



Title	A measurement-based congestion alarm for self-similar traffic
Author(s)	Chan, TK; Lau, WC; Li, VOK
Citation	IEEE International Conference On Communications, 2001, v. 5, p. 1528-1533
Issued Date	2001
URL	http://hdl.handle.net/10722/46242
Rights	©2001 IEEE. Personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to reuse any copyrighted component of this work in other works must be obtained from the IEEE.

A Measurement-based Congestion Alarm for Self-Similar Traffic

Tat-Keung Chan

Wing-Cheong Lau

Victor O. K. Li

Department of Electrical and Electronic Engineering

The University of Hong Kong

Pokfulam, Hong Kong

email: {tatkeung,wclau,vli}@eee.hku.hk

Abstract—Self-similar traffic is distinguished by positive correlation, which can be exploited for better traffic management. Inspired by measurement-based admission control schemes, a measurement-based congestion alarm is proposed. The aggregate traffic at an output port of a switch or router in a high-speed network is modeled by a fractional Gaussian noise process. Traffic measurements are performed in regular time intervals to determine the current traffic loading. This information is then used to predict the loading situation in the near future. If congestion is likely to occur, a congestion alarm is set off and appropriate network management functions taken to alleviate the possible congestion. The above constitutes a closed loop feedback control mechanism that maintains high resource utilization. Simulation results show that the proposed scheme, when used with dynamic bandwidth allocation, reduces bandwidth requirements by more than 20%.

I. INTRODUCTION

Traffic measurements [1], [3], [12], [14], [16] over the past few years have shown a common phenomena about packet traffic over a wide variety of modern networks: they are self-similar. One distinctive characteristic of self-similar traffic is the existence of positive correlation over long period of time. This nontrivial correlation can be exploited such that future traffic load can be predicted statistically. In [7], the authors studied the continuous-time optimal predictor for the fractional Brownian motion (fBm) process. From the analysis, a rule of thumb is found which says that one should predict the next second with the latest second, the next minute with the latest minute and so on. In other words, if one wants to predict the next second, history of more than one second does not reduce the prediction error significantly. The result for the continuous-time fBm is of theoretical interests. In practice, only the samples of the continuous-time process are available. Hirchoren et al.[9] considered this practical discrete-time version, the optimal predictor for the increment process of fBm, that is, the fractional Gaussian noise (fGn). They verified the rule of thumb mentioned. This discrete-time estimator provides a practical tool for traffic prediction.

Duffield et al. [4] proposed a framework for network design and control based on the likelihood that the aggregate demand will exceed capacity, where the demand is predicted using a specification of the sources and their corresponding source models. In this paper, we consider an alternative approach whereby *a priori* traffic specification is not required for sources. This is made possible by employing the idea of traffic measurement in measurement-based admission control schemes [2], [6], [10], [11]. Consider an output port of a switch or router in a high-speed network, the actual aggregate traffic load is measured periodically. Modeling this aggregate traffic by an fGn process and using the measured traffic load in the past, the traffic load in the near future is predicted. If

congestion is likely to occur, appropriate network management functions will be taken to alleviate the congestion. In particular, we study the use of dynamic bandwidth allocation with the measurement-based congestion prediction. Simulation results show that the proposed scheme reduces bandwidth required by about 20%.

This paper is organized as follows. Section II explains the system model of the measurement-based congestion alarm and discusses the fGn predictor. In Section III, the congestion alarm is modified to work with dynamic bandwidth allocation. Simulation results are presented in Section IV. Section V concludes this paper.

II. SYSTEM MODEL

To capture the characteristics of self-similar traffic, we propose to model the aggregate traffic as seen at an output port of a network router or switch by a fractional Gaussian noise (fGn) process. Although a single source may not be modeled accurately by the fGn process, the aggregation of a large number of sources can be. The fGn model is a parsimonious self-similar model that can be defined by three parameters, namely the mean, the peak, and the Hurst parameter. Figure 1 shows the system architecture of the proposed measurement-based congestion alarm. Basically, we model the aggregate traffic by an fGn process. Traffic measurements are performed in short time intervals periodically. These traffic measurements will then be used to predict the traffic intensity in the near future using the fGn predictor. If it is found that the system capacity is insufficient to handle the expected amount of traffic, the congestion alarm will be set off. When this happens, appropriate network control functions are carried out so as to alleviate the possible congestion.

