

# A GREEDY POLICY FOR THE ADMISSION CONTROL OF SELF-SIMILAR TRAFFIC

Cesar A. V. Melo and Nelson L.S. da Fonseca

**Abstract** - Admission control is a fundamental mechanism for the provision of Quality of Services in multiservice networks. The control of long-range dependent traffic demands specially designed mechanism. In this paper, a greedy policy for the admission control of self-similar traffic is introduced. This policy accepts flows into the network according to a decreasing order of revenue. It is shown that results produced by this new policy are close to the one produced by an optimum policy.

**Keywords:** Admission control, self-similar traffic, envelope process

**Resumo** - Controle de admissão é um mecanismo fundamental para o provimento de QoS em redes multimídia. Controlar tráfego que apresenta dependência de longa duração requer mecanismos especialmente projetados para este fim. Neste artigo, uma política gulosa para controle de admissão de tráfego que apresenta dependência de longa duração é introduzida. Essa política aceita novos fluxos na rede seguindo uma ordem decrescente de recompensa. Mostra-se que os resultados produzidos por esta nova política aproximam-se de resultados obtidos com uma política ótima.

**Palavras-chave:** Controle de Admissão, Tráfego Auto-similar, Processo envelope

## 1. INTRODUCTION

The advent of traffic integration has allowed the deployment of new multimedia services such as IP Telephony and Video-on-Demand. The provision of Quality of Service for such services requires special new traffic control mechanisms. Admission control decides on the acceptance of new flows into the network by considering the possibility of supporting the QoS required by them as well as the maintenance of QoS for the flows already admitted into the network.

Several studies [1]-[2] have claimed that different types of network traffic can be accurately modeled using a self-similar process. A self-similar process is able to capture the long-range dependence (LRD) phenomenon exhibited by such traffic. Moreover, a series of simulation and analytical studies [3]-[5] has demonstrated that the long-range dependencies may have a pervasive effect on queueing performance. In fact, there is clear evidence that it can potentially cause massive packet losses. Furthermore, such a queueing system suffers from the buffer inefficiency phenomenon [3]-[7]. Increasing the buffer size is not effective for decreasing

the buffer overflow probability significantly. Although various papers have analyzed the self-similar nature of traffic [1]-[2], control mechanisms for self-similar traffic have not yet been fully investigated.

The concept of equivalent bandwidth is intimately connected with admission control and associated service requirement [6]. The equivalent bandwidth of a connection (flow) is a characterization of the minimum required bandwidth of that connection such that QoS requirements are met. There has been a great interest in the concept of equivalent bandwidth, since it promises to bridge the design of statistical multiplexing networks to the familiar design of circuit-switched networks. Although there is a remarkable collection of results of equivalent bandwidth [6]-[9], very few results are available for traffic with long-range dependencies [7]-[8].

A new traffic model called the fractional Brownian motion (fBm) envelope process for the characterization of an LRD flow has been proposed [13]. A new framework for computing probabilistic delay bounds for a deterministic queueing system has also been developed. It was shown that the delay bounds were in agreement with known results obtained via large deviation theory. This traffic characterization has made a more intuitive understanding of the dynamics of queueing systems possible. Time-scales for the complete characterization of the behavior of a queueing system fed by self-similar flows has also been derived [14]. The accuracy of the fractional Brownian motion (fBm) envelope process for modeling real network traffic has been extensively validated [12].

Recent studies [10] [11] have shown that IP traffic, as well as video traffic, presents multi-scaling characteristics which are properly modeled by multifractal processes. For time scales larger than a cut-off value, IP and video traffic should be considered monofractal (self-similar). For IP traffic, the cut-off time scale is of the order of a round trip time (512ms to 1 sec); while for video traffic, the cut-off time scale is of the order of the transmission time of a frame.

Besides the provision of QoS requirements, network providers are mostly interested in increasing revenues obtained from the acceptance of new flows (connections). In this paper, a new admission control policy for flows presenting LRD is proposed. This policy accepts new flows in decreasing order of their associated revenue. Results given by this new policy are compared to ones provided by an optimal policy which maximizes the long-term revenue by using a knapsack type of solution. It is shown that the use of the proposed policy generates revenues close to the one generated by the optimal policy.

This paper is organized as follows. In Section II, an envelope process for self-similar traffic is presented. In Section III, the time scale of interest for a queueing system fed by a self-similar process is derived. In Section IV, the statistical

Cesar A. V. Melo and Nelson L.S. da Fonseca are with State University of Campinas Institute of Computing, Campinas, SP, Brazil. E-mails: cavneto@ic.unicamp.br, nfonseca@ic.unicamp.br.

multiplexing of heterogeneous self-similar flows is shown. In Section V, a new admission control policy is introduced. Finally, Section VI presents the conclusions.

