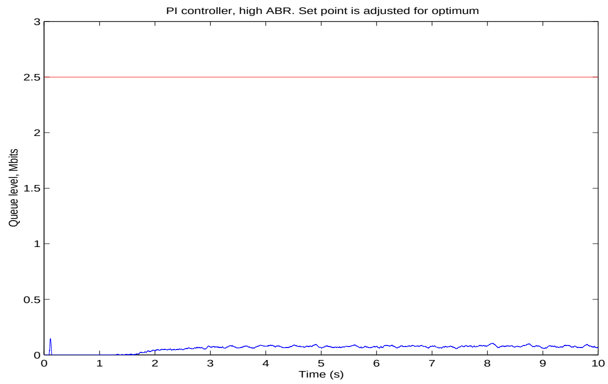

Figure 9: Queue level, PI controller, ABR_1[36-44 Mbps]

Now if the ABR suddenly falls to ABR0, the same reference r_3 (optimal for ABR_1) gives a reasonably low queue level for ABR0, as shown in Fig.(10). This level is lower than the one in Fig.(8), obtained with a P controller, under the same circumstances, i.e. a constant reference for different traffic loads.

As a conclusion, bursts in the ABR conditions can be more efficiently compensated for when a PI controller is used, with respect to the P controller. It is not necessary to continuously readjust the set point to keep a good performance (lowest delay and full channel utilization).

Nevertheless, the queue level for low ABR values are still relatively high, and sensitivity to ABR conditions remains high. The controller can be improved to further reduce this sensitivity.

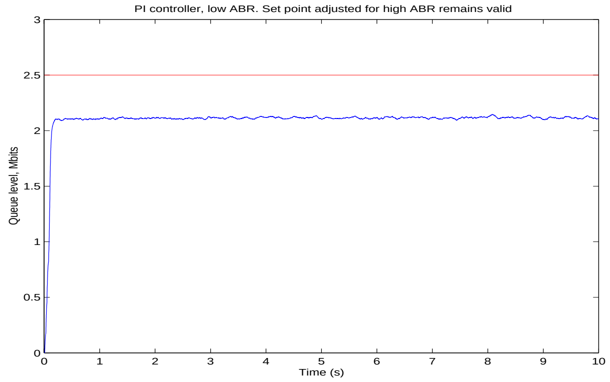

Figure 10: Queue level, PI controller, ABR0[2-10 Mbps]

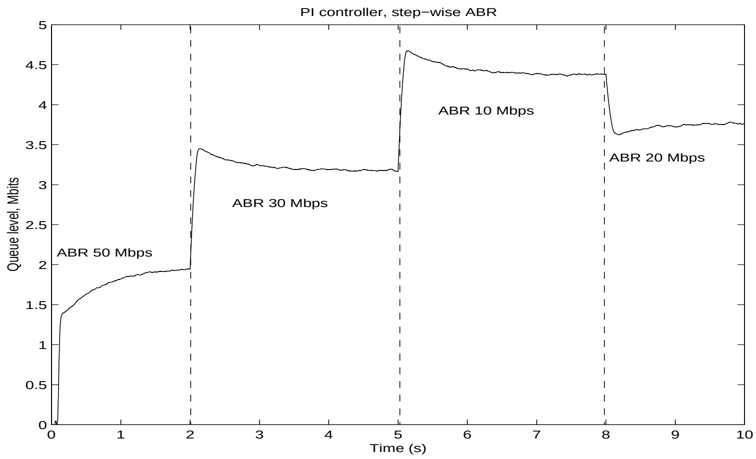

Figure 11: Queue variations with a step-wise ABR, PI controller

4.3 PI controller with feedforward

In order to compensate for the control delay caused by the NRM cells, a Smith predictor has been proposed. A PI controller inside attempts to lower the average queue level but it will not fully compensate for it, as is shown by the disturbance-rejection transfer function in Smith predictors (see [9]). This has been stated in equation (1). The previous section illustrates the effect of ABR variations on the average queue level. As Fig.(11) shows, steps appear in the average queue level, as ABR conditions change in a step-wise manner.