One of the possible network management function that can be used is Call Admission Control (CAC). In case of possible congestion, the CAC function can start to drop new calls that arrive to the system. In this study, however, we consider the use of dynamic bandwidth allocation whereby the system bandwidth is renegotiable. Therefore, in case of possible congestion, the system can request larger bandwidth. When the amount of traffic is predicted to be smaller, excess bandwidth can be relinquished accordingly. In the following, we will discuss the predictor for fGn. The best predictor is actually the conditional mean given the past history. Moreover, the random variable under prediction is actually Gaussian distributed given the past history. As a result, we could construct the congestion alarm based on statistical guarantee that the probability of the network load being more than the system capacity is less than a predefined threshold. In the following, we first describe the optimal predictor for fGn, and then explain the congestion alarm that is constructed based on this predictor.

This research is supported in part by the University Grants Committee, Hong Kong, Area of Excellence in Information Technology, Grant No. AOE 98/99.EG01.

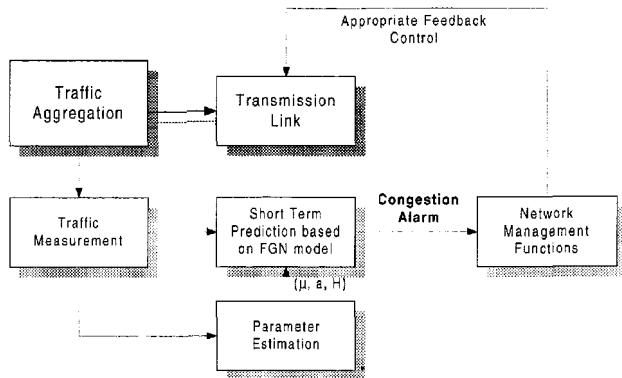


Fig. 1. System model for the feedback mechanism with the congestion alarm.

A. Predictor for Fractional Gaussian Noise

Consider a discrete time fractional Gaussian noise (fGn) process whose sample path in the past is known. Based on this known history, we want to predict a sample in the future. In [9], the authors derived the best estimator of the future value of a discrete time fGn process, which is actually the conditional mean of the future value conditioned on the past history. Let $\{Z_0, Z_1, \dots, Z_{m-1}, Z_i\}, n \geq m$ be a stationary discrete time Gaussian process with mean μ , variance σ^2 and autocorrelation function given by $R(k)$ ($R(0) = \sigma^2$). Denote $\{Z_0, Z_1, \dots, Z_{m-1}\}$ by \mathbf{Z}_m . The vector \mathbf{Z}_m represents the past history of the process. Z_i with $i \geq m$ is the future value that we want to predict. It can be shown that the best estimator is given by the conditional mean, $E[Z_i|\mathbf{Z}_m]$. In fact, it is a well-known result that conditioned on \mathbf{Z}_m , Z_i is still Gaussian. In particular, the conditional mean and variance are given by:

$$\mu_{Z_i|\mathbf{Z}_m} = \mathbf{R}_i^T \Sigma_m^{-1} \mathbf{Z}_m + (1 - \mathbf{R}_i^T \Sigma_m^{-1} \mathbf{1}) \mu \quad (1)$$

and

$$\sigma_{Z_i|\mathbf{Z}_m}^2 = R(0) - \mathbf{R}_i^T \Sigma_m^{-1} \mathbf{R}_i \quad (2)$$

respectively, where

$$\Sigma_m = \begin{bmatrix} R(0) & R(1) & \dots & R(m-1) \\ R(1) & R(0) & \dots & R(m-2) \\ \vdots & \vdots & \ddots & \vdots \\ R(m-1) & R(m-2) & \dots & R(0) \end{bmatrix},$$

$$\mathbf{R}_i = [R(i)R(i-1) \dots R(i-m+1)]^T$$

and $\mathbf{1}$ is an i by one vector whose elements are all equal to one.

The predictive power is small for small Hurst parameter, however, it increases rapidly. For $H = 0.95$, the predictor is able to resolve half of the variance of Z_i . To illustrate the utility of the predictor, Figure 2 shows an example with a simulated fGn trace whose Hurst parameter is equal to 0.9. The top figure shows the trace and the predicted values together with the 95% confidence intervals based on the first 64 sample points of the trace. It can be seen that as we predict further ahead, the variance (and hence the confidence interval) gets larger. The middle figure shows the case when we predict only ten samples based on a moving window of 64 samples.

