Predicted Sum: A Robust Measurement-Based Admission Control with Online Traffic Prediction

Matyas W. Egyhazy and Yao Liang, *Senior Member, IEEE*

***Abstract*— We present a new measurement-based admission control (MBAC) scheme with online traffic prediction, referred to as Predicted Sum (PS). A simple adaptive online traffic predictor based on normalized least mean squares (NLMS) algorithm is employed. Our scheme has two unique merits: (1) robustness to traffic characteristics and network environment, and (2) easy and accurate control for the provisioning of network utilization and quality-of-service. The simulation results demonstrate its significant performance enhancement over the well-known MBAC measured sum (MS) scheme.**

***Index Terms*— Admission control, multimedia networking, network traffic prediction, robustness, quality-of-service.**

I. Introduction

MEASUREMENT based admission control (MBAC) is essential to support quality-of-service (QoS) and high network utilization simultaneously for soft real-time services [1] [2]. In this letter, we propose a new MBAC scheme that employs online aggregate traffic prediction, referred to as Predicted Sum (PS), for admission control in multimedia packet networks. The term *online* is used to indicate that the traffic predictor is trained automatically in real time. The fundamental difference from traditional MBACs is that the admission tests of our approach are not based on the direct measurement of traffic history, instead they are based on online traffic prediction. Since variable-bit-rate (VBR) video traffic is highly bursty, using online traffic prediction is more capable of accurately estimating the future traffic load for the subsequent admission test. Our traffic predictor is based on the normalized least mean squares algorithm (NLMS), due to its simplicity and efficiency [3].

Recent rigorous comparisons between traditional MBAC schemes revealed that they all yield similar performance [4]. This indicates that, even though they may be dramatically different in nature, most traditional MBACs are basically equivalent regarding the set of operating points, given the traffic characteristics. However, as noted in [5], a great challenge for MBACs is their ability to achieve targeted network utilization and QoS *robustly* without excessive parameter tuning. This motivated our work. The robustness of our proposed PS scheme is rigorously investigated and compared to traditional MBAC Measured Sum (MS) [2] [1].

Manuscript received July 20, 2006. The associate editor coordinating the review of this letter and approving it for publication was Prof. Iakovos Venieris.

M. W. Egyhazy is with the Department of Computer Science, Virginia Polytechnic Institute and State University, Falls Church, VA.

Y. Liang is with the Department of Electrical and Computer Engineering, Virginia Polytechnic Institute and State University, Arlington, VA (email: yaliang@vt.edu).

Digital Object Identifier 10.1109/LCOMM.2007.061127.

II. Predicted Sum (PS) Scheme

A. Online Traffic Predictor

Let $x(n)$ denote the observed load, and $\mathbf{w}_n$ denote the predictor's weight vector $[w(0), w(1), \ldots, w(p-1)]^T$, at time n respectively. A 1-step pth-order linear predictor is given as

$$\hat{x}(n+1) = \sum_{l=0}^{p-1} w_n(l) x(n-l) \tag{1}$$

and the prediction error $e(n) = x(n+1) - \hat{x}(n+1)$. The Least Mean Square (LMS) algorithm uses error feedback to make successive updates to the weight vector $\mathbf{w}_n$ to reduce the mean square error as follows:

$$\mathbf{w}_{n+1} = \mathbf{w}_n + \mu e(n)\mathbf{x}(n) \tag{2}$$

where $\mathbf{x}(n) = [x(n), x(n-1), \ldots, x(n-p+1)]^T$ and μ is the step size parameter.

NLMS is a modification to the LMS algorithm that uses a normalization term for weight vector updating as follows:

$$\mathbf{w}_{n+1} = \mathbf{w}_n + \frac{\mu e(n-1)\mathbf{x}(n-1)}{\|\mathbf{x}(n-1)\|^2} \tag{3}$$

The error $e(n-1)$ is used instead of $e(n)$ in Equation 3 because $e(n)$ is not available at time n, and accordingly $\mathbf{x}(n-1)$ is used instead of $\mathbf{x}(n)$ [3]. We choose $\mu = 1$ as $0 < \mu < 2$ is necessary for convergence [6]. The coefficients of the NLMS predictor will be continuously adapted after every new sample becomes available.