## 2. A FRACTAL BROWNIAN MOTION ENVELOPE PROCESS

It is well known that for a Brownian motion (Bm) process  $A(t)$  with mean  $\bar{a}$  and variance  $\sigma^2$ , the envelope process  $\hat{A}(t)$  can be defined as follows [15]:

$$\hat{A}(t) = \bar{a}t + \kappa\sqrt{\sigma^2 t} = \bar{a}t + \kappa\sigma t^{1/2}.$$

The parameter  $\kappa$  determines the probability that  $A(t)$  will exceed at time  $t$ . Since  $A(t)$  is a Brownian motion process, the following is also true:

$$P\left(\frac{A(t) - \bar{a}t}{\sigma t^{1/2}} > \kappa\right) = \Psi(\kappa),$$

where  $\Phi(y)$  is the residual distribution function of the standard Gaussian distribution. Using the approximation  $\Psi(y) \approx (2\Pi)^{-1/2}(1+y)^{-1} \exp(-y^2/2) \approx \exp(-y^2/2)$   $\kappa$  is found such that  $\Phi(\kappa) \leq \epsilon$ . Hence,  $\kappa$  is given by  $\kappa = \sqrt{-2 \ln \epsilon}$ .

It is claimed here that  $\Pr(A(t) > \hat{A}(t)) \approx \epsilon$ , where  $k = \sqrt{-2 \ln \epsilon}$ . This approach can then be extended to deal with LRD traffic. Let be a fractional Brownian motion process with mean  $\bar{a}$ . Hurst's law states that the variance of the increment of this process is given by  $Var[A_H(t+s) - A_H(t)] = \sigma^2 s^{2H}$ , where  $H \in [1/2, 1)$  is the Hurst parameter. A fBm envelope process can then be defined by the following:

$$\hat{A}_H(t) \stackrel{\text{def}}{=} \bar{a}t + \kappa\sqrt{\sigma^2 t^{2H}} = \bar{a}t + \kappa\sigma t^H. \quad (1)$$

The Brownian motion envelope process is only a special case of  $H = 1/2$ . Similarly,  $\kappa$  determines the probability that  $A_H(t)$  exceeds  $\hat{A}_H(t)$ . Moreover, since the process exhibits LRD, if  $A_H(t)$  exceeds  $\hat{A}_H(t)$  at time  $t$ , it is possible that it will stay above it for a long period of time.

It should be noted that the flow does not necessarily need to be self-similar in order to fit this characterization; what is important is that it matches the behavior of the envelope process for the time scale of interest. The accuracy of the fBm envelope process representation is investigated by inspecting how well it can model the worst-case behavior of real network traffic. Assuming that the input traffic is characterized by a trace with  $N$  sample points, defined by  $A(t)$ , where  $A(t)$  represents the cumulative number of cells arriving up to time  $t, t \in [1, 2, \dots, N]$ , a very simple method for computing the fBm envelope process parameters for this trace is proposed, which computes the optimal envelope process of the trace. The advantage of this approach lies in the fact that it is not necessary to estimate the Hurst parameter of the trace accurately. The optimal envelope process for this trace is defined by  $Y(t-s) = \max_{s < t} (A(t) - A(s))$ . Assume that the process is stationary, yields  $Y(\tau), \tau = t - s$ , which defines the maximum number of cells arriving in a given interval of size  $\tau$ . Therefore, the fBm envelope process parameters  $\hat{A}(\cdot)$  can be chosen to match the behavior of  $Y(\cdot)$ .

Figure 1 shows the accuracy of the fBm envelope when compared to Bellcore's LAN trace. The sample average arrival rate and the sample variance are computed and inserted

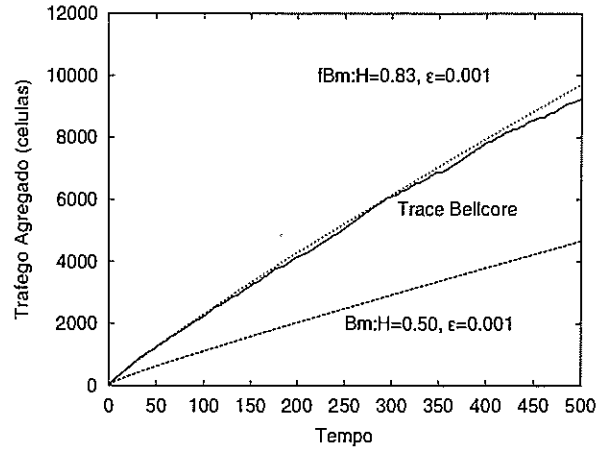


Figure 1. The Accuracy of the fBm Envelope Process

in Equation 1. The optimal envelope process, i.e  $Y(\cdot)$ , is then computed and  $H$  chosen so that  $\hat{A}(\cdot)$  matches the behavior of  $Y(\cdot)$ . This example shows that the fBm envelope process matches closely the behavior of the LRD traces. The effectiveness of the fractal Brownian motion envelope process was extensively validated by utilizing synthetic traces as well as traces derived in real networks. A report on the validation can be found in [12].

