ABR Traffic Control Using Feedback Information and Algorithm

Kwang-Ok Lee[a] and Young –Su Son[b] and Hyeon-ju Kim[a] and Sang-Hyun Bae[a]

[a]*Dept. of Computer Science & Statistics, Chosun University, Korea*

375, Seosuk-Dong, Dong-Gu Kwangju 501-759

Tel: +82-62-230-6623, Fax: +82-62-230-6623, E-mail : csskwang@orgio.net, onlywom@empal.com

[b]*Dept. of Multimedia Informaiton & Telecommunication Engineering, Hanlyo University, Korea*

199-4, Dokryeri, Kwangyang 545-704

Tel: +82-62-230-6623, Fax: +82-62-230-6623, E-mail : son@hlu.hanlyo.ac.kr

Abstract *ATM ABR service controls network traffic using feedback information on the network congestion situation in order to guarantee the demanded service qualities and the available cell rates. In this paper we apply the control method using queue length prediction to the formation of feedback information for more efficient ABR traffic control. If backward node receive the longer delayed feedback information on the impending congestion, the switch can be already congested from the uncontrolled arriving traffic and the fluctuation of queue length can be inefficiently high in the continuing time intervals.*

The feedback control method proposed in this paper predicts the queue length in the switch using the slope of queue length prediction function and queue length changes in time-series. The predicted congestion information is backward to the node. NLMS and neural network are used as the predictive control functions, and they are compared from performance on the queue length prediction. Simulation results show the efficiency of the proposed method compared to the feedback control method without the prediction. Therefore, we conclude that the efficient congestion and stability of the queue length controls are possible using the prediction scheme that can resolve the problems caused from the longer delays of the feedback information.

1. Introduction

ABR service should use an appropriate control for an unpredictable congestion due to a feature of data traffic. A feedback mechanism is used for a dynamic control of the transmission rate of each source to a present network state in order to guarantee the quality of a required service[1][2]. ABR service has been also devised for a fair distribution of an available bandwidth for ABR users. As it were, it should maintain a better packet loss rate and a fair share of given resources by an adaptive adjustment to a network state. In addition to ATM cell, the identity management cell having a control information is called RM(Resource Management) cell in ATM network. The feedback mechanism for ABR service uses RM cell in order to provide a traffic control information. This RM cell having a detailed description of a control information is transmitted to a source, which adjusts a cell transmission rate suitable for a present network by using the information cell[2].

Most studies of feedback congestion control schemes for a ABR traffic control tend to focus on the control algorithms using a threshold of internal queue of ATM switch[2][3].

However, an effective control of a source traffic already transmitted before controlled would be impossible in the existing algorithms, because the transmission time of a backward RM delayed due to the congestion between a source and a destination[6][7][8]. A congestion at the switch can occur due to a control information delay, and thus a variation of queue length can also occur over time. A variation of queue length impedes an efficient ATM traffic control. A delay of feedback information transmission can be caused not only by a long physical transmission time of a network but also by a network congestion.

This paper proposes a predictive control function and feedback algorithm improved for an even more effective traffic control than the algorithms[2] for a long feedback delay within a time-out period after the establishment of a dynamic connection. The algorithm which is implemented at a switch predicts a future value of queue length, sends a queue length of a switch to a source in advance, and prevents a congestion. It also controls a variation of a queue length to the utmost. That is, it uses a feedback information, as it increases or decreases a transmission rate of a source beforehand in a computation of a future queue length at a switch.

In order to predict a future queue length, it monitors periodically a cell input rate to a switch and a recent queue length. It adapts periodically a predictive function of a future queue length to an optimized value using NLMS(Normalized Least Mean Square)[4][5] and an optimized adaptation of a neural network[9]. A new transmission rate of a source is computed with a feedback algorithm using an existing threshold value and a predictive function of a future queue length. As a predictive function of a future queue length, NLMS method and an optimized adaptation of a neural network method predict a queue length using a linear function and non-linear function respectively. I studied a predictive control method of ABR traffic which was even more efficient through a simulation using the two methods described above.

