AN APPROACH TO ESTIMATING THE TRAFFIC OF A DYNAMIC PACKET-SWITCHED NETWORK

Yuming Liu and Peter J. Moylan
Dept. of Electrical & Computer Engineering
University of Newcastle
NSW 2308, Australia

Abstract— To control the performance of a packet-switched network efficiently, some knowledge on dynamic behavior of the network needs to be investigated. In this paper, we apply an adaptive control technique, *Recursive Least Square (RLS)* estimation, to propose a new approach to estimating the traffic parameters of the network. Based on an introduced dynamic model of the network, this approach can, through adaptively adjusting forgetting factors, not only filter out the noise from estimates while the network operates in a steady state, but also track parameter variations while in a transient state particularly following a step change on the parameter. Since the mean values of the parameters are mainly considered, the dynamic model is suitable to any type of traffic. Therefore, the methodology presented in this paper can then be applied to the issues of B-ISDN network control, such as routing problem and congestion control.

I. Introduction

In a packet-switched network, the good performance means to keep the network operating at a stage where the average delay per packet through the network or the network congestion is minimized subject to individual quality of service (QOS) requirement for a broad diversity of traffic classes [1],[2]. To approach such stage, however, one needs to understand the network quite well; and one needs some sort of dynamic tracking, due to the stochastic property of the network. Parameters such as packet mean arrival rate λ and the link service rate μ, generally determining the performance of the network, play a crucial role in formulating a desirable route table or a flow control strategy for the network. Unfortunately, these two parameters are not measurable, i.e. cannot be obtained from the network directly. This situation forces us to estimate these parameters in real time, based on a mathematical model which expresses the dynamic behavior of the network.

Previous studies of parameter estimation have been mainly restricted to time invariant parameters. Moylan in [3], [4] investigates a dynamic model to smooth, to predict the traffic, and then to complement a routing strategy in circuit-switched network. For a packet-switched network, maximum likelihood estimation of λ is formulated in [5] based on the data for k observed packets.

To estimate time-varying parameters, a standard algorithm is Recursive Least Squares(RLS) estimation, with a forgetting factor to discount the effect of old data. The forgetting factor can simply be a constant over a period of time, or deliberately be adjusted from time to time, depending on variation pattern in parameters. A number of ways to vary the forgetting factor have been investigated. Fortescue [6] suggests a time-varying forgetting factor chosen to keep a measurement of information content at a constant value. Parkum [7] introduces a selective forgetting method in which an individual forgetting factor to each of the eigenvalues of variance of estimate errors is assigned in order to make it possible to adjust the forgetting profile according to the information pattern. These algorithms, however, are only suitable for tracking the slow changes of parameters, and may appear some drawbacks while applied to the packet-switched network, since the real traffic of the network may fluctuate dynamically and step changes are likely to occur frequently. Holst [8] develops a scheme to deal with these abrupt changes and achieves a good performance by switching the system works among different models. This switch is controlled by an argument called detector. But this algorithm is limited only for a system with a normal distribution of measurement noise. Millnert [9] presents an adaptive regulator to handle a dramatically changing system based on the assumption that the system alternates among some fixed grade levels. Obviously, the above assumption doesn't hold for the packet-switched network.

In this paper, based on some priori knowledge on the variance of measured noise, we introduce a new approach to adapting the forgetting factor subject to network parameter variations, even step changes. A careful analysis of the measured noise reveals that the individual sampling value of measured noise is conditional bounded. By defining some confidence limits, one can easily check whether the measured noise fluctuates within a bound or beyond it at a sequence of sampling times so as to know if there has been a step change in the parameters. This is because a certain percent of samples are supposed to vary within the bound unless a true change does happen. This criterion can immediately be applied to the design of a time-varying forgetting factor. Moreover, since we adjust the forgetting factor more precisely, the approach can deal with both abrupt and slow changes.

The paper is outlined as follows. In section II, a dynamic model of a packet-switched network is introduced,

0-7803-2509-5/95 US$4.00 © 1995 IEEE

which is suitable for the application of RLS estimation. Then the measured noise is studied and a time-varying forgetting factor is proposed in section IV and section V, respectively. Finally, section VI gives a simulation study and some conclusions appear in section VII.