B. Admission Control Algorithm

The proposed PS scheme samples the traffic load at certain intervals. These samples are the inputs for the online traffic predictor. After each sampling, a prediction is made for the traffic load of the next sampling period. This load forecast is then used in the future admission decision test to admit or reject new flows. In contrast, in traditional MBACs such as MS, the measurement of past traffic is used for the future admission decision test. The PS admits or rejects a new incoming flow α based on the following test:

$$\hat{x}(n+1) + r_\alpha \leq \eta C \tag{4}$$

where $\hat{x}(n+1)$is the predicted aggregate traffic load for the future sampling period $n+1$; r_α is the specified rate (peak rate or token rate) for the new incoming flow α; C is the capacity of the link; and η is a utilization parameter called the target utilization. The target utilization η is generally used to control a proper balance between the QoS target of the admitted flows

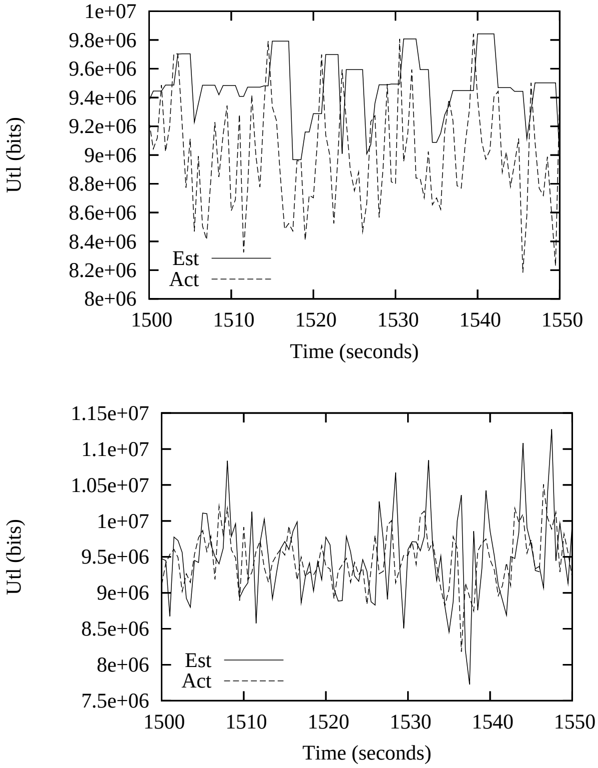

Fig. 1. A typical snapshot of: MS behavior (top), and PS behavior (bottom).

and the overall network utilization target. If the above test on the new flow α is successful, then the new flow α is admitted and otherwise it is rejected.

III. Performance Analysis

We adopt the bufferless network model used in [5] for the performance analysis in our study. The bufferless single link with capacity C in the model represents the bottleneck link in an actual network system. The metrics used are loss rate and utilization (loss-load). The similar structures of the PS and MS algorithms enable us to conduct a fair comparison. We focus on investigating the robustness of the PS scheme regarding: (1) the measurement window size T for the past aggregate traffic, (2) the sampling period size S, (3) the target utilization, and (4) the average hold time of admitted VBR video flows in the network.

Simulations are conducted using ns2. The base simulation environment is a simple topology with multiple flows connecting over a single link. The flows use a token bucket traffic shaping model. Our simulation uses real-world MPEG-4 VBR video traces as the traffic sources [7]. Five heavily-bursty MPEG-4 movies Jurassic Park, Silence of the Lambs, South Park, The Firm, and Star Wars IV are used to create a heterogeneous environment.

Figure 1 illustrates the PS and MS from a network bandwidth estimate vs. actual utilization perspective. The solid and dotted lines indicate the estimated and actual aggregate traffic load, respectively. The PS algorithm with traffic prediction closely tracks the state of the system, while the MS is overly conservative and does not track traffic load fluctuations.