2. A predictive feedback control model

2.1 A proposed predictive feedback control model

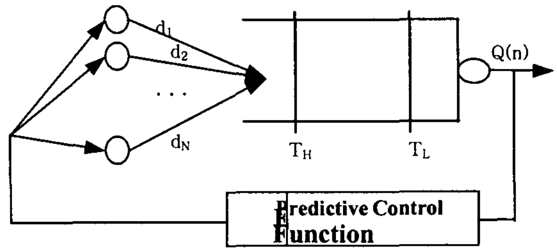

Figure. 1- Feedback predictive control model

A proposed predictive feedback control model is presented in Fig. 1 above. N sources transmit packet data cells in a single switch, a cell transmission rate is constant, and a queue state is monitored regularly. It is assumed that a transmission delay time of packet data between a source and a switch is di, and that sources is added or deleted randomly for ABR service. A network state is specified in time n by Q(n) of a queue length at a switch node. For a given ABR traffic processing buffer, TH and TL show high and low thresholds respectively. A predictive control function computes a future queue length in response to time-series by a queue length. When a future predictive queue size exceeds the high threshold TH, the switch is considered to be in a congestion and the switch computes the explicit rate(ER) at which sources have to send a backward RM cell to the switch in order to avoid a congestion. If it is less than the high threshold T_H, however, a source changes its transmission rate in its computation of ACR(Available Cell Rate) by being informed of non-congestion situation instead of ER.

2.2 A predictive control function using NLMS

NLMS control estimates buffer size in the next k steps using a linear function with a current value of the buffer size and weighting factor(slope) at time n.

Let Q(n) denote the buffer size at time n. The k-step predictor is formulated such that the buffer size at k steps in the future is estimated from the Q(n), as given by

$$Q(n+k) = a^k(n)Q(n) \qquad (1)$$

where $a(n)$ is an estimated weighting factor at time instant n, and $k = 1,2,3,\text{K},t$ and t is a maximum prediction interval. Error of the prediction at time n is

$$e(n) = Q(n) - \hat{Q}(n) \qquad (2)$$

where

$$\hat{Q}(n) = a(n-1)Q(n-1) \qquad (3)$$

The prediction scheme uses the error to modify the weighting factor whenever the error is available at each time step. Furthermore, the weighting factor $a(n)$ is affected in time as sources are added or removed and as the activity levels of source changes. We thus put the problem into the one of estimating the weighting factor and use the normalized least mean square error (NLMS) linear prediction algorithm. Given an initial value for $a(0) = 0$, the weighting factors are updated by

$$a(n) = a(n-1) + \frac{\mu e(n)Q(n-1)}{|Q(n-1)|^2} \qquad (4)$$

where μ is a constant. If $Q(n)$ is stationary, $a(n)$ is known to converge in the mean squared sense to the optimal solution[1][4][5]. The NLMS is known to be less sensitive to the factor μ. The estimated weighting factor $a(n)$ in each time- step will be used to predict the buffer size $\hat{Q}(n)$. Therefore each time step, the weighting factor indicates the direction of evolution of the functions for buffer size increases/decreases in term of recent residual $e(n)$ computed by the estimated buffer size $\hat{Q}(n)$ and actual buffer sizes $Q(n)$.

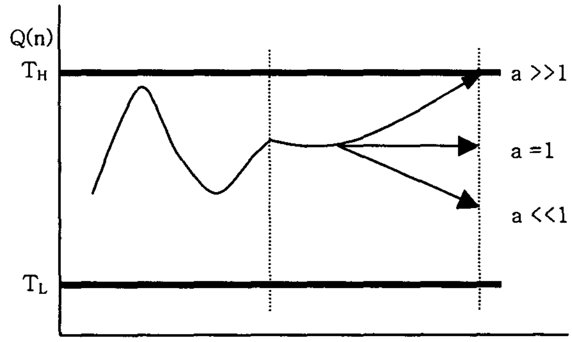

Figure 2- Changed graph in applying a predictive algorithm

Therefore Fig. 2 presents a predictive scheme described in this paper[6]. If $a >> 1$, a predictive queue length increased by the expression $Q(n+k) = a^k(n)Q(n)$. Therefore k, a time to hit the T_H, is predicted at time n using $Q(n)$ and $a(n)$ which are clearly known at time n [6].