II. Dynamic Model

Consider measurable quantities $a(t_k)$, $m(t_k)$ and $q(t_k)$ at sampling time t_k in discrete-time system. Let $a(t_k)$ be accumulated number of packet arrivals up to time t_k, $m(t_k)$ be the number of packets in the system including both those in the queue waiting to served and that is being served, and $q(t_k)$ denote the queue length in the buffer. There is little difficulty to understand that each of terms $a(t_k)\, m(t_k)$ and $q(t_k)$ can then be formulated by a sum of its mean and a noise term. That is,

$$\begin{aligned} a(t_k) &= \bar{a}(t_k) + v_a(t_k) && (2.1)\\ m(t_k) &= \bar{m}(t_k) + v_m(t_k) && (2.2)\\ q(t_k) &= \bar{q}(t_k) + v_q(t_k) && (2.3) \end{aligned}$$

where $\nu_a(t)$, $\nu_m(t)$ and $\nu_q(t)$ are measured noises of zero mean.

A further investigation on the dynamic behavior of the queue indicates that a dynamic model, based on the quantities $a(t_k)$, $m(t_k)$ and $q(t_k)$ can be built as [10].

$$y_k = \varphi_{k-1}^T \theta + \nu_k \qquad (2.4)$$

where

$$\begin{aligned} y_k &= \begin{bmatrix} \delta a(t_k) \\ \delta m(t_k) \end{bmatrix} \\ \varphi_{k-1}^T &= \begin{bmatrix} 1 & 0 \\ 1 & -\tilde{p_0} \end{bmatrix} \\ \theta &= \begin{bmatrix} \lambda \\ \mu \end{bmatrix} \\ v_k &= \begin{bmatrix} \delta v_a(t_k) \\ \delta \nu_m(t_k) \end{bmatrix} && (2.5) \end{aligned}$$

and furthermore,

$$\begin{aligned} \tilde{p_0} &= m(t_{k-1}) - q(t_{k-1}) \\ \delta a(t_k) &= \frac{a(t_k) - a(t_{k-1})}{\Delta} \\ \delta m(t_k) &= \frac{m(t_k) - m(t_{k-1})}{\Delta} \\ \delta v_a(t_k) &= \frac{v_a(t_k) - v_a(t_{k-1})}{\Delta} \\ \delta v_m(t_k) &= \mu[v_m(t_{k-1}) - v_q(t_{k-1})] \\ &\quad + \frac{v_m(t_k) - v_m(t_{k-1})}{\Delta} \end{aligned}$$

where $\Delta = t_{k+1} - t_k$ is the sampling interval, and $\Delta < 2/\mu$ is required to guarantee the model is stable [11]. In the model (2.4), It can be seen that y_k and φ_{k-1}^T are two vectors of quantities depending on the number of packet arrivals during the current sampling interval, and the number of packets in the buffer of the node at present and previous sampling instants. ν_k is a vector of zero-mean entries. Our goal is to obtain a filtered estimate of the parameter vector θ. Given such a model, one can develop a parameter estimation algorithm to deal with the case, of considerable practical importance, where the arrival intensity λ and service rate μ are unknown and may be allowed to be time-varying even abruptly.

III. Recursive Least Squares(RLS) Estimation

Applying the well-known recursive least square estimation technique into our model, an estimator can be described by the following equations [12].

$$\hat{\theta}_k = \hat{\theta}_{k-1} + L_k \varepsilon_k \qquad (3.1)$$

where the gain vector L_k is given by

$$L_k = P_{k-1}\varphi_k (1 + \varphi_k^T P_{k-1} \varphi_k)^{-1}$$

and the prediction error is gained by

$$\varepsilon_k = y_k - \varphi_k^T \hat{\theta}_{k-1}$$

the matrix P_k is updated according to

$$P_k = \alpha^{-1}\left(P_{k-1} - \frac{P_{k-1}\varphi_k \varphi_k^T P_{k-1}}{\alpha I + \varphi_k^T P_{k-1} \varphi_k}\right) \qquad (3.2)$$