TABLE I

Effect of T on PS.

T	Drop Rate	Utilization
2	0.001818	0.957197
4	0.001312	0.952968
12	0.001648	0.954684
16	0.001923	0.956988
20	0.002442	0.958311
24	0.003432	0.966778

TABLE II

Effect of T on MS.

T	Drop Rate	Utilization
1	0.000439	0.937216
3	$4.59276e^{-6}$	0.889605
10	0.0	0.840539

1) Measurement Window T : The measurement window T specifies the window size in terms of the number of samples of past traffic load. The T variable controls the amount of traffic history measurement to be used in MBAC approaches. In our PS scheme, T is determined by the chosen order p of the NLMS predictor. That is, $T = p$. Our simulation results are shown in Tables I and II for PS and MS respectively, in which S, the sampling period size, is $5e^3$ packet transmissions (125 bytes per packet) as used in the MS simulations [2] [1]. For this simulation, the utilization target is set to 0.95; the average flow interval is set to 400 ms and the average hold time for each flow is 300 seconds. This simulation environment setup will also be applied to the other simulations except when explicitly indicated otherwise.

Table II shows that for the MS scheme, as T increases, utilization drops dramatically and the loss rate also falls. Our simulation confirms the results from [1] that the value of measurement window T has a significant impact on the MS performance. When T becomes larger, the MS becomes significantly more conservative. More detail regarding the effects of tuning T in the MS can be found in [1].

In contrast, the T in the PS scheme does not have a significant effect on the performance; increasing T does not make PS more conservative (Theoretically, an increase in T for the PS algorithm would increase the accuracy of the prediction as more recent historical information is used.). This is important as the choice of the best value of T in the MS algorithm is a difficult task and depends on variables that can change rapidly in the network.

2) Sampling Period S : The sampling period size S variable controls the sampling rate at which the admission control algorithm executes in ns2. In MS [2] [1], a value of $5e^3$ packet transmissions with a packet size of 125 bytes was used to ensure that at least 100 packets were sampled per sampling period. We conducted simulations with different values of S (with the same packet size of 125 bytes) in the PS and MS algorithms and compare its effect on performance. T was set to 3 for both PS and MS. The simulation results are shown in Table III. For MS, as S increases, the drop rate decreases rapidly and finally goes to 0.0 and the utilization

TABLE III
EFFECT OF S ON PS VS. MS.

Algorithm	S	Drop Rate	Utilization
PS	1000	0.001966	0.958045
PS	5000	0.001880	0.956982
PS	10000	0.001639	0.950750
PS	50000	0.001850	0.950204
MS	1000	$6.373692e^{-5}$	0.916335
MS	5000	$4.592763e^{-6}$	0.889605
MS	10000	$7.749943e^{-8}$	0.881532
MS	50000	0.0	0.788966

TABLE IV
CHANGING UTILIZATION THRESHOLD IN PS VS. MS.

Algorithm	Target Utilization	Drop Rate	Utilization
PS	0.85	0.0	0.861158
PS	0.90	$8.05999e^{-5}$	0.909258
PS	0.95	0.001648	0.954684
MS	0.85	0.0	0.798944
MS	0.90	0.0	0.847587
MS	0.95	$4.59276e^{-6}$	0.889605

falls. However, the PS scheme is clearly less sensitive to S, where only slightly decreasing drop rate and utilization are observed. This indicates that the PS scheme is more robust to the effect of changing S while it gradually becomes more conservative as S increases.

3) Target Utilization η : For a given value of η, an MBAC scheme would set an operating point on the loss-load curve at which the system stays for the given network conditions. The system operating point can be adjusted by choosing appropriate value of η. For instance, if it is known that a utilization of 0.9 is the highest possible load at which acceptable loss rate can occur with an admission control scheme, then η can be set to 0.9, realizing this desirable operating point of the system. We further examine the "provisioning accuracy" of target utilization η regarding the actual system utilization achieved with respect to various provisioned target utilizations as shown in Table IV. The actual network utilization in the MS scheme is significantly less than the chosen target utilization η for every value of η tested. In contrast, with the PS scheme, the actual network utilization adheres closely to its provisioned target utilization.

