

Tail Probability of a Gaussian Fluid Queue under Finite Measurement of Input Processes

Kazutomo Kobayashi

C&C Media Research Laboratories, NEC Corporation

4-1-1 Miyazaki, Miyamae-ku, Kawasaki, Kanagawa 216, Japan

tel:+81 44 856 2316, fax:+81 44 856 2235, email: koba@ccm.cl.nec.co.jp

Yukio Takahashi

Department of Mathematical and Computing Sciences,

Tokyo Institute of Technology

2-12-1 Ookayama, Meguro-ku, Tokyo 112, Japan

tel:+81 3 5734 3203, fax:+81 3 5734 2752, email: yukio@is.titech.ac.jp

Abstract

Recently, there has been frequent discussion of whether communication traffic is long-range dependent or not. This paper gives another insight to this issue by discussing the problem of estimating the tail probability $P(Q > x)$ of a Gaussian fluid queue under finite measurement of input processes.

We show that if the mean m and the autocovariance function $\{\Lambda(t)\}_{0 \leq t \leq T}$ of an input rate process can be estimated from traffic data of a finite length, $P(Q > x)$ for x in a finite interval can be evaluated by an approximation formula determined only from m , $\{\Lambda(t)\}_{0 \leq t \leq T}$ and the output rate c . This result implies that as long as we evaluate $P(Q > x)$ in a finite region of x , it is not important whether the input rate process is long-range dependent or not.

We also apply the approximation formula to the performance evaluation of an ATM multiplexer with VBR Video traffic. We see that $\Lambda(t)$ can be estimated in a sufficient range to evaluate $P(Q > x)$ for x in the practical range and that the formula provides a good approximation except for a scale parameter.

Keywords

Gaussian processes, fluid queues, long-range dependence, finite measurement, ATM multiplexers

1 INTRODUCTION

Recently it has been reported that local area network traffic and VBR (Variable Bit Rate) video traffic exhibit long-range dependence [Beras 95, Lela 94]. Many traffic models proposed so far, for example Markovian-type traffic models [Ster 91, Heym 92, Addi 94] and Gaussian-type input rate processes [Kobas 95, Debi 95], are not long-range dependent. They are essentially Markovian and short-range dependent. *Self-similar* processes such as *fractional Gaussian noises* [Mand 68] are typical long-range dependent processes, and they have therefore been proposed as models of such traffics [Beras 95].

Generally speaking, however, it cannot be determined from traffic data of a finite length whether the traffic is long-range dependent or not in a rigorous sense. We can only guess whether it is long-range dependent or not by using *R/S* analysis or variance-time analysis.

In this paper, we discuss this problem from another point of view by using a fluid queue with a stationary Gaussian input rate process R_t and a constant output rate c . We call this queue a *Gaussian fluid queue*.

According to [Bene 63, Duffs 95], the tail probability of the queue-length distribution $P(Q > x)$ is represented as

$$P(Q > x) = P(\sup_{t>0} W_t > x), \quad (1)$$

where

$$W_t = \int_{-t}^0 R_s ds - ct.$$

The stochastic process W_t is completely characterized by $\mu = c - m$ and $\{\Lambda(t)\}_{0 \leq t < \infty}$, where m is the mean of R_t and $\Lambda(t)$ is its autocovariance function. Therefore $P(\sup_{t>0} W_t > x)$ can be thought as a function of x , μ and $\{\Lambda(t)\}_{0 \leq t < \infty}$:

$$P(\sup_{t>0} W_t > x) = F(x, \mu, \{\Lambda(t)\}_{0 \leq t < \infty}).$$

If $\sup_{t>0} W_t > x$ for a large x , then under a mild condition it is likely that W_t reaches the level x in a neighborhood of t^* , say, at which $P(W_t > x)$ attains its supremum [Duffs 95, Kou 97]. We denote by $\kappa(x)$ the point t^* as a function of x :

$$\kappa(x) = \arg \sup_{t>0} P(W_t > x). \quad (2)$$

Since the probability that W_t reaches the level x out of the neighborhood of $t^* = \kappa(x)$ is negligible, $P(\sup_{t>0} W_t > x)$ for x in a finite interval (a, b) can be estimated from the behavior of $\{W_t\}_{0 \leq t \leq T}$ with T such that $\sup_{a < x < b} \kappa(x) < T < \infty$. Note that the process $\{W_t\}_{0 \leq t \leq T}$ is characterized by μ and $\{\Lambda(t)\}_{0 \leq t \leq T}$. Thus if $\kappa(x)$ is also determined from μ and $\{\Lambda(t)\}_{0 \leq t \leq T}$, then we see that the tail probability $P(Q > x)$ is approximated by a function of x , μ and $\{\Lambda(t)\}_{0 \leq t \leq T}$, namely,

$$P(Q > x) \approx \tilde{F}(x, \mu, \{\Lambda(t)\}_{0 \leq t \leq T}), \quad a < x < b. \quad (3)$$

We note that Ryu and Elwalid [Ryu 96] and Grossglauser and Bolot [Gros 96] made

similar discussions, independently, but their mathematical expositions were not so clear as ones given here.