where the forgetting factor α can be chosen as $0 < \alpha \leq 1$, and may be adjusted dynamically. For example, when the network performs in a steady state, λ and μ must have remained at some fixed levels. Therefore, the main task of the estimator will be to filter out the measured noise so as to smooth the estimates. This goal can be achieved by raising the α to a higher level to allow the estimator to include more old data; On the other hand, whenever there have been some changes on the parameters particularly like step changes, the α needs to be reduced to a low level to discount the old data as quick as possible so as to eliminate their effects on the new estimates. As a result, the estimator can promptly track those variations. This turns out that the key step to improve the filtering and tracking ability of the estimator is to employ a state-dependent detector. The detector will continuously monitor the variations from the network, and consequently instruct the adjustment of the forgetting factor. More details about the design of the detector are discussed in the following section.

IV. Study of Measured Noise

The study of measured noise aims at seeking an effective detector to monitor the network. Let $a(\Delta)$ be a stochastic variable representing the number of packet arrivals during the time interval Δ. Recalling the quantity $a(t_k)$ in equation (2.1), we have $a(\Delta) = a(t_k) - a(t_{k-1})$ in kth sampling interval. The measured noise $v_a(\Delta)$ is evaluated by

$$v_a(\Delta) = a(\Delta) - \bar{a}(\Delta)$$

there $\bar{a}(\Delta)$ denotes the mean value of $a(\Delta)$ which has

$$\bar{a}(\Delta) = E\{a(\Delta)\} = \lambda \Delta \qquad (4.1)$$

Theorem 1 *Let η be a percentage level, which is known as a confidence limit. Then we have,*

(a). $v_a(\Delta)$ is conditional bounded, that is $|v_a(\Delta)| \leq \mathcal{B}$ with probability greater than η.

(b). the bound $\mathcal{B}$ is governed by following criterion,

$$\mathcal{B} = \sqrt{\frac{V_{a(\Delta)}}{1-\eta}} \tag{4.2}$$

where $V_{a(\Delta)}$ represents the variance of the measured noise $v_a(\Delta)$

Proof:

The proof is immediate from Chebyshev's Inequality theorem, which is stated as

$$P\{|\xi - \bar{\xi}| \geq \epsilon\} \leq \frac{V_\xi}{\epsilon^2} \tag{4.3}$$

where $\bar{\xi}$, V_ξ are respectively the mean and variance of a stochastic variable ξ, and ϵ can be any positive real number.

For a given traffic model and confidence limit, say 95%, the bound $\mathcal{B}$ can be obtained from equation (4.2) at every sampling instant for the current estimated λ. Now let $v_a(\Delta)$ be observed by the detector to see whether it varies within the bound or beyond the bound. If it is noticed that there have been some sampling times at which $v_{a(\Delta)}$ fluctuates beyond the bound, it implies the balance of relation (4.3) has been broken down due to a dramatical change of $\bar{\xi}$. The detector therefore reports that a step change might have occured on parameter λ. Accordingly, the forgetting factor will be dropped down to a lower level to allow the estimator to discard the old data rapidly so that it enables the estimates to follow that change promptly. Since the detector reacts much quickly with the occurrence of the step change, proper values for the time-varying forgetting factor can immediately be assigned. The accuracy of estimates will definitely be improved.

V. Periodic Forgetting Factor

Periodic Forgetting Factor(PFF) was first presented in authors' previous work [10]. We would like to give a brief explanation here. Considering time interval T, which is assumed to be much longer than sampling interval Δ, one can construct a time sequence which is marked as T_0, t_{01}, t_{02}, $\cdots$, t_{0n}, T_1, t_{11}, t_{12}, $\cdots$, t_{1n}, $\cdots$, T_m, $\cdots$, where $t_{ij+1} - t_{ij} = \Delta$, $T_{m+1} - T_m = T$ with $T = n\Delta$, $(n >> 1)$. The PFF is defined by.

$$\alpha_{pff} = \begin{cases} \kappa & t = T_m \, (m = 1, 2, \ldots) \\ 1 & otherwise \end{cases}$$

where κ will be modified at each moment $T_i (i = 1, 2, \cdots)$ subject to recommendations from the detector. During a period of time T, α_{pff} keeps at a high value to allow more old data 'alive' so as to smooth the estimates. This smoothness is attributed to the increment of measurements during long time observation. While at each time $T_i (i = 1, 2, \cdots)$, since κ is changeable between 0 and 1, α_{pff} can be adjusted to whatever it requires, such as a low value.