Nevertheless, the structure can be modified by injecting the disturbance signal (if available) into the inner control loop of the Smith predictor. The disturbance-rejection transfer function then becomes:

$$\frac{X(s)}{D(s)} = P_n(s)\frac{-1}{1 + C(s)G_n(s)} \tag{7}$$

Here, the PI controller $C(s)$ will compensate the unstable pole of $P(s)$. Whereas this is true whenever the disturbance is known (see Fig.12). As a result, the steady-state queue level would not depend on the average ABR, and the set point can be reduced to the optimal level for any ABR conditions. This is illustrated in Fig.(13).

However, the disturbance is often difficult to measure and a noisy estimation should be enough. In those cases, a properly adjusted feed-forward filter can be added in order to handle estimation errors. In the problem treated here, the real ABR capacity is unknown to the switch, but the switch can observe the movements of the queue level. Differentiating the queue level with respect to time and subtracting the control signal from it yields the estimation. Nevertheless, this will be a very noisy estimation because control signals are delayed and

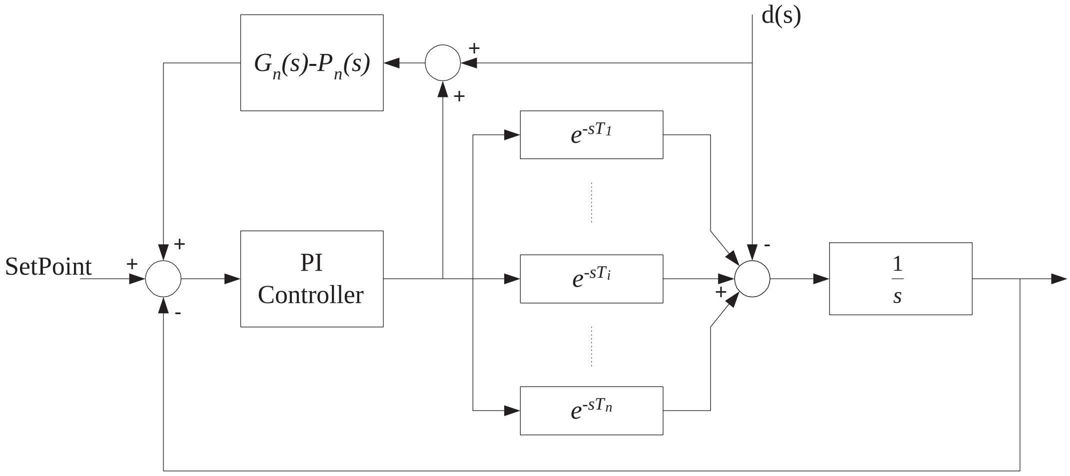

Figure 12: Smith predictor with feed-forward

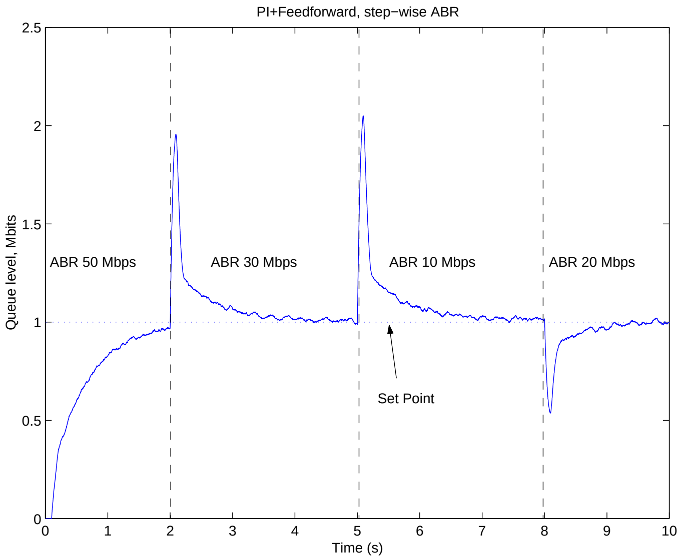

Figure 13: Queue variations with a step-wise ABR, PI controller with feedforward

ABR capacity is not. This justifies applying a low-pass filter to the result before using it. The speed of the pole of this filter is crucial to the overall performance of the controller and must be adjusted in light of the expected slew rate in the ABR signal. The resulting control architecture is shown in Fig.(14) The low-pass filter cutoff frequency is chosen in view of the rate of change of the average traffic demand. The time variations in ABR traffic can be ideally split in two parts, a high rate of change, due to the background noise caused by many simultaneous VBR connections, and slow variations due to a moving average of the traffic demand. The slow variations are caused by changes in business activity throughout the dayly schedule. Hence, the slow network activity change is characterized on a one-day basis, by measuring the traffic demand at intervals of 10 min or more.