4) Network Conditions: Intuitively, when the flows' average hold time increases, the dynamics of flows' departure and arrival is less intensive. Simulation results are listed in Table V, in which the target utilization is set to 0.9. The PS scheme is quite robust to the change of flows' average hold time in terms of the system's operating point (i.e., actual utilization and drop rate), while the MS scheme is sensitive to changes in traffic characteristics. This is expected because MS is supposed to perform better when the flows' dynamics are less intense.

IV. DISCUSSION AND CONCLUSION

Robustness: Although most MBACs are capable of searching the set of operating points in terms of QoS and utilization, a critical challenge for MBAC schemes is to reach and maintain a controllable and desirable operating point robustly under various network traffic conditions without excessive tuning. Indeed, robustness is the key for any MBAC scheme to be practical. As shown in Section III, our proposed PS scheme is insensitive to sampling period S and/or measurement window T. Furthermore, PS is not sensitive to traffic characteristics such as flows' average holding time. Therefore, the PS scheme demonstrates desirable robustness.

TABLE V
EFFECT OF FLOW HOLD TIME ON PS VS. MS.

Algorithm	Hold Time(s)	Drop Rate	Utilization
PS	300	$1.36249e^{-5}$	0.910320
PS	900	$1.26463e^{-5}$	0.909920
PS	3000	$1.15781e^{-5}$	0.908520
MS	300	0.0	0.869605
MS	900	0.0	0.880936
MS	3000	$1.11897e^{-5}$	0.900310

Operating Control: A desirable operating point for soft real-time services depends on application domain. In practice, an ideal MBAC scheme should be able to provide an explicit and reliable means to control performance. It turns out that with the PS scheme, a single parameter – target utilization η can elegantly serve this purpose. When η is set to a value, no matter how the traffic characteristics may change, the expected actual utilization closely adheres to the chosen η. This is significant as it provides network operator/administrator with a convenient means to accurately control the provisioning of the system's operating point.

Our proposed PS admission control is also simple and inexpensive to implement, and there is no need to tune the T and/or S parameters in its practical use. In particular, the T can be as small as 2, which dramatically minimizes the computing and memory overhead of the PS scheme and makes it a viable alternative to traditional MBAC schemes. Future work includes optimizing μ and introducing background network traffic into the simulation study.

REFERENCES

[1] S. Jamin, P. B. Danzig, S. J. Shenker, and L. Zhang, "A measurement-based admission control algorithm for integrated service packet networks," *IEEE/ACM Trans. Networking*, vol. 5, no. 1, pp. 56–70, 1997.
[2] S. Jamin, S. J. Shenker, and P. B. Danzig, "Comparison of measurement-based admission control algorithms for controlled-load service," in *Proc. INFOCOM '97*, p. 973.
[3] A. M. Adas, "Using adaptive linear prediction to support real-time vbr video under rcbr network service model," *IEEE/ACM Trans. Networking*, vol. 6, no. 5, pp. 635–644, 1998.
[4] L. Breslau, S. Jamin, and S. Shenker, "Comments on the performance of measurement-based admission control algorithms," in *Proc. INFOCOM 2000*, pp. 1233–1242.
[5] M. Grossglauser and D. N. C. Tse, "A time-scale decomposition approach to measurement-based admission control," *IEEE/ACM Trans. Networking*, vol. 11, no. 4, pp. 550–563, 2003.
[6] S. Haykin, *Adaptive Filter Theory Third Edition*. Upper Saddle River, NJ: Prentice-Hall, 1996.
[7] F. H. P. Fitzek and M. Reisslein, "Mpeg-4 and h.263 video traces for network performance evaluation," *IEEE Network*, vol. 15, no. 6, pp. 40–54, Nov.-Dec. 2001.