2.3 A predictive control function of neural network using BP

A non-linear predictive function using neural network adjusts to predict a optimized value using BP algorithm[9]. It computes optimized variables of a non-linear equation(sigmoid) included in neural network nodes, and adjusts to get minimal errors to be occurred in a predictive value. That is, as in Fig. 3, BP is a kind of delta learning method to adjust adaptively the degree of a connection in order to minimize the differential error between required output and predictive output. Input layer x_i got continuously changing queue length $Q(n), Q(n-1), \cdots, Q(n-m-1)$ in time units, and output layer got a predictive value of queue length $Q(n+k)$ after n+k.

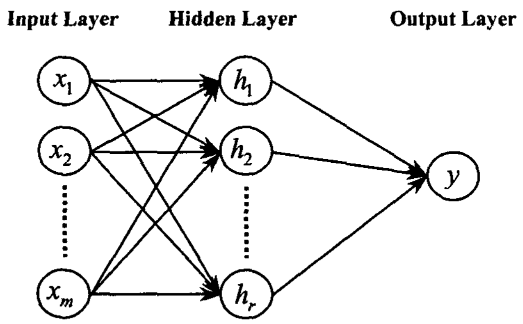

Figure 3- Multi-layer Neural Network Structure

A case using neural network as in using NLMS also predicts a future queue length through monitoring queue length at a switch. However, the case is more complicated than the case of NLMS, because a weighted value for each connection link should be computed in advanced for optimal adaption. The detailed computation processing of BP algorithm is consulted in Reference [9].

2.4 A feedback algorithm using a predictive value

A feedback algorithm is explained for an implementation of ABR feedback control using a predictive control function. A predictive queue length is computed using a predictive control function with consulting high and low thresholds and with monitoring present queue length at ATM switch. If a predictive value is over high threshold, it sends minimal cell transmission rate to each source in advance after its computation as a congestion is impending. By performing a prediction, it prevents a congestion due to cell inflow from sources having long transmission delay of feedback information. Fig. 4 is presented as a predictive algorithm for a congestion.

1) **Predict a Normal State:** When no congestion is detected the $\hat{Q}(n+k) < T_H$ and

[1] : Initialize ABR parameter
 FS = Link_speed_Switch / Number_sources
 Interval ← initial value
[2] : For n do [3-4] until max n
 [3] : IF n=1 *Initialize training parameter*
  ELSE *NLMS or BP algorithm*
 [4] : IF $\hat{Q}(n+k) \ge Q_H$
  Congestion=1
  ER←FS×ERF
  ACR←min[ER, ACR(1-RDF)]
 ELSE IF $\hat{Q}(n+k) < Q_H$
  and $\hat{Q}(n+k) > Q_L$
  ACR←min[ER, PCR, ACR+RIF×PCR]
  goto [2]
[5] : Stop condition

Figure 4- Feedback algorithm using a predictive function

$\hat{Q}(n+k) > T_L$, the value of the ACR for the sources is computed at the switch by the linear increase algorithm as specified in the source behavior in the ATM Forum standard[2][8].

2) **Predict a Underload State:** When no congestion is detected and $\hat{Q}(n+k) < T_L$ then the ACR for the sources is computed at the switch using the exponential increase.[8].

3) **Predict a Overload State:** If the congestion is detected in terms of the predicted queue length, $\hat{Q}(n+k)$ at the time n, ER is computed by

ER=Fair_ShareExplicit_Reduction_Factor,

Where

Fair_Share=Link_speed_at_switch/Number_of_sources.

In Fig. 4, at a congestion, if all the sources get cell transmission rate computed at ATM switch through RM cell, the cell transmission rate cannot be over ER. ACR is the next cell transmission rate computed at each source, in case a congestion does not occur at the switch(it is not specified at RM cell).

3. Simulation

3.1 Simulation environment

As in Fig. 5, the simulation model of a control algorithm presented in this paper is that the link speed of the switch is set to 150 Mbps, and the link speed for each source to the switch is set to 150 Mbps/N for N sources. The control algorithm is experimented in Visual C++ for a single bottleneck switch with a buffer of high and low thresholds of 5000 and 1000 cells respectively.