As time constants related to the slow process are clearly higher than one minute, a filter with a cutoff frequency greater than 0.1 rad/s will allow the system to properly compensate for slow changes while rejecting the high frequencies which could cause instability in the presence of delays.

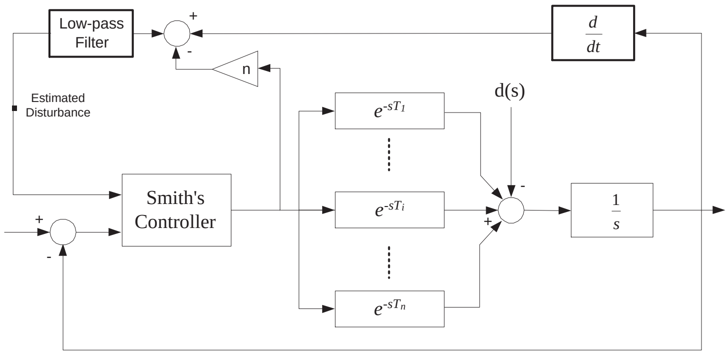

Figure 14: Control scheme with disturbance compensation

4.4 Anti-windup

As seen in section 2, the saturation problem must be dealt with. This implies a strong nonlinearity so that linear PI controllers can decrease in performance. In Fig.(16), the switch works in conditions near to the saturation point at zero. The anti-windup mechanism will prevent the integral part of the controller from accumulating memory of the error signal when it is caused by nonlinear phenomena like saturation. The anti-windup mechanism can be easily understood by looking at Fig.(15), where the system includes a new feedback loop, injecting the signal generated by the difference between the controller output and the real input to the plant. This difference is introduced into an integrator with integral constant T_t and is added to the controller output. The signal through the new loop is zero if the control signal is not saturated, but if the actuator is saturated at its maximum value, then $v(t) > u(t)$ and the signal injected to the new loop is negative, subtracting a certain amount from the accumulated error. In the case of saturation at the minimum the effect is the opposite.

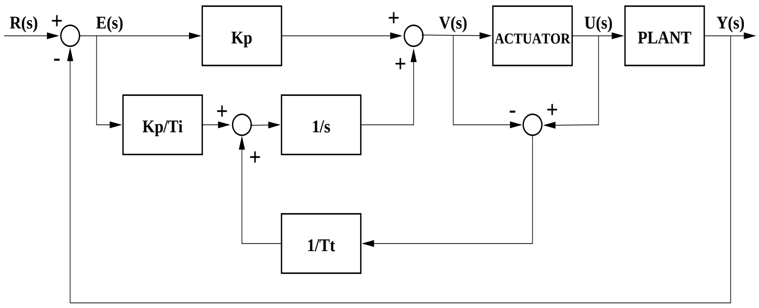

Figure 15: PI control with anti-windup mechanism

In Fig.(16) the control signal is observed when the ABR traffic falls drastically from 50 to 10 Mbit/s. The control signal reaches zero for a short while and the saturation prevents it from having a negative value. As the signal is saturated at zero, the excess under zero is fed back to the integrator to avoid an uncontrolled growth of the integral term.