A. New Bound

In PFF, since the forgetting factor will be updated in every time interval T rather than Δ, the inspection point of the detector will instead be measured noise of the packet arrivals during time T, i.e.

$$v_a(T) = a(T) - \lambda T$$

The corresponding bound $\mathcal{B}_T$ is found

$$\mathcal{B}_T = \sqrt{\frac{V_{a(T)}}{1-\eta}} \tag{5.1}$$

where $V_{a(T)}$ refers to the variance of $v_a(T)$. As mentioned in last section, we prefer to know if $|v_a(T)|$ is less than $\mathcal{B}_T$ and how big a gap exists between them. In practice, we find that the measured noise still needs to be trimmed before being examined by the detector. Therefore, a low pass filter is applied, which is of following recursive form.

$$\sigma_{k+1} = \beta \sigma_k + (1 - \beta) v_a(T) \tag{5.2}$$

where σ_k denotes the filtered measured noise and β is a coefficient. It is clear that the mean value of term σ_k still remains the same as the mean value of $v_a(T)$. But V_n, the variance of σ_k, will be scaled down

$$V_n = \frac{1-\beta}{1+\beta} V_{a(T)} \tag{5.3}$$

Accordingly, the new bound $\mathcal{B}_n$ for the filtered measured noise can be roughly estimated as.

$$\mathcal{B}_n = \sqrt{\frac{1-\beta}{1+\beta}} \mathcal{B}_T \tag{5.4}$$

The parameter β determines, roughly speaking, how many $\sigma_{a(T)}$ values that should be included. E.g. $\beta = 0.95$ corresponds to about 20 values, which is a reasonable choice in many applications. A small β allows a fast detection of step changes, although at the price of less security against false alarms. This trade-off is typical for all of such kind of detections.

B. Dynamic Property of κ

The quantity κ plays a key role in controlling the forgetting factor. Intuitively, we believe that an ideal κ subject to measured noise and its bound should be particularly designed as shown in Fig (1). The graph can be divided into two areas: steady area and transient area, which correspond to two cases: measured noise varies within its bound or breaks through its bound. Consider the case that the measured noise fluctuates beyond the bound. This implies there might have been a step change on the parameters. The forgetting factor κ, therefore, should be reduced to a lower level so as to enable the estimates to track the step change as quick as possible. Generally speaking, the more significant the measured error exceeds its bound, the more certain the estimator feels that a step change has occurred.

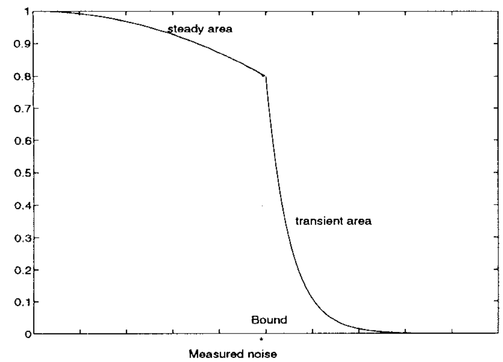

Figure 1: Ideal Forgetting Factor Versus Measured Noise

Consequently, the more urgent to forget old data the estimator becomes, and the more dramatically the forgetting factor is required to be reduced. On the other hand, however, if the noise varies within the bound, it means a steady state has been built, therefore, κ needs to be increased to a high level so as to keep more data alive and to filter out the noise from estimates. At this moment, for a security reason, any change on forgetting factor must be very careful, since the steady state may be interrupted due to the choice of a wrong value for forgetting factor. In the standard Recursive Least Square algorithm, forgetting factor is usually recommended varying around 0.8 to 0.99. The reason one never chooses too low value is that, it is believed, if the forgetting factor has moved close to 0.8, discounting speed may become quick enough for estimates to attempt to follow any variations except step change. This rule is also applicable to the design of κ in its steady area. summarizing above two aspects, κ can be deliberately designed as

$$\kappa = \begin{cases} \gamma_l e^{\mathcal{B}_n - \sigma_s} & |\sigma_s| \geq \mathcal{B}_n \\ 1-(1-\gamma_u)(\frac{\sigma_s}{\mathcal{B}_n})^2 & \text{otherwise} \end{cases} \tag{5.5}$$