## 3. TIME SCALE OF INTEREST

This section characterizes the time until a queue reaches its maximum occupancy, in a probabilistic sense. The queue size at this time provides a simple delay bound [13]. A rigorous mathematical derivation of the delay bound can be found in [14], and a heuristic derivation is introduced in keeping with the precepts explored here. Consider a continuous-time queueing system, with deterministic service given by  $c$ . The cumulative arrival process is represented by  $A_H(t)$ . Let  $\hat{A}_H(t)$ , continuous and differentiable, be the probabilistic envelope process of  $A_H(t)$ , such that  $\Pr(A_H(t) > \hat{A}_H(t)) \leq \epsilon$ .

During a busy period, which starts at time 0, the number of cells in the system at time  $t$  is given by  $q(t)$ . Thus,  $q(t) = A_H(t) - ct \geq 0$ .

By defining  $\hat{q}(t)$  as

$$\hat{q}(t) = \hat{A}_H(t) - ct \geq 0, \quad (2)$$

we can see that  $\Pr(q(t) > \hat{q}(t)) = \Pr(A_H(t) > \hat{A}_H(t)) \leq \epsilon$ .

The maximum delay in a FIFO queueing system is given by the maximum number of cells in the queue during the busy period, which can be defined as follows:

$$q_{max} \stackrel{\text{def}}{=} \max(\hat{q}(t)), t \geq 0.$$

Therefore,

$$\Pr(q(t) > q_{max}) \leq \Pr(q(t) > \hat{q}(t)) \leq \epsilon,$$

$$\Pr(q(t) > q_{max}) \approx \epsilon.$$

The queue length at time  $t$ ,  $q(t)$ , will thus only exceed the maximum queue length  $q_{max}$  with probability  $\epsilon$ . In other words, only when the arrival process exceeds the envelope process, will the maximum number of cells in the system exceed the estimated value. Intuitively, by bounding the behavior of the arrival process, it is possible to transform the problem of obtaining a probabilistic bound for the stochastic system defined by  $q(t) = A_H(t) - ct \geq 0$ , into an easier problem of finding the maximum of a deterministic system, described by  $\hat{q}(t) = \hat{A}_H(t) - ct \geq 0$ .

For the fBm process, the envelope process defined above is inserted into Equation 2, which gives the following:

$$\hat{q}(t) = \hat{A}_H(t) - ct = \bar{a}t - \kappa\sigma t^H - ct. \quad (3)$$

In order to compute  $q_{max}$ , it is necessary to find  $t^*$  such that

$$\frac{d\hat{q}(t^*)}{dt} = 0$$

or equivalently,

$$\frac{d\hat{A}_H(t^*)}{dt} = c. \quad (4)$$

Hence,  $t^*$  is given by

$$t^* = \left[ \frac{\kappa\sigma H}{(c - \bar{a})} \right]^{\frac{1}{1-H}}.$$

The time scale of interest is defined by the time until a queue size reaches its peak, i.e.,  $t^*$ . This is denoted the Maximum Time Scale (MaxTS), and it defines the point in time where the unfinished work in the queueing system achieves its maximum in a probabilistic sense. This means that the average arrival rate has dropped below the link capacity so that the queue size will start decreasing. The average arrival rate converges to mean arrival rate of the flow on the basis of the law of large numbers. Consequently, there is no need to worry about anything other than the time scale for which the flow rate still exceeds the link capacity. In other words, after a period of time, the probability that the average arrival rate exceeds the link capacity is negligible; hence, the arrival model need no longer reproduce the flow behavior for these time scales. This is the most important time scale in terms of traffic modelling. As a rule of thumb, to choose the parameters of an input flow to match the fBm envelope process, it is necessary to find MaxTS analytically, while the parameters of the fBm process must be chosen to match the flow optimal envelope process of the flow at MaxTS.