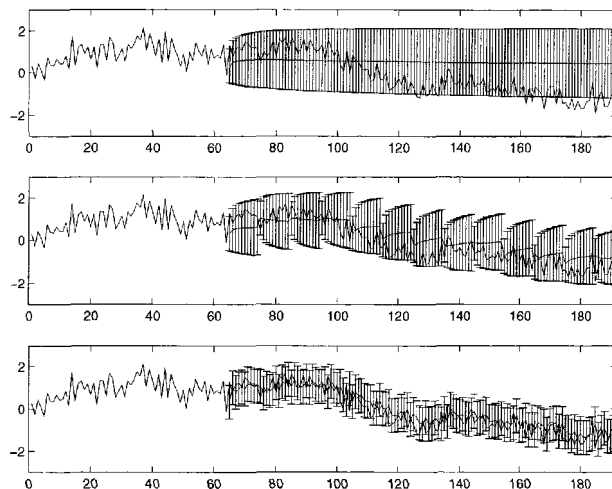


Fig. 2. An example of short term prediction for fGn. ($H = 0.9$)

This tracks the trajectory of the trace better. The bottom figure shows the extreme case whereby only one sample is predicted based on a history of 64 samples. It can be seen that the prediction is able to track the trace closely. However, the final scenario is of less practical interest. If we predict one sample ahead only, even if congestion is highly probable in the coming time slot, we may not have time to respond to such congestion. So in practice, we want to predict a short period of the future whose length should be determined by the time scale on which the appropriate traffic management is capable of responding. The time scale may range from hundreds of millisecond to seconds, depending on the particular traffic management function and the supporting network.

B. Congestion Alarm

Based on the short term prediction of the fGn process described in the previous section, we design a congestion alarm as follows. The aggregate traffic at an output port of the switch or router under consideration is modeled by a discrete-time fGn process. Every time slot, the traffic intensity is measured. A moving window of m samples of historical values is used to predict all of the coming n samples in the future. We want to make sure that the probability of the aggregate traffic exceeding the capacity, also called the *outage probability*, is less than a preset threshold θ , that is,

$$P[Z_i > C | \mathbf{Z}_m] < \theta \quad i = 1, 2, \dots, n \quad (3)$$

This is illustrated in Figure 3. Since the conditional value of Z_i is Gaussian with mean and variance given by (1) and (2), respectively, the above probability can be easily calculated. The congestion alarm is set off when one or more of the predicted values fail to satisfy the above inequality. When this happens, appropriate network management procedures should be taken to alleviate the situation. In the following section, we will discuss dynamic bandwidth allocation in which the system bandwidth can be renegotiated based on the varying traffic requirement.

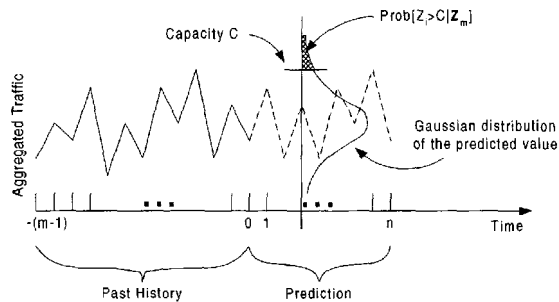


Fig. 3. The congestion alarm.

III. DYNAMIC BANDWIDTH ALLOCATION

Saito [15] studied dynamic virtual path (VP) bandwidth allocation in ATM [13] networks. This corresponds to the first target of the proposed “self-sizing network” concept [5]. Dynamic allocation implies that network engineers do not need to worry about detailed resource allocation at any moment. Instead, the network continuously adjusts its own resource allocation to adapt to the current traffic load. They implemented a dynamic VP bandwidth control system on the telecommunications management network (TMN) platform [18]. Field studies showed that dynamic bandwidth allocation is effective. In this paper, we also consider a similar dynamic bandwidth allocation scheme whereby the system bandwidth can be renegotiated from time to time. Next, we extend the congestion alarm discussed above to determine the amount of extra bandwidth to request or excessive bandwidth to relinquish in case of renegotiation.

Let $\tilde{Z}_i, i = 1, 2, \dots, n$, be the random variable of Z_i given the m historical samples $\{z_{-(m-1)}, z_{-(m-2)}, \dots, z_0\}$. Then, the mean and variance of \tilde{Z}_i , denoted by $m\tilde{z}_i$ and $\sigma_{\tilde{Z}_i}^2$, are given by (1) and (2), respectively. The congestion alarm ensures that

$$P[\tilde{Z}_i > C] < \theta \quad (4)$$

Let C_i be the minimum capacity satisfying the above inequality. Since \tilde{Z}_i is Gaussian, we have:

$$C_i = \sigma_{\tilde{Z}_i} Q^{-1}(\theta) + m\tilde{z}_i \quad (5)$$

where $Q(\cdot)$ is the complementary cdf of a standard Gaussian random variable. Therefore, $\max_{1 \leq i \leq n} C_i$ will be the minimum capacity that will satisfy the quality of service requirement for all the n steps ahead. Suppose the current capacity is C , and let ΔC be defined as follows:

$$\Delta C = \max_{1 \leq i \leq n} C_i - C \quad (6)$$

If ΔC is greater than zero, it means that the current system capacity is probably insufficient to handle the amount of traffic expected to arrive in the following n time slots. In this case, ΔC represents the amount of extra bandwidth to request. On the other hand, if ΔC is less than zero, it means that the current system capacity is more than enough for the expected traffic. In this case, it represents the amount of bandwidth that can be relinquished.