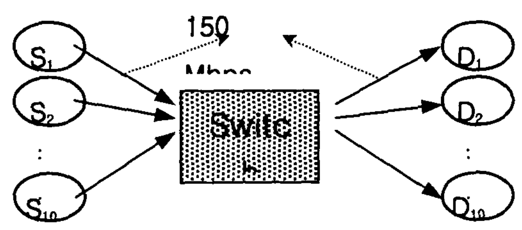

Figure 5- Simulation Model

Following parameters are used for the simulation: peak cell rate(PCR) is set to 150 Mbps, additive increase rate(AIR) is set to 0.1 Mbps, and explicit reduction factor(ERF) is defined to 4/5. Ten active sources with various packet data cell generation times are used for simulation. In order to examine the transitional behaviors, an abrupt change of active sources is made. Initially, sources with cell generation times at {2, 4, 6, 9, 11, 13, 16, 20, 21, 23} are active, which the numbers represents the time-delay d_i from current time-unit n at the switch to the sources. At time-unit 1000, sources with time-delay d_i {14, 16, 17, 19, 20, 22, 23, 26, 28, 30} are active, which it includes the active sources with long delays. Two cases are compared in terms of a stabilization and a congestion of queue length at the switch through the change of transmission delay. The first case uses only feedback control method, and the second one does feedback predictive control method.

3.2 Simulation results

Fig. 6 presents the change of queue length size at each switch, one of which uses a feedback predictive control algorithm using NLMS proposed in this paper, and the other of which uses only a feedback control one. A predictive interval(k) for NLMS was 10. Fig 6 presents that A feedback control algorithm only always brings about a congestion, and that a variation of queue length is considerably severe. It also shows that a variation of the length size Q(n) is severe after time 1000. it means that the sources with much longer delay time than other ones are incoming at the same time.

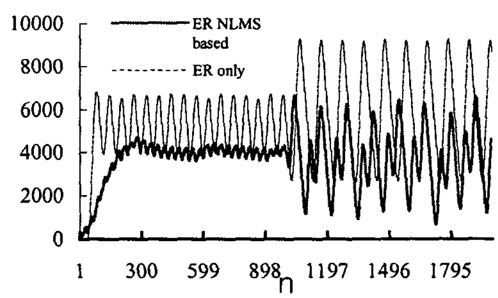

Figure 6- Comparison of Q(n) for ER only and ER with NLMS

Fig 6 shows that NLMS feedback predictive control algorithm presents no variation of queue length before time unit 1000, close to high and low threshold, compared with a feedback control one only. After time unit 1000, however, as the sources with much longer delay time than other ones are incoming at the same time, the variation occurs more severely than before 1000 even with the predictive control method, and cases exceeding over high and low threshold also occur. The reason is that feedback delay of transmission sources is longer from time unit 1000, and that as a worst condition any traffic does not occur during time delay 1 through 13. That is, a predictive control function responds inappropriately to constant long-term interval, or sudden and random change. However, It is concluded that a predictive control algorithm caused a stability of the change and not severe congestion during simulation, compared with a feedback control one only. The use of neural network structure brought about similar results. It also responds inappropriately after n=1000.

Neural network structure needs a training using long-term BP, compared with NLMS, to adapt non-linear predictive control function. As in Fig. 7, a variation of queue length can be stabilized by a rapid drop of error rate. A systematic establishment for neural network structure should be preceded[9]. The decision on how many nodes of hidden layer are needed is required. In order to solve the problem about an inappropriate response of a predictive control function to occur after n=1000, a following method was tested. In the algorithm proposed in Fig. 4, increase rate control computation was used with constant instead of linear increase in ACR computation when predicting normal queue state. Therefore, traffic occurrence could not be detected, and a constant was increased in case normal queue state predicted. A simulation result for it is presented in Fig. 8. A change of queue length occurred when neural network structure had input node 10, hidden layer node 10, and k=10. NLMS predictive control function also had k=10.