Substituting  $t^*$  back into Equation 2, it can be concluded that:

$$q_{max} = \hat{A}_H(t^*) - ct^*, \quad (5)$$

$$q_{max} = (c - \bar{a})^{\frac{H}{H-1}} (\kappa\sigma)^{\frac{1}{1-H}} H^{\frac{H}{1-H}} (1 - H).$$

Since the fBm process does not exceed with probability  $1 - \epsilon$ , the maximum number of cells will be bounded by  $q_{max}$  with the same probability.  $\hat{c}$  must be found such that  $q_{max}$  is equal to  $K$ . In other words, a buffer of size  $K$  will overflow with probability  $\epsilon$  if the link capacity is  $\hat{c}$ . Therefore,  $\hat{c}$  is given by

$$\hat{c} = \bar{a} + K^{\frac{H-1}{H}} (\kappa\sigma)^{1/H} H (1 - H)^{\frac{H-1}{H}}.$$

This result was previously obtained by Norros [4] [19] and Duffield [5]. However, the framework here makes it possible to achieve the same results predicted by the large deviation theory without making restrictive assumptions. It has also been possible to reduce the sensitivity of the estimation process by using a bound rather than attempting to estimate the parameters directly from the full trace.

A comparison between the MaxTS and the so-called Critical Time Scale (CTS) [20] is provided in [14]. It is shown that this CTS is the most likely time scale when buffer overflow occurs; but, it does not furnish information about the marginal distribution of the random variables which describes the overflow process.

#### 4. STATISTICAL MULTIPLEXING OF SELF-SIMILAR FLOWS

In this section, MaxTS is used to derive expressions for predicting the equivalent bandwidth and buffer requirements of an aggregate of self-similar flows. Essentially, a method for computing the bandwidth necessary to support requirements of buffer overflows is proposed, as well as for determining the maximum probabilistic delay for an aggregate of heterogeneous flows. The problem in this section can be stated as follows:

*Given a set of flows with mean  $\bar{a}_i$ , standard deviation  $\sigma_i$  and Hurst parameter  $H_i$ , what is the link capacity needed so that the maximum queue size will be bounded by  $q_{max}$  with probability  $\epsilon$ ?*

Assume  $N$  independent flows  $A_H^i(t)$  defined by the following parameters: mean  $\bar{a}_i$ , standard deviation  $\sigma_i$  and Hurst parameter  $H_i$  for  $i \in [1, N]$ . Let the aggregate traffic be denoted by  $A_H(t) = \sum_{i=1}^N A_H^i(t)$ . The envelope process of each flow is given by  $\hat{A}_H^i(t)$ , and the envelope process of the aggregate traffic is given by  $\hat{A}_H(t)$ . The  $q_{max}$  of a queue with heterogeneous flows can be computed by finding  $t^*$  for the envelope process of the aggregate stream. The mean of the aggregate traffic is given by the sum of the mean of the individual flows. Similarly, since the flows are independent, the variance of the aggregate traffic is also furnished by the sum of the variance of the individual flows. Hence, the envelope process of the aggregate traffic is defined as follows:

$$\hat{A}_H(t) = \sum_{i=1}^N \bar{a}_i t + \kappa \left( \sum_{i=1}^N \sigma_i^2 t^{2H_i} \right)^{1/2}.$$

Replacing in Equation 4, gives the following:

$$\kappa \frac{1}{2} \left( \sum_{i=1}^N \sigma_i^2 t^{2H_i} \right)^{-1/2} \left( \sum_{i=1}^N \sigma_i^2 2H_i t^{2H_i-1} \right) = c - \sum_{i=1}^N \bar{a}_i. \quad (6)$$

Equation 6 can be solved numerically to find  $t^*$ , which is then inserted in Equation 5 to compute  $q_{max}$ . Moreover,

combining Equations 5 and 6 results in the following:

$$\kappa \frac{1}{2} \left( \sum_{i=1}^N \sigma_i^2 t^{2H_i} \right)^{-1/2} \left( \sum_{i=1}^N \sigma_i^2 2H_i t^{2H_i-1} \right) - \kappa \left( \sum_{i=1}^N \sigma_i^2 t^{2H_i-2} \right)^{1/2} + \frac{q_{max}}{t} = 0. \quad (7)$$