5 Results

The introduction of the feedforward term into the controller allows to fully control the queue level, as shown in Fig.(13). However this is made under the assumption that the disturbance $d(t)$ is known. The last controller presented considers $d(t)$ unknown and which will be estimated from the queue level and filtered in order to avoid potential oscillations caused by time delays. Fig.(17) depicts the behavior of the final controller with estimation feedback, anti-windup and

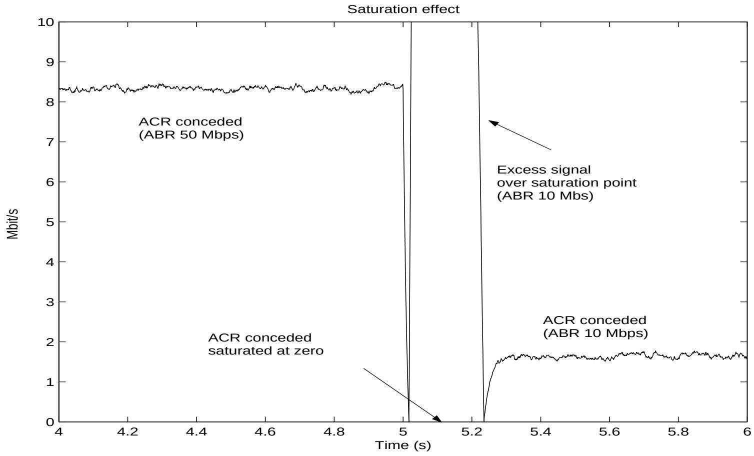

Figure 16: Anti-windup mechanism. Control signal and saturation feedback

a Smith Predictor with internal PI controller. As stated above, there is a filter preventing oscillations in the estimation. These oscillations appear at the queue level due to delays and delay estimation errors. To distinguish between the average queue level which is directly related to the average ABR and the oscillations themselves, a time constant near the shortest delay is assigned to the variations and a long time constant based on daily traffic schedule is assigned to the average ABR variations. These two time constants are sufficiently apart to only use a first order filter to get rid of the oscillations. In Fig.(17) the effect of setting the ABR conditions to 50, 30, 10, and 20 Mbit/s is successfully regulated by the proposed controller with the filter mentioned above.

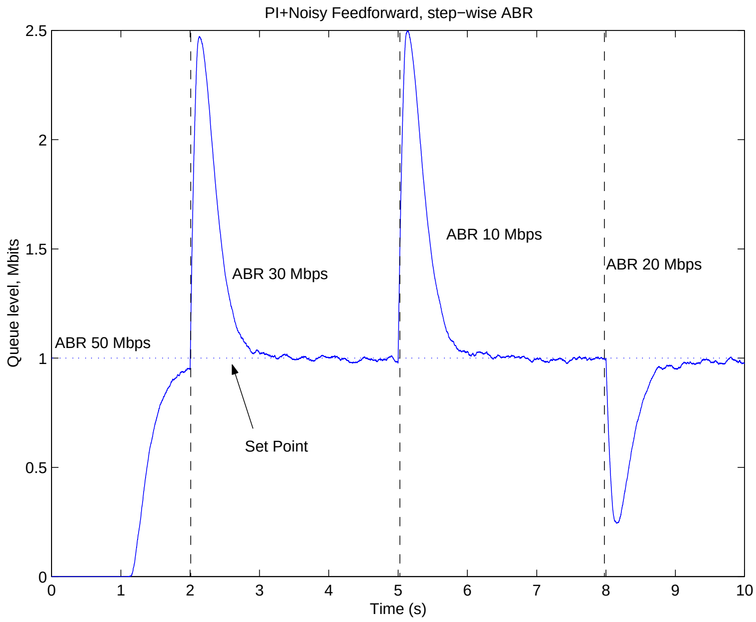

Figure 17: Queue variations with a step-wise ABR, PI controller with noisy feedforward

As a final study, the output signal will be analysed in depth. The average value of the queue, its standard deviation as a measure of the oscillations, the sensitivity to parameter estimation errors (mainly in delays) and hence robustness will be considered below.

The main properties of the output signal when analysing the behavior of the closed loop system by simulation are

- The average occupancy level. As expected, it follows the set point of the system. In

Fig.(17) it can be observed that there is no relation between the steady-state average level of the queue and the average ABR signal.