Before investigating the performance of the proposed scheme, the source model used in our simulation, namely, the Pareto ON/OFF model will be discussed.

IV. SIMULATION RESULTS

According to [16], we use the Pareto ON/OFF process as the source model in our simulation. The Pareto distribution is a heavy-tailed distribution which can be defined by the shape parameter and location parameter. The Pareto ON/OFF process consists of strictly alternating ON and OFF periods. During the ON period, packets are generated at a peak rate of p packets per second. The number of packets during each ON period is independent and identically distributed according to a Pareto distribution with a mean of M packets and shape parameter β . After each ON period, a source will transit to an OFF state during which no packets will be generated. The lengths of the OFF state are independent and identically distributed according to a Pareto distribution with a mean of I seconds and shape parameter γ . A single Pareto ON/OFF source does not generate long-range dependent traffic; the aggregation of them does. Willinger et al. [17] have shown that the superposition of a large number of ON/OFF sources with strictly alternating ON and OFF periods and whose ON and/or OFF periods exhibit high variability or infinite variance produce aggregate traffic that exhibits self-similarity. Moreover, they have derived a simple relationship between the parameters describing the high variability and self-similarity. For aggregation of Pareto ON/OFF sources, the Hurst parameter is related to the shape parameters of the ON and OFF periods by $H \doteq \frac{3 - \min(\beta, \gamma)}{2}$.

To justify the use of fGn to model the aggregate traffic, we performed experiments on aggregating 20, 50, and 100 Pareto ON/OFF sources. The generated traces, each of length one million samples, are assessed by using the method of quantile-to-quantile plot to see how well each can be modeled by a Gaussian random variable. Figures 4 to 6 show the quantile-to-quantile plots for aggregation of 20, 50, and 100 Pareto ON/OFF sources, respectively. For the trace to be accurately modeled by the Gaussian distribution, the points should be as close to a 45 degree line as possible. From the figures, we can see that aggregation of 20 sources is not quite Gaussian, 50 sources are better, and for 100 sources, the aggregation can be accurately approximated by a Gaussian distribution. Therefore, we will use 100 sources in the following simulations.

We perform simulations to study the congestion alarm with dynamic bandwidth allocation. The system under consideration is a single link with capacity C Mbps (which is renegotiable) and buffer of size B packets. N Pareto ON/OFF sources with the same parameters ($M = 10$, $\beta = 1.2$, $I = 360$ msec, $\gamma = 1.1$, $p = 256$ packets/sec)¹ are loaded into the system. Packets are assumed to be 10 Kbits long and traffic measurement is performed every 50 msec. The performance threshold θ is set to 10^{-5} .

Three scenarios are compared. Scenario 1 corresponds to the case with fixed bandwidth assignment. The capacity required is determined by the Gaussian assumption with mean and variance measured from the whole trace of the simulation. If the trace mean and variance are given by m and σ^2 respectively, the fixed capacity required to satisfy the given threshold θ is therefore:

¹chosen according to the observations in [16].

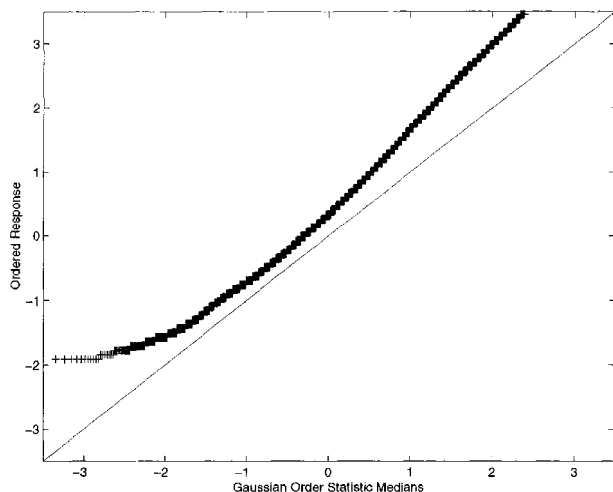


Fig. 4. Quantile-to-quantile plot for aggregation of 20 Pareto ON/OFF sources.