Using Equations 6 or 7 makes it possible to answer the fundamental question posed at the beginning of this section. For the special case of multiplexing  $N$  identical flows, the envelope process is given by:

$$\hat{A}_H(t) = N\bar{a}t + \sqrt{N}\kappa\sigma t^H.$$

since the Hurst parameter is preserved when aggregating  $N$  identical flows [18]. In this case, Equation 6 is reduced to the following:

$$\frac{k(N\sigma^2 2H t^{2H-1})}{2(\sqrt{N}\sigma t^H)} = N(c - \bar{a}).$$

Using the same approach as above, it is possible to find  $t^*$  and  $q_{max}$ :

$$t^* = \left[ \frac{\sqrt{N}\kappa\sigma H}{N(c - \bar{a})} \right]^{\frac{1}{1-H}} = N^{\frac{1}{2(H-1)}} t_i^*,$$

$$q_{max} = N(\bar{a} - c)N^{\frac{1}{2(H-1)}} t_i^* + N^{\frac{H}{2(H-1)}} N^{1/2} \kappa\sigma (t_i^*)^H = N^{\frac{(H-1/2)}{H-1}} \hat{q}_{max},$$

$$t_i^* = \left[ \frac{\kappa\sigma H}{(c - \bar{a})} \right]^{\frac{1}{1-H}},$$

$$\hat{q}_{max} = \hat{A}(t^*) - ct_i^*.$$

where  $t_i^*$  and  $\hat{q}_{max}$  corresponds to a queueing system fed by a single flow.

This specific case of a single flow is shown first. Figure 2 shows the overflow probability as a function of the maximum buffer size for a link utilization of 60% and for  $H = 0.63$  and  $\sigma^2 = 0.64$ . When the overflow probability given by the analytical models is compared to that with the overflow probability observed in the simulation experiments, it can be seen that the two models provide reasonably consistent results. It can be observed that the precision of results found here increases with the Hurst parameter as well as with the link utilization.

To evaluate the effectiveness of the equivalent bandwidth expressions (Equations 6/7), it is necessary to define multiplexing gain as the ratio between  $N$  times the equivalent bandwidth of a single flow and the equivalent bandwidth of  $N$  identical flows. It can be seen that a significant multiplexing gain can be achieved when multiplexing homogeneous flows. In Figure 3, the gain for a link capacity of 150 Mbps, and flows with a mean arrival rate of 1.1Mbps for various Hurst parameters is plotted. Figure 3 considers flows with  $\sigma^2 = 0.3$ . Results indicate that the multiplexing gain also increases with the Hurst parameter, especially for streams with moderate to high variance. This can be explained by the

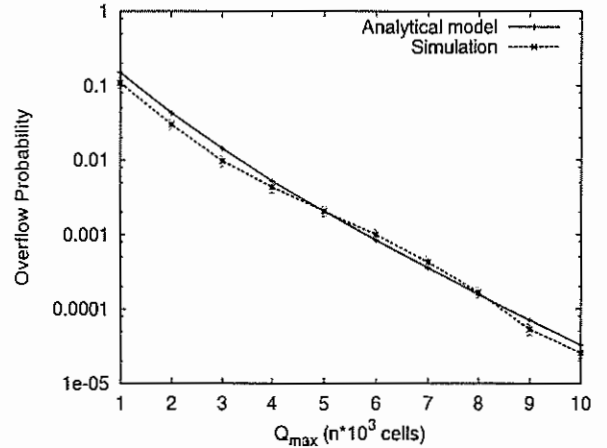


Figure 2. Accuracy of the Predicted Overflow Probability by Equations 6/7 for a Single flow as a Function of the Buffer Size

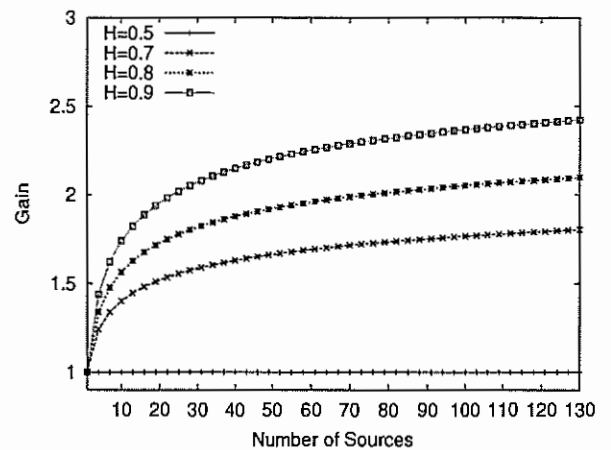


Figure 3. Multiplexing Gain Obtained by Using Equations 6/7

fact that Equations 6/7 take into consideration the existence of long periods with no arrivals in streams with high Hurst parameters. As a consequence, Equations 6/7 demand less bandwidth when multiplexing several flows than would be necessary in a non-multiplexing approach. It was also shown that the multiplexing gain increases not only with the variance and the Hurst parameter, but also with that of the number of multiplexed flows.