- The standard deviation of the queue level, that is, the magnitude of the oscillations. It is related to the fast variations of the ABR introduced and the oscillations caused by delays. Adding 20% of white noise to a step-wise ABR signal gives the results shown in table 1. Again, there is no dependence between the oscillations of the queue and neither the average ABR value nor the set point used. This is an important point for the controller, where the properties remain constant in a wide range.
- The change in the standard deviation when errors in the delay estimation appear. Delay estimation errors produce oscillations due to the incorrect balance of the Smith Predictor. These oscillations are measured using the standard deviation, because of the stochastic nature of the ABR. Large oscillations may force the controller designer to raise the set point (and the transmission delays) in order to ensure full channel utilization. As can be seen in Fig.(18), the magnitude of oscillations in the queue increases slowly as the real delays differ from the estimations used in the controller. A maximum delay estimation error of 50% is considered for all sources simultaneously. This error increases the standard deviation of the queue level in less than 50 cells.
- The portion of time the queue is empty, i.e. the channel not being fully used. This can be made near zero in a steady state when assigning the set point twice the expected standard deviation.
- For a specific controller, the maximum step size in a bursting ABR signal is limited. An instant ABR change higher than this maximum will produce a queue overflow.
- The new oscillations that can appear near the saturation points. This has been studied in table 1, to obtain the maximum step size in a step-wise ABR signal that the controller can handle without exceeding the queue size.

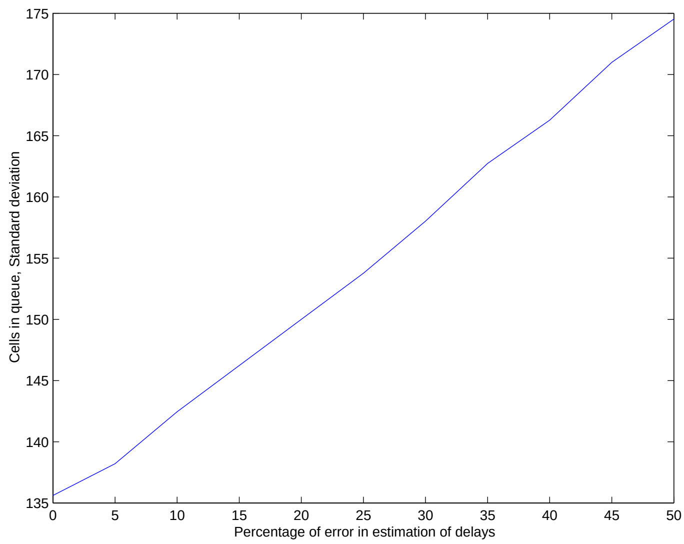

Figure 18: Std. deviation of the queue vs. the percentage of error in estimations

6 Conclusions

In this paper we have presented the problem of ABR bandwidth assignment in ATM networks with dead-times. Improving the control algorithm for traffic assignment can increase the practical applications and the interest in this service.

Set Point	Cells	10 Mbit/s	20 Mbit/s	30 Mbit/s	40 Mbit/s	50 Mbit/s
2 Mbits	4946 cells	0,0568	0,0568	0,0611	0,0511	0,0499
1,5 Mbits	3710 cells	0,0503	0,0503	0,0507	0,0533	0,0499
1 Mbits	2473 cells	0,0503	0,0503	0,0502	0,0526	0,0499
0,5 Mbits	1237 cells	0,0503	0,0506	0,0558	0,053	0,0499

Table 1: Standard deviation of the queue level

First, the problem has been addressed with a classic control scheme based on a Smith predictor with a proportional controller inside, proposed in [5]. The closed loop stability of this system is guaranteed. The simulation shows a strong variation in the queue level when the available bit rate in the network changes. Not only did the average queue level present oscillations due to the delays, but its average value was too sensitive to working conditions. This value is a critical parameter when offering a traffic service as it is proportional to the transport delay introduced by the switch. It also determines the size of the buffers used in the ATM switches. This suggested the addition of new techniques to gain control over the queue level.

A better approach has been proposed in [8] where it was proven that the interdependence between the queue level and the ABR conditions is reduced when the P controller is substituted by a proportional-integral block.

While reducing the average queue level, the Smith predictor with a PI controller is still unable to take it to the desired set point. This has been clearly seen in simulation in Fig.(11) for different ABR conditions. The reason for the inability of the Smith predictor to fully compensate the disturbance of the system has been analysed in Eq.(1). It fails when the plant is unstable as it has an integrator in the model presented here.