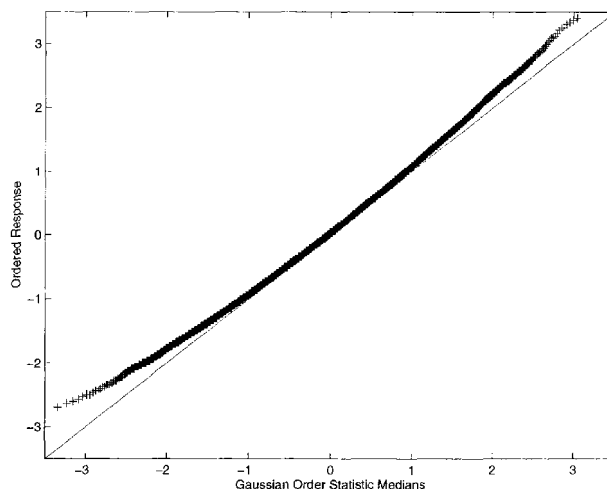


Fig. 6. Quantile-to-quantile plot for aggregation of 100 Pareto ON/OFF sources.

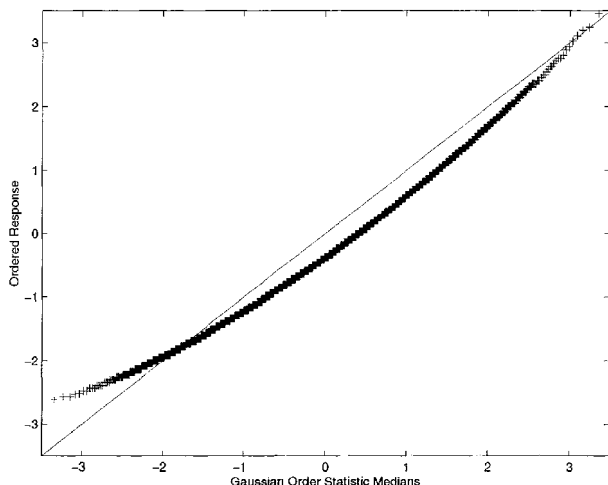


Fig. 5. Quantile-to-quantile plot for aggregation of 50 Pareto ON/OFF sources.

$$C_{\text{fixed}} = \sigma Q^{-1}(\theta) + m \quad (7)$$

In scenario 2, the sample mean and variance of a moving window of W samples are used to dimension the bandwidth based on the Gaussian assumption. In this case, the predictive power utilizing the Hurst parameter is not exploited. This will be called “prediction off” in the following discussion. Finally in scenario 3, the proposed congestion alarm is used. The predictive power for self-similar traffic based on the Hurst parameter is utilized. We call this “prediction on.” The average capacity required in all three scenarios are compared.

In experiment one, we study the effect of the length of the predictive window in performing parameter estimation, that is, the number of measured samples used in estimating the mean and variance of the traffic trace. 100 sources are loaded into the system, and a buffer size of 100 packets is assumed. The simulation time is 15000 seconds, with statistics mea-

sured from the latest 10000 seconds. The first 5000 seconds is a long enough warmup period for the long range dependent effect to be seen. Figure 7 plots the percentage savings in capacity for when prediction is turned on and off, compared to the fixed bandwidth assignment. The mean and standard deviation of the aggregate traffic are found to be 32.6885 Mbps and 6.86683 Mbps, respectively. Based on the Gaussian assumption, a fixed capacity of 61.97 Mbps is required to satisfy the preset threshold of $\theta = 10^{-5}$.

From the figure, a 5% to 8% savings is observed with prediction on compared to prediction off. The percentage savings with both prediction on and off get smaller when the window gets larger. This is because with a larger window, more variation is expected which results in a larger estimated variance. In the extreme case, when the whole trace is used for performance estimation, the estimated variance approaches the variance of the trace and no savings will be observed. However, with a small window, the inaccuracy of the predicted parameters will induce a larger packet loss rate. This can be seen in Figure 8. We can see that for too small an estimation window, the packet loss rate will be unacceptably high. A window of length 40 to 60 samples is therefore appropriate.

The next set of experiments investigates the effect of the length of the history in performing prediction. In these experiments, we use 100 samples for parameter estimation and perform prediction for 20 steps ahead based on a sample history of length from 10 to 100. Figure 9 shows the result. We see that the savings will increase with the number of samples in the history. However, the savings will level off very quickly. This is in line with the rule of thumb, i.e., to predict a second, one needs no more than one second of history. Therefore, in our case, a history of 20 samples is enough. The packet loss rates for all cases are in the order of 10^{-6} .