A comparison of the equivalent bandwidth expressions shown here with those derived by Kelly [7], and by Stathis and Maglaris [21] is made, using Equation 3.35 in [7] and Equation 6 in [21]. The real flow traffic parameters as described in [21] are utilized. In Figure 4, the number of accepted connections as a function of the overflow probability are presented, considering flows with a mean rate of 1.4 Mbps,  $\sigma = 0.28$  Mbps and an  $H = 0.85$ . The buffer size was set to 1000 ATM cells. It can be observed that the number of flows admitted as predicted by the present approach is the same as that predicted by Kelly (overlapping curves). In other words, the same result predicted by the large deviation theory is achieved with much less computational effort.

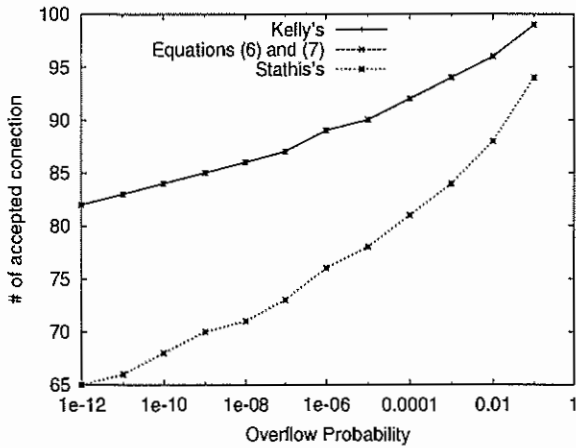


Figure 4. Comparison of the Number of Admitted Flows Given Equations 6/7 and the Work in [7] and [21] for Different Overflow Probability Requirements

## 5. A GREEDY POLICY FOR ADMISSION CONTROL

Whenever a request for admission arrives, an admission controller needs to verify both if it will be possible to provide the required QoS and to continue to furnish QoS to all other already accepted flows. In order to maximize revenue, the admission controller should collect requests during time intervals, to decide which requests should be granted.

The policy proposed here accepts requests according to their decreasing order of associated revenues. However, this simple policy may not maximize the total revenue, since it considers arrivals during short periods and not in the long run. In the long run, a single flow with an associated high revenue can consume the same amount of resources and provide lower total revenue than a set of flows each with an associated low revenue. An optimal policy should find a combination of flows that, in the long run, gives the highest possible revenue subject to the available bandwidth. This is a classic knapsack problem in which the available bandwidth is the knapsack, and the bandwidth demand of each flow is the size of each object. This knapsack problem can be stated as:

$$\max \sum_{i=1}^M R_i x_i$$

$$\sum_{i=1}^M \bar{a}_i t^* x_i + \kappa \left( \sum_{i=1}^M \sigma_i^2 t^{*2H_i} x_i \right)^{1/2} \leq q_{max}, \quad (8)$$

$$x_i \in \{0, 1\},$$

where  $R_i$  - is the revenue of flow  $i$ ;  $H_i$  - is the Hurst parameter of flow  $i$ ;  $\bar{a}$  - is the mean arrival rate of flow  $i$ ;  $\sigma_i$  - is the standard deviation of flow  $i$ ;  $t^*$  - is the Maximum Time Scale (MaxTS);

To verify the extent to which the proposed greedy policy leads to the same results produced by the knapsack approach, extensive simulation experiments have been carried out. The revenue function used is  $aT + bV$ , where  $T$  is time duration,

|   | $\bar{a}$ | $\sigma$ | H    |
|---|-----------|----------|------|
| A | 0.16      | 1.01     | 0.67 |
| B | 0.12      | 1.12     | 0.78 |
| C | 0.2       | 0.9      | 0.85 |
| D | 0.22      | 0.84     | 0.91 |

Table 1. Traffic parameters of video streams.

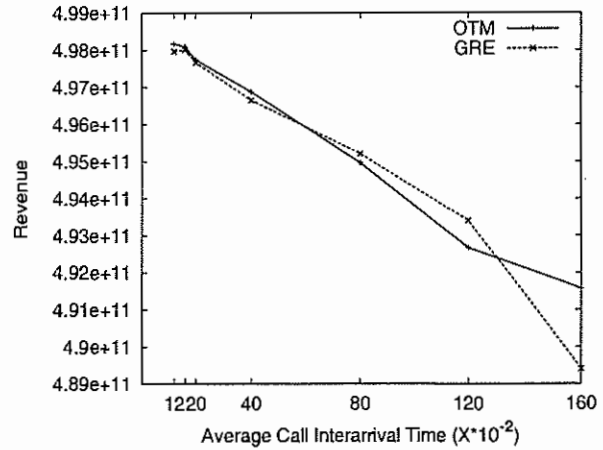


Figure 5. Revenue x Average Flow Interarrival Time

of a connection and  $V$  is the traffic volume [23][21]. Figures 5 and 6 show typical findings. In these examples, traces with the traffic parameters shown in Table 1 were used. Note that the mean arrival rate is normalized to the channel capacity. The mean duration of a flow varies between 20 and 500 seconds, and the buffer size corresponds 1000 ATM cells.