In the next section we show that the inverse function $\phi(t)$ of $\kappa(x)$ exists and is given by

$$\phi(t) = \mu \left(\frac{2}{\beta_t} \int_0^t \beta_s ds - t \right), \quad \beta_t = \int_0^t \Lambda(s) ds \quad (4)$$

under the condition that $\sigma_t = \sqrt{\text{Var}[W_t]}$ is increasing and concave and that $\phi(t)$ is a monotone increasing function with $\phi(+0) = 0$. Since the function is monotone, for a, b such that $0 < a < b < \phi(T)$ we have $\sup_{a < x < b} \kappa(x) \leq \kappa(b) < T$. Thus the value of $\kappa(x)$ for $x \in (a, b)$ is uniquely determined from μ and $\{\Lambda(t)\}_{0 \leq t < T}$. We also show that the approximating function $\tilde{F}(x, \mu, \{\Lambda(t)\}_{0 \leq t < T})$ is given by

$$\tilde{F}(x, \mu, \{\Lambda(t)\}_{0 \leq t \leq T}) = B \exp \left\{ -\frac{\mu^2}{\beta_{\kappa(x)}^2} \int_0^{\kappa(x)} \beta_s ds \right\}. \quad (5)$$

Therefore, if the mean m and the autocovariance function $\{\Lambda(t)\}_{0 \leq t \leq T}$ of R_t can be estimated from traffic data of a finite length, the tail probability $P(Q > x)$ for $x \in (a, b)$ with $0 < a < b < \phi(T)$ can be evaluated using (4) and (5). Conversely, if we need to evaluate $P(Q > x)$ for $x \in (a, b)$, we can easily obtain its approximation, except for B , by estimating m and $\Lambda(t)$ from 0 to T such that $b < \phi(T)$. We note that the approximation formula (5) is consistent with previous results obtained by Norros [Norr 94], Duffield and O'Connell [Duffs 95] and Kouchi *et al.* [Kou 97]. However, our result differs from them on the point that we discuss $P(Q > x)$ for x in a finite interval under *finite measurement*, while they discuss the tail behavior of $P(Q > x)$ as x increases to infinity.

In the discussion of whether traffic is long-range dependent or not, the behavior of $\Lambda(t)$ as $t \rightarrow \infty$, namely, $\{\Lambda(t)\}_{T < t < \infty}$, is the most important factor [Bera 95]. However, the formula (5) indicates that $\{\Lambda(t)\}_{0 \leq t \leq T}$ is a big factor of the tail probability $P(Q > x)$ for x in a finite region. We check this result by simulation for a Gaussian fluid queue with a long-range dependent input and for a queue with a short-range dependent one. These inputs have almost the same $\{\Lambda(t)\}$ in the range $(0, 10)$ and almost the same β_t in the range $(0, 15)$. The simulation result indicates that in both cases $P(Q > x)$ is almost the same in the region of $x \in (0, \phi(15))$ and is well approximated by the formula (5) except for the constant B . This result indicates that as long as we discuss $P(Q > x)$ in a finite region of x , it is not important whether an input rate process is long-range dependent or not. It is more important for practical applications whether $\Lambda(t)$ in a sufficient range can be estimated or not from measured traffic data and whether the formula (5) provides a good approximation or not.

Finally, we apply the approximation formula (5) to the performance evaluation of an ATM multiplexer with VBR video traffic. In practical applications the formula (5) may contain various kinds of errors, such as an estimation error of $\Lambda(t)$, the difference between an actual input traffic and a Gaussian process, an approximation error indwelled in the formula (5) and so on. Nevertheless, the simulation results show that the formula (5) provides a good approximation except for the multiplicative constant B and that $\Lambda(t)$ can be estimated in a sufficient range for the performance evaluation of ATM multiplexers.

The rest of the paper is organized as follows. In Section 2 we discuss our queueing

model and state the main result. We explain the Gaussian fluid queue in Subsection 2.1 and derive the inverse function (4) and the approximation formula (5) in Subsection 2.2. In Section 3 we present numerical results. In Subsection 3.1 we apply the formula (5) to fluid queues with a fractional Gaussian noise input and with an autoregressive process one. In Section 3.2 we apply the formula (5) to the performance evaluation of an ATM multiplexer with VBR Video traffic.