where parameters γ_u gives the minimum value of the forgetting factor in the steady area, and γ_u determines the maximum value in the transient area. Both need a tradeoff for choice. The lower the γ_u or γ_l are picked, the fast the estimate follows the variations of parameters, with the cost of less security against false measurements. The parameter σ_s is referred to a successive increment or decrease of $v_a(T)$ immediately prior to the moment κ is updated. The reason of using σ_s instead of $v_a(T)$ directly is that, as σ_s records a peak-peak value of $v_a(T)$, it shows the tendency of variation of $v_a(T)$, which reveals the step changes more apparently than $v_a(T)$ does.

It is note that the equation (5.5) is not the unique one. More solutions could be possible in practice.

VI. Simulation Study

To confirm the superbly performance of *PFF* in the estimation of parameters in the network, simulation study is implemented.

The following example is typical. λ is allowed to be time-varying during the 24000 seconds of simulation time, which jumps up from 0.1 to 0.3 at the 8000th second and back from 0.3 to 0.1 at the 16000th second. Let $1/\mu$ keep as constant 3. We pick up 0.4 for γ_u and 0.2 for γ_l as its extreme values of κ, and let T be equal to 20 seconds and Δ equal to 1 second. In order to distinguish its good performance of *PFF* clearly from that of the standard *RLS*, we first depict the estimation results of using *RLS* in Fig (2) and (3), where the forgetting factor is set as a constant 0.98. By *PFF*, the estimated result and reference value of λ are