Next, the effect of the number of prediction steps is studied. With a history of 20 samples, the number of prediction steps is varied from 20 to 100. The resulting percentage savings is plotted in Figure 10. We see that the percentage savings drops rapidly when the number of prediction steps is larger than 20. The packet loss rate in all cases are in the 10^{-4} range. The

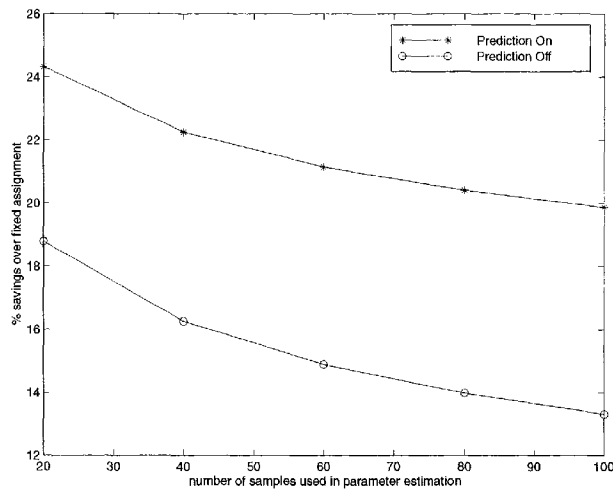


Fig. 7. Effect of window length of parameter estimation on capacity savings.

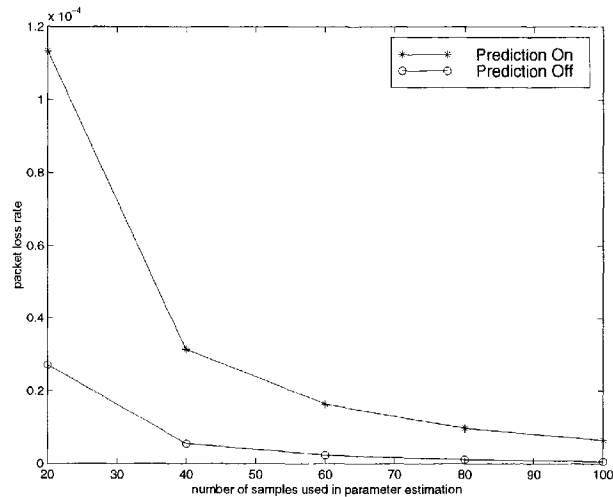


Fig. 8. Effect of window length of parameter estimation on packet loss rate.

larger packet loss rate experienced here is mostly due to the larger “jump window” (of 100 samples) to be discussed next.

To save computation load, after each prediction (or parameter estimation in case of prediction off), the congestion alarm will be idle for a *jump window* of J samples. This jumping reduces the amount of bandwidth renegotiation. However, a large jump window means slower response to traffic loading variation. In the last set of experiments, we study how the jump window affects the performance. With a parameter estimation window of 20 samples, 20 samples are predicted from 20 samples of the history. The jump window size is varied from 20 to 100. The resulting percentage savings is around 24% for prediction on, and 19% for prediction off regardless of the jump window size. However, from Figure 11, which plots the associated packet loss rate, we see that the packet loss rate increases with the jump window size. This is a trade off between the computational complexity and the performance. In practice, it is recommended that the jump

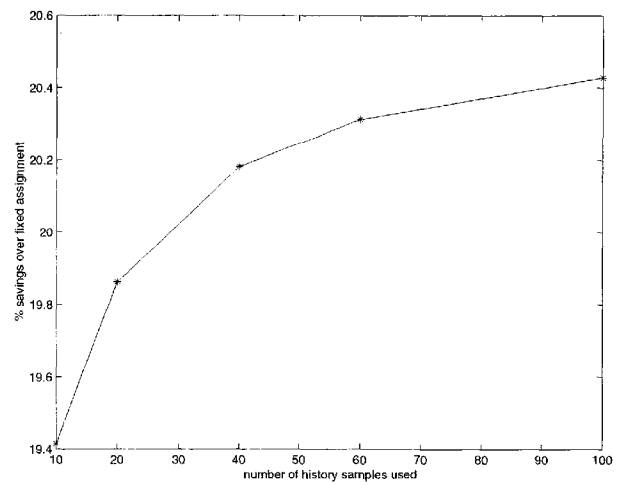


Fig. 9. Effect of number of historical samples used for prediction.