A new modification has been proposed in this paper so that the dependence of the PI Smith controller on the disturbance completely disappears. This is accomplished by adding a feedforward of the estimation of the disturbance into the controller input. This estimation is easily obtained by observing the initial and final size of the queue on a time interval and dividing their difference by the time elapsed. Subtracting the control signal from the result yields the estimation of the real ABR capacity. This estimation can be very noisy, so a low-pass filter is added before introducing it into the controller.

The low pass filter cutoff frequency is chosen in view of the time constant of changes in the average traffic demand. The time variations in ABR traffic can be ideally split in two parts, a high rate of change, due to the background noise caused by many simultaneous VBR connections, and slow variations due to a moving average of the traffic demand. The slow variations are caused by changes in business activity throughout the day schedule. Hence, the slow network activity change is characterized on a daily basis, by measuring the traffic demand at intervals of 10 min or more.

As time constants related to the slow process are clearly higher than one minute, a filter with a cutoff frequency greater than 0.1 rad/s will allow the system to properly compensate slow changes while rejecting the high frequencies which could cause instability in the presence of delays.

The queue level of the system with the feedforward follows the desired set point, as seen in Fig.(17). At this stage the set point can be tuned to minimize the queue level in terms of the expected standard deviation of the output. This parameter and also the admisible variations of the queue level must be considered to ensure the full utilization of the link. With this purpose, a practical study of the oscillations has been included in this article.

In order to deal with more realistic situations, saturations in the control signal and the queue size are considered in the simulations. The inclusion of anti-windup techniques in the controller prevents side-effects from saturations. Finally, sensitivity to delay estimation errors is analysed, and the limitations of the proposed controller are discussed with respect to ABR traffic parameters.

The applications of the ABR service can increase with the use of this new controller. Communications with restrictions on time delay or oscillations can now rely on this service as it is no longer indeterminate during normal operation. The reduction achieved in the queue size permits the use of lower buffer sizes at the switch, reducing its cost. It can also reduce the cell loss rate. As a future study, we are developing new controllers in order to improve the performance of video and voice transmission over ABR ATM service.

Acknowledgements. The authors would like to acknowledge CICYT for funding this work under grant TAP-98-0541, and the FEDER project 1FD97-0783. The authors also acknowledge many helpful suggestions by the reviewers and editors which have been incorporated into the manuscript.

References

[1] S. Mascolo. "Smith's Principle for Congestion Control in High Speed ATM Networks" Proc. of 36th IEEE CDC, San Diego, Ca. Dec 1997.

[2] S. Mascolo "Congestion Control in High Speed communication networks using the Smith Principle", Automatica 35, pp 1921-1935, 1999.

[3] S. Mascolo, D. Cavendish and M. Gerla. "ATM rate-based congestion control using a Smith Predictor: an EPRCA Implementation" Proc. of IEEE Infocom 96, San Francisco, CA, March 1996.

[4] D. Cavendish, S. Mascolo and M. Gerla. "SP-EPRCA: an ATM Rate-Based Congestion Control Scheme based on a Smith Predictor", Technical report, UCLA, 1997.

[5] S. Mascolo and M. Gerla. "Classical Control Approach to congestion Avoidance in ATM High Speed Networks" IEEE ATM'97 Workshop, pages 361–367, Lisbon, 1997

[6] Flavio Bonomi and Kerry W. Fendwick "The Rate-Based Control Framework for the Available Bit Rate ATM Service", IEEE Network, pp.25-39, Mar./Apr. 1995

[7] J. Jormaka "ABR with Bursty Traffic Sources", IEEE/ATM Workshop 97. pp.427-436, Lisbon, 1997

[8] F.R. Rubio and J.Fornés. "ABR Control for Video Transmission over ATM Networks". International Congress on Modelling, Identification and Control (MIC'99), pp 514-517. 1999.

[9] J. Normay, "Prediction for control", PhD. Thesis, June 1999

[10] Z.J. Palmor, "Time-Delay Compensation. Smith Predictor and its Modifications". The Control Handbook. Ed. W.S Levine. pp 224-237. IEEE Press. 1996.

[11] ATM Forum, "ATM Traffic Management Specification Version 4.0", April 1996, af-tm-0056.000