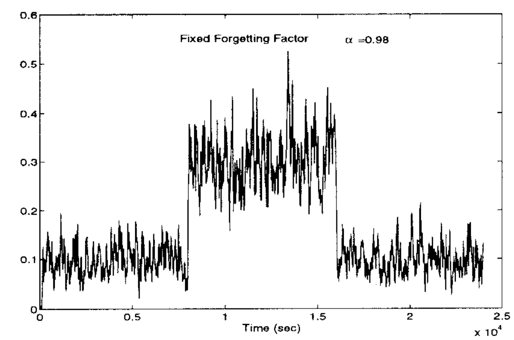

Figure 2: Estimate of λ Using RLS

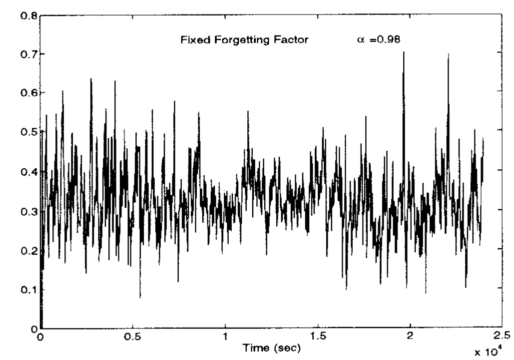

Figure 3: Estimate of μ Using RLS

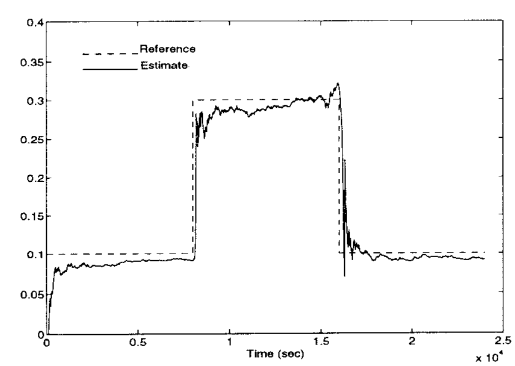

Figure 4: Estimate of λ

sketched In Fig (4), and those of μ in Fig (5). From Fig

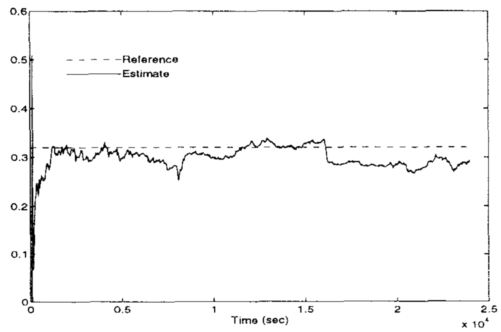

Figure 5: Estimate of μ

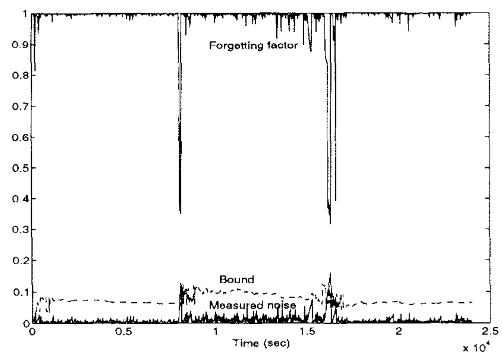

Figure 6: Measured noise, Bound and Forgetting Factor

(2) and (3), it is obvious that the estimates are seriously interfered of measured noise. The only way to smooth the results is to choose a high value for the forgetting factor. Unfortunately, the price is leading a sluggish response near the time where λ changes. Fig (4) shows that new scheme using time-varying PFF performs quite well for the system either in a steady or transient state. To show how the forgetting factor varies in different situations such as λ remains in a fixed level or jumps from one level to another, Fig (6) gives a graphical view along with the related measured noise and its bound. It can be seen that during those three assumed steady states, measured noise never grows up beyond its bound, so the forgetting factor is kept at higher levels almost close to 1; While in those two transient states, measured noise suddenly breaks through its bound, the forgetting factor in turn drops down markedly. These results are just what we expected.

VII. Conclusions

In this paper, a new time-varying forgetting factor is formulated. This factor is updated subject to variation pattern of the traffic in every regular time. Since the adaptive forgetting factor enables the estimates either to remain in a stable level or to track the variations quickly, it is particularly useful in the estimation of time-varying parameters of a packet-switched network. Both theoretical and simulation studies illustrate that the time-varying PFF can handle a system either in a steady state or a transient moment following parameter variations even step changes.

To estimate the parameters in a packet-switched network is an essential step but not our final destination. Based on such a result, a more efficient way to control the network can be achieved if a dynamic routing strategy or flow control technique is proposed.

References

[1] J. Kurose. Open issues and challenges in providing quality of service guarantees in high-speed networks. *Computer Communication Review*, 23(1):6–15, January 1993.

[2] D. Yates, J. Kurose, D. Towsley, and M. G. Hluchyj. On per-session end-end delay distribution and the call admission problem for real-time applications with QoS requirements. *Computer Communication Review*, 23(4), Oct. 1993.

[3] P. J. Moylan and B. H. Soong. The expected mean demand in a telephone network. *A.T.R.*, 23(1), 1989.

[4] P. J. Moylan. A network model for use in adaptive call routing studies. First Australian Teletraffic Research Seminar, Melbourne, 1986.

[5] T. M. Chen, J. Walrand, and D. G. Messerschmitt. Parameter estimation for partially observed queues. *IEEE Transactions on Communications*, 42(9), September 1994.

[6] T. R. Fortescue, L. S. Kershenbaum, and B. E. Ydstie. Implantation of self-tuning regulators with variable forgetting factors. *Automatica*, 17(6), 1981.

[7] J. E. Parkum, N. K. Poulsen, and J. Holst. Recursive forgetting algorithms. *Int. J.Control*, 55(1), 1992.

[8] J. Holst and K. Poulsen. Self tuning control of plants with abrupt changes. *Preprints 9th IFAC World Congress,Budapest*, 7:144–149, 1984.

[9] M. Millnert. Adaptive control of abruptly changing systems. In *Proceeding of 9th IFAC world Congress,Budapest*, pages 208–212, 1984.

[10] Y. Liu and P. J. Moylan. On-line estimation of parameters in packet-switching networks. In *Proceedings of Australian Telecommunication Networks and Applications Conference*, pages 255–260, Melbourne, December 1994.

[11] R. H. Middleton and G. C. Goodwin. *Digital Control And Estimation: A Unified Approach.* Prentice-Hall, Englewood Cliffs, NJ, 1990.

[12] G. C. Goodwin and R. C. Payne. *Dynamic System Identification: Experiment Design and Data Analysis.* Academic Press, New York, 1977.