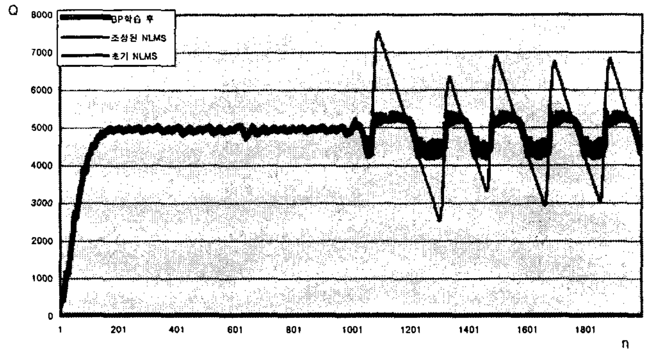

Figure 7- Buffer size prediction learning result using BP

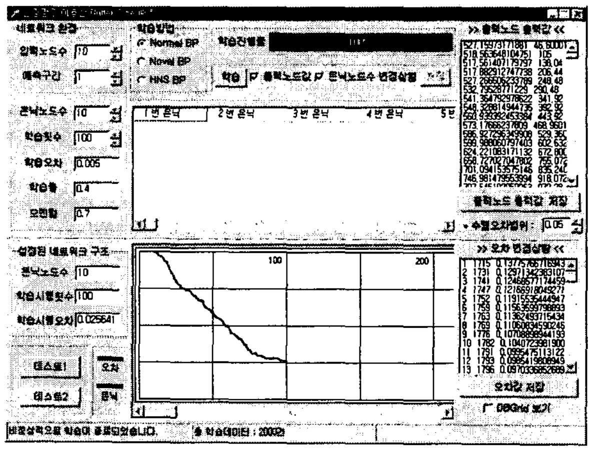

Figure 8- Simulation result using constant increase

For comparison in Fig. 8, NLMS presents the most severe change of queue length, and normal predictive case of queue state is the result of using linear increase method. The other two graphs of queue length change represent the use of NLMS and neural network respectively as a predictive control function, and a constant increase method is used in normal predictive state of queue length. The figure shows a stability of queue length and a non-congestion. NLMS represents a traffic control result similar to neural network.

4. Conclusion

This paper studied that a congestion at switch was predicted in advance and thus traffic could be controlled. Also, by making an active use of the result of predicted queue length as feedback control information, the sources could be informed of prompt and precise congestion situation. A predictive algorithm based on NLMS prediction scheme estimated the buffer size in the next k steps by using NLMS, and the control algorithm based on ER algorithm was applied. The case of NLMS predictive algorithm used proved to be effective.

Neural network structure also proved to be effective in controlling a congestion and queue length variation as in case NLMS predictive control function. In order to apply neural network, the establishment of an optimal neural network structure should be preceded, and many variables in it requires much more computation time needed for a training than in NLMS. Therefore, it is not suitable for ATM switch requiring a real-time processing.

However, the experimental comparison of a control algorithm based on a predictive algorithm with the other ER control algorithm is different in terms of the establishment of the input variables. Thus, it is difficult to assert that the algorithm presented in this paper proves to be effective by a simple comparison. The simulation in a different environment is required for the experimental verification of its effectiveness.

Reference

[1] Adas, A., "Supporting Real Time VBR Video Using Dynamic Reservation Based on Linear Prediction," Infocom 1996.

[2] ATM Forum, Traffic Management v.4.0 Aug. 1996.

[3] Black, Uyless., ATM Volume I Foundation for Broadband Networks, New Jersey: Prentice Hall PTR, Feb. 1999.

[4] Hayes, M. H, "Statistical Signal Processing and Modeling," John Wiley & Sons, 1996.

[5] Haykin, S., "Adaptive Filter Theory," Prentice Hall, 1991.

[6] Jang, B., B. G. Kim and G. Pecelli "A Prediction Algorithm for Feedback Control Models with Long Delays" IEEE BSS 1997.

[7] Mascolo, S., D. Cavendish and M. Gerla, "ATM Rate Based Congestion Control Using a Smith Predictor: an EPRCA Implementation," Infocom 1996.

[8] Ritter, M., "Network Buffer Requirements of the Rate-based Control Mechanism for ABR Services," Infocom 1996.

[9] Zurada, J.M, "Introduction to Artificial Neural Systems", WEST, 1992.