2 MODEL AND MAIN RESULT

In this section we introduce our queueing model and discuss the main result. All the symbols used in Introduction are re-defined here.

2.1 Model description

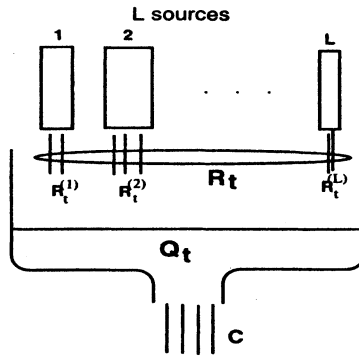


Figure 1 A fluid queue

We consider a fluid queue model (see Figure 1) in which the buffer content Q_t varies according to the differential equation

$$\frac{dQ_t}{dt} = \begin{cases} R_t - c, & \text{if } Q_t > 0 \text{ or } R_t > c, \\ 0, & \text{otherwise,} \end{cases}$$

where c is a constant output rate and R_t is a stationary Gaussian input rate process. We denote by m the mean, by σ_R^2 the variance and by $\Lambda(t)$ the autocovariance. We refer to this fluid queue model as a *Gaussian fluid queue*.

Let A_t be the input process of R_s in the interval $[-t, 0)$, namely,

$$A_t = \int_{-t}^0 R_s ds. \quad (6)$$

We define by β_t the covariance between R_0 and A_t , and by σ_t^2 the variance of A_t . Using

the autocovariance function $\Lambda(t)$ of R_t , they are written as

$$\beta_t = E[(R_0 - m)(A_t - mt)] = E[(R_0 - m) \int_{-t}^0 (R_s - m) ds] = \int_0^t \Lambda(s) ds, \quad (7)$$

$$\begin{aligned} \sigma_t^2 &= E[\int_{-t}^0 \int_{-t}^0 (R_u - m)(R_v - m) dudv] = 2 \int_0^t \int_0^s \Lambda(u) dudv \\ &= 2 \int_0^t \beta_s ds. \end{aligned} \quad (8)$$

From (7) and (8) the standard deviation σ_t is two times differentiable.

In this paper, we assume that $m < c$ and that σ_t is strictly increasing and concave, i.e., $\sigma_t' > 0$ and $\sigma_t'' < 0$.

2.2 The tail behavior of $P(Q > x)$ under finite measurement

Suppose that the autocovariance function $\Lambda(t)$ of the input rate process in the region $[0, T]$ and the mean m are estimated from traffic data of a finite length. Then we show that the tail probability $P(Q > x)$ for x in a finite interval can be evaluated by an approximation formula determined from μ and $\{\Lambda(t)\}_{0 \leq t \leq T}$. This result suggests that as long as we evaluate $P(Q > x)$ in a finite region of x , it is not important whether the input rate process is long-range dependent or not.

The tail behavior of $P(Q > x)$

Duffield analyzed the tail probability of a single queue with an input whose class covers long-range dependent processes [Duff 95, Duffs 95]. Applying his result to a Gaussian fluid queue, under a mild condition the tail probability $P(Q > x)$ is approximated by

$$P(Q > x) \approx B \exp \left\{ - \inf_{t>0} \frac{(x + \mu t)^2}{2\sigma_t^2} \right\}, \quad (9)$$

where $\mu = c - m$ and B is a constant. See Appendix for more details.

We can interpret the formula (9) as follows. Let $W_t = A_t - ct$. Then $P(Q > x)$ is given as $P(Q > x) = P(\sup_{t>0} W_t > x)$ [Bene 63, Duffs 95]. We define $\kappa(x)$ as the point t at which $P(W_t > x)$ attains the supremum; $\kappa(x) = \arg \sup_{t>0} P(W_t > x)$. Since W_t for a fixed t is a normal variable with mean $-\mu t$ and variance σ_t^2 , $\kappa(x)$ is given by

$$\kappa(x) = \arg \sup_{t>0} P \left(\frac{W_t + \mu t}{\sigma_t} > \frac{x + \mu t}{\sigma_t} \right) = \arg \inf_{t>0} \frac{x + \mu t}{\sigma_t}. \quad (10)$$

If $\sup_{t>0} W_t > x$ for a large x , then it is very likely that the process W_t reaches x in a neighborhood of $t^* = \kappa(x)$ [Duff 95, Duffs 95]. That is, the probability that W_t reaches x out of the neighborhood of $t^* = \kappa(x)$ is negligible. Therefore $P(\sup_{t>0} W_t > x)$ behaves as

$$\sup_{t>0} P(W_t > x) = \sup_{t>0} \frac{1}{\sqrt{2\pi}} \int_{\frac{x+\mu t}{\sigma_t}}^{\infty} e^{-\frac{y^2}{2}} dy. \quad (11)$$