In Figure 5, the total revenue is displayed as a function of the mean interarrival time of the flows. The total revenue is computed by the accumulated revenue of all accepted flows during the simulation experiment. The exact value of the revenue is immaterial in the sense that it changes according to the choice of  $a$  and  $b$ . The lower the interarrival time, the higher the load. Note that for moderate to high loads, the revenue produced by the greedy policy is close to the one produced by the optimal approach. Under low load conditions, this trend is not observed.

In Figure 6, the total revenue as a function of the mean flow duration is shown for a utilization value of 0.7. It can be seen that the greedy policy provides revenues close to the one given by the optimal approach irrespective of the flow duration (for moderate to high loads).

## 6. CONCLUSIONS

Admission control is a fundamental mechanism for the provision of Quality of Service in multiservice networks such as IP Diffserv and ATM networks. Such provision is sensitive to the long-range dependencies existing in several types of traffic. In this paper, a simple admission control policy was introduced. This new policy admits flows into the network according to the decreasing order of revenue provided by these flows. It was shown that for moderate to high loads this simple policy produces revenues in the same range as rev-

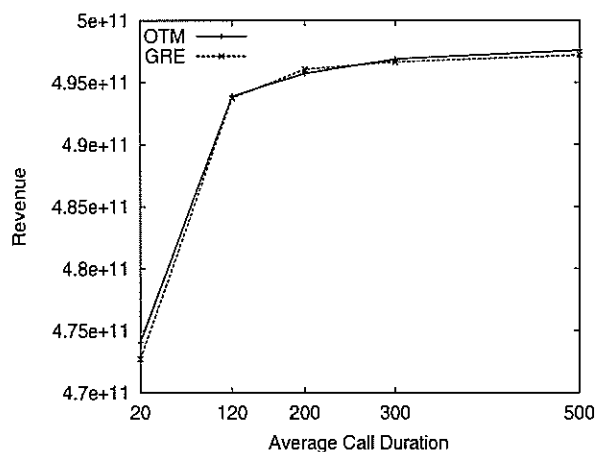


Figure 6. Revenue x Average Flow Duration

enues produced by an optimal policy. The simplicity and ease of policy implementation introduced here make it a potential candidate for realtime admission control in QoS-oriented networks.

## 7. ACKNOWLEDGEMENTS

This work was partially sponsored by CNPq (process # 30006495-0) and FAPESP(process # 00 09772-6).