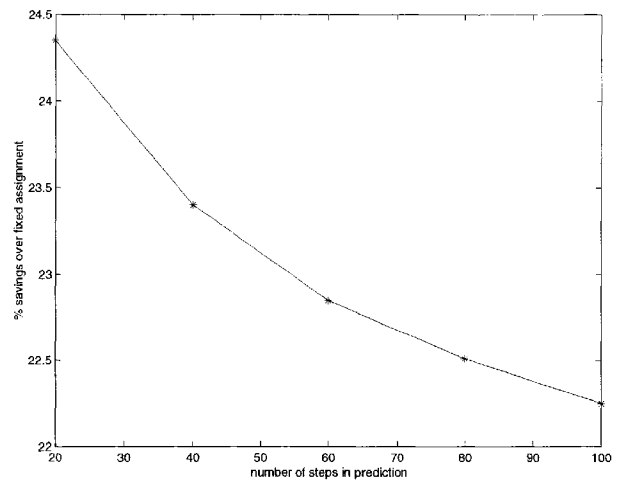


Fig. 10. Effect of number of prediction steps.

window size should not be larger than the number of samples predicted.

In the above experiment, we focus on the capacity savings and packet loss probability without considering the actual outage probability, which is the probability that the aggregated traffic rate is greater than the capacity. The target outage probability is in fact equal to θ . An experiment is performed whereby 200 Pareto On/Off sources are loaded into the system and the outage probability is actually measured. With a target outage probability of 10^{-5} , the simulated outage probability is found to be 0.0059, where there is a two-order of magnitude difference. To see the reason for this discrepancy, another experiment is performed in which a single trace with the same statistics (mean, variance, and Hurst parameter) is generated using an fGn generator. This trace is loaded into the same system and the outage probability is again measured and is found to be 0.0049, which is of the same order as in the previous case. Hence it is clear that the discrepancy is not due to the Gaussian assumption, but probably due to errors intro-

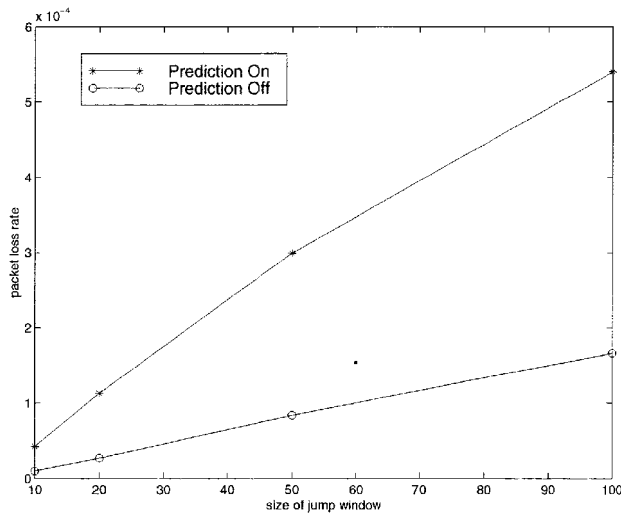


Fig. 11. Effect of the size of the jump window on packet loss rate.

duced in the process of traffic measurements as well as traffic parameter estimations. This kind of difference between target and actual performance is not uncommon with algorithms associated with traffic measurement. For instance, the authors in [8] established a relationship between target and actual performance in their measurement-based admission control framework. We are currently working on quantifying this difference in our congestion alarm.

To sum up, several experiments are performed to study the effects of the various system parameters of the congestion alarm. With prediction on, where the mean, the variance, and the Hurst parameter are used to predict future traffic intensity, the savings is about 20% when compared with fixed bandwidth allocation, and 5% to 8% better than with prediction off, in which case only the mean and variance are used to characterize the traffic loading. In fact, with more stringent outage probability requirement, the savings will be even larger. This is due to the fact that the conditional variance associated with fGn prediction is always smaller than the unconditional variance. The required capacities in both cases are equal to the mean plus the corresponding standard deviation times $Q^{-1}(\theta)$.

V. CONCLUSIONS

In this paper, we studied a measurement-based congestion alarm based on short term prediction for the fractional Gaussian noise model. The utility of the congestion alarm combined with dynamic bandwidth allocation has been investigated through simulations. For a target outage probability requirement of 10^{-5} , there is about 20% savings in capacity over the fixed bandwidth assignment based on the Gaussian assumption of network traffic. Compared to dynamic bandwidth allocation without using prediction based on the Hurst parameter, our approach requires about 5% to 8% less bandwidth on the average. The savings is expected to be even larger for more stringent outage probability requirement. In our simulations, we also verified the rule of thumb which says that to predict one second of the future, we need to use only one second of the recent history. We also find that dy-

namic bandwidth allocation need not be performed frequently to obtain satisfactory results. With the low renegotiation frequency, administrative cost is acceptable. Therefore, the proposed measurement-based congestion alarm with dynamic bandwidth allocation is a promising approach to achieve high system utilization when traffic is self-similar.