Since for $y > 0$

$$\frac{1}{y+1} e^{-\frac{y^2}{2}} \leq \int_y^\infty e^{-\frac{u^2}{2}} du \leq \frac{1}{y} e^{-\frac{y^2}{2}},$$

the exponent of (11), and hence the exponent of $P(Q > x)$, is represented as

$$-\inf_{t>0} \frac{(x + \mu t)^2}{2\sigma_t^2}. \quad (12)$$

The approximation of $P(Q > x)$ under finite measurement

We show that *the formula (9) can be written as*

$$P(Q > x) \approx B \exp \left\{ -\frac{\mu^2}{\beta_{\kappa(x)}^2} \int_0^{\kappa(x)} \beta_s ds \right\}. \quad (13)$$

The function β_s for $s \in [0, T]$ can be calculated from $\{\Lambda(t)\}_{0 \leq t \leq T}$ using (7), but $\kappa(x)$ can not be calculated directly by using the definition (10). Fortunately, as will be proved later, *the inverse function $\phi(t)$ of $\kappa(x)$ exists and is given as*

$$\phi(t) = \mu \left(\frac{2}{\beta_t} \int_0^t \beta_s ds - t \right), \quad (14)$$

and it is a monotone increasing function with $\phi(+0) = 0$.

Since $\phi(t)$ is a monotone increasing function with $\phi(+0) = 0$, the value of $\kappa(x)$ for x in any interval (a, b) where $0 < a < b < \phi(T)$ can be easily obtained from μ and $\{\Lambda(t)\}_{0 \leq t \leq T}$ through (14). Thus, the exponent of (13) for $x \in (a, b)$ can be calculated from x , μ and $\{\Lambda(t)\}_{0 \leq t \leq T}$.

Remark 1. In this paper we do not discuss the multiplicative constant B in (13). However, from the discussion on (3) in the introduction, it is clear that B is determined from μ and $\{\Lambda(t)\}_{0 \leq t \leq T}$.

Remark 2. Note that the parameters governing (14) and the exponent of (13) are μ and $\{\beta_t\}_{0 \leq t \leq T}$. Thus, for two input rate processes with the same m and the same $\{\beta_t\}_{0 \leq t \leq T}$, we expect that the exponents of the tail probabilities are almost the same for $x \in (a, b)$ where $0 < a < b < \phi(T)$.

Let

$$f(x, t) = \frac{(x + \mu t)^2}{2\sigma_t^2} \quad (15)$$

and t^* be a value of t satisfying

$$\frac{\partial f(x, t)}{\partial t} = 0. \tag{16}$$

From (8) we have $\sigma_t \sigma'_t = \beta_t$. It follows that

$$\mu \sigma_{t^*}^2 = (x + \mu t^*) \beta_{t^*} \quad \text{or equivalently} \quad x = \mu \left(\frac{\sigma_{t^*}^2}{\beta_{t^*}} - t^* \right). \tag{17}$$

Under the assumption that σ_t is concave, $f(x, t) > f(x, t^*)$ holds for $t \neq t^*$ since

$$\frac{x + \mu t^*}{\sigma_{t^*}} - \frac{x + \mu t}{\sigma_t} = \frac{\mu \sigma_{t^*}}{\beta_{t^*} \sigma_t} \{ \sigma_t - (\sigma_{t^*} + \sigma'_{t^*} (t - t^*)) \} < 0,$$

where $\sigma_{t^*} + \sigma'_{t^*} (t - t^*)$ represents a tangent line of σ_t at t^* . Thus the value t^* , the solution of (16), is unique for a given x so that $\kappa(x) = t^*$ and

$$f(x, t^*) = \inf_{t>0} f(x, t). \tag{18}$$

Conversely for a given t^* the value x such that $\kappa(x) = t^*$ is uniquely determined from (17). Therefore the inverse function $\phi(t)$ of $\kappa(x)$ exists and is given as

$$\phi(t) = \mu \left(\frac{\sigma_t^2}{\beta_t} - t \right) = \mu \left(\frac{2}{\beta_t} \int_0^t \beta_s ds - t \right). \tag{19}$$

From (7), (8) and (19) we have $\lim_{t \rightarrow 0} \phi(t) = 0$ and

$$\phi'(t) = - \frac{\sigma_t}{(\sigma'_t)^2} \sigma_t'' > 0.$$

Thus the function $\phi(t)$ is a monotone increasing function with $\phi(+0) = 0$.

Using (17) and (18), the formula (9) is rewritten as