## REFERENCES

- [1] W. Leland, M. Taqqu, W. Willinger and D. Wilson, "On the Self-Similar Nature of Ethernet Traffic (Extended Version)", *IEEE/ACM Transaction on Networking*, vol 2, no 1, pp. 1-15, February 1994.
- [2] V. Paxson and S. Floyd, "Wide-Area Traffic: The Failure of Poisson Modeling", *IEEE/ACM Transaction on Networking* vol 3(3), pp. 226-244, 1995.
- [3] G. Mayor and J. Silvester, "An ATM Queueing System with a Fractional Brownian Noise Arrival Process", in *Proc. of IEEE International Conference Communications*, pp. 1607-1611, 1996.
- [4] I. Norros, "A Storage Model with Self-Similar Input", *Queueing Systems* 16, pp. 387-396, 1994.
- [5] N. Duffield, J. T. Lewis, N. O'Connell, R. Russel, F. Toomey, "Predicting Quality of Service with Long-range Fluctuations", in *Proc. of IEEE International Conference Communications*, pp. 473-477, 1995.
- [6] A. I. Elwalid and D. Mitra, "Effective Bandwidth of general Markovian Traffic sources and Admission Control of High Speed Networks", *IEEE/ACM Transaction on Networking*, 1(4), pp. 329-343, 1993.
- [7] F. Kelly, "Notes on Effective Bandwidth", in *Stochastic Networks: Theory and Applications*, F. Kelly, S. Zachary and I. Ziednis ed., Oxford Press, 1996.
- [8] N. G. Duffield and N. O'Connell, Large Deviation and Overflow probabilities for the general single-server with application, in *Proc. Cambridge Philosophical Soc.*, pp. 363-374, 1995.
- [9] G. Kesidis, J. Walrand and C.S. Chang, Effective Bandwidth for Multiple Class markov Fluid and other ATM Sources, *IEEE/ACM Transaction on Networking*, pp. 424-428, Aug, 1993.
- [10] A. Erramilli and O. Narayan and A. Neidhardt and I. Saniee, "Multi-scaling models of sub-frame VBR video and TCP/IP traffic", *IEEE Journal of Communications and Networks*, 3(4) December 2001
- [11] A. Erramilli and O. Narayan and A. Neidhardt and I. Saniee, "Performance impacts of multi-scaling in wide-area TCP/IP traffic". In *Proc IEEE INFOCOM 2000* pages 352-359.
- [12] N. L. Fonseca and G. S. Mayor and C. A. V. Neto, "On the Equivalent Bandwidth of Self-Similar Sources", *ACM Transactions on Modeling and Computer Simulation*, 104-124, April 2000, 10(2).
- [13] G. S. Mayor and J.A. Silvester, "Time Scale Analysis of an ATM Queueing System with Long-range Dependent Traffic", in *Proc of Infocom'97*, pp. 205-212, 1997.
- [14] G.S. Mayor and J. A. Silvester, "Providing QoS for Long-Range Dependent Traffic", the 7th IEEE Computer-aided Modeling Analysis and Design of Communications Links and Networks, pp. 19-28, 1998.
- [15] A. Papoulis, Probability, Random Variables, and Stochastic Processes, McGraw-Hill, 1991.
- [16] M. Chi, E. Neal and G. Young, "Practical Application of Fractional Brownian Motion and Noise to Synthetic Hydrology", *Water Resources Research*, Vol. 9, pp. 1523-1533, December 1973.
- [17] B.B. Mandelbrot, "Long-run Linearity, Locally Gaussian Process, h-spectra and Infinite Variance, *International Economic Review*, 10, pp. 82-113, 1969.
- [18] A. Erramilli, O. Narayan and W. Willinger, "Experimental Queueing Analysis with Long-range Dependent Packet Traffic". *IEEE/ACM Transactions on Networking*, 4(2), pp. 209-223, April, 1996.
- [19] I. Norros, "The management of large Flows of Connectionless traffic on the basis of Self-similar, in *Proc. of IEEE International Conference Communications*, pp. 451-455, 1995.
- [20] B. Ryu and A. Elwalid, "The Importance of Long-Range Dependence of VBR Video Traffic in ATM Traffic Engineering: Myths and Realities", in *Proc. of ACM SIGCOMM*, pp. 13-14, 1996.
- [21] M. Falkner, M. Devetsikiotis, and I Lambadaris, "An Overview of Pricing Concepts for Broadband IP Networks", *IEEE Communications Surveys & Tutorials*
- [22] C.Stathis and B. Maglaris, "Modeling the self-similar behavior of network traffic", *IFIP 6th Workshop on Performance Modeling and Evaluation of ATM Networks*, 1998.
- [23] C. Courcoubetis, V. A. Siris and G.D. Stamoulis, "Comparing Usage-Based Pricing Schemes for Broadband Networks", *IEE Colloquium on "Changing for ATM - the Reality Arrives"*, London, UK, Nov 1997.

**Cesar A V Melo** received his B.Sc. degree in Computer Science from the University of Amazonas and his M.Sc. degree in Computer Science (1999) from the State University of Campinas. He is currently is pursuing a Ph.D. degree in Computer Science at the State University of Campinas - Brazil.

**Nelson L.S. da Fonseca** received his Electrical Engineer (1984) and MSc. in Computer Science (1987) degrees from The Pontifical Catholic University at Rio de Janeiro, Brazil, and the MSc (1993) and Ph.D (1994) degrees in Computer Engineering from The University of Southern California. Since 1995, he has been affiliated with the Institute of Computing of The State University of Campinas, Campinas - Brazil where is currently an Associate Professor. Currently, he is Director of on-line Services of the IEEE Communications Society and ice-chairman for South America of the IEEE ComSoc Technical Committee on Multimedia Communica-

tions, and vice-chair of ComSoc Communications Systems Integration and Modeling technical committee He served as editor-in-chief of IEEE Communications Newsletter (1999-2002), and is currently an editor for Revista Brasileira de Telecomunicações, IEEE Transactions on Multimedia Computer Networks, and IEEE Communications Magazine. He is the recipient of Elsevier Editor of the Year 2000, USC International Book award and of the Brazilian Computing Society First Thesis and Dissertations award. His graduate student have received five awards in Latin-American Thesis Contests.