The proposed congestion alarm can work with other traffic management functions. Dynamic bandwidth allocation is just one of them. Admission control, for instance, is also a possible candidate. Rather than making an admission control decision on a call-by-call basis, it is possible to perform the decision based on the anticipated congestion level. The system starts out admitting any calls that arrive, and very soon, the system utilization will become high and congestion is likely to occur. When the congestion alarm is set off, the admission control can start to drop some of the new calls. This feedback control mechanism will maintain high system utilization. This approach, as compared to other measurement-based admission control schemes, is simpler since admission control decision is not done on a call-by-call basis, but rather on the anticipated congestion level. The utility of this scheme is worth further investigation.

REFERENCES

- [1] J. Beran, R. Sherman, M. S. Taqqu, and W. Willinger. Long-range dependence in variable-bit-rate video traffic. *IEEE Transactions on Communications*, 43:1566–1579, 1995.
- [2] S. Crosby, I. Leslie, B. McGurk, J. T. Lewis, R. Russel, and F. Toomey. Statistical properties of a near-optimal measurement-based cac algorithm. *Proceedings IEEE ACM '97*, Jun. 1997.
- [3] M. E. Crovella and A. Bestavros. Self-similarity in World Wide Web traffic: evidences and possible causes. *IEEE Trans. on Networking*, 5(6):835–846, Dec. 1997.
- [4] N. G. Duffield and W. Whitt. Network Design and Control Using On-Off and Multi-level Source Traffic Models with Long-Tailed Distributions. *RTCSA '99*, pages 128–132, 1999.
- [5] H. Saito et al. Innovations of circuit/path operations in ATM networks - self-sizing network. *NTT Rev.*, 8(1):56–65, 1996.
- [6] R. Gibbens and F. Kelly. Measurement-based connection admission control. *15th International Teletraffic Congress*, Jun. 1997.
- [7] G. Gripenberg and I. Norros. On the prediction of fractional Brownian motion. *J. Appl. Prob.*, 33:400–410, 1996.
- [8] M. Grossglauser and D. N. C. Tse. A Framework for Robust Measurement-based Admission Control. *IEEE/ACM Trans. on Networking*, 7(3):293–309, Jun 1999.
- [9] G. A. Hirchoren and D. S. Arantes. Predictors for the discrete time fractional Gaussian processes. *IEEE ITS '98*, pages 49–53, 1998.
- [10] S. Jamin, S. Shenker, and P. Danzig. Comparison of measurement-based admission control algorithms for controlled-load service. *IEEE Infocom '97*, pages 973–980, Apr. 1997.
- [11] W. W. Knightly and J. Qiu. Measurement-based admission control with aggregate traffic envelopes. *IEEE ITWDC '98*, Sep. 1998.
- [12] W. E. Leland, M. S. Taqqu, W. Willinger, and D. V. Wilson. On the self-similar nature of Ethernet traffic. *Proc. of the ACM Sigcomm '93, San Francisco, CA*, pages 203–213, 1993.
- [13] S. Ohta and K. Sato. Dynamic bandwidth control of the virtual path in an asynchronous transfer mode network. *IEEE Trans. Commun.*, 40(7):1239–47, 1992.
- [14] V. Paxson and S. Floyd. Wide area traffic: The failure of Poisson modeling. *Proc. of the ACM Sigcomm '94, London, UK*, pages 257–268, 1994.
- [15] H. Saito. Dynamic resource allocation in ATM networks. *IEEE Comm. Magazine*, pages 146–153, May 1997.
- [16] W. Willinger, M. S. Taqqu, W. E. Leland, and D. V. Wilson. Self-similarity in high-speed packet traffic: analysis and modeling of Ethernet traffic measurements. *Statistical Science*, 10:67–85, 1995.
- [17] W. Willinger, M. S. Taqqu, R. Sherman, and D. V. Wilson. Self-Similarity through high-variability: statistical analysis of Ethernet LAN traffic at the source level. *IEEE/ACM Trans. on Networking*, 5(1):128–132, Feb. 1997.
- [18] I. Yoda, K. Yata, and N. Fujii. Object oriented TMN-based operations systems development platform. *Proc. ICC '94*, pages 1824–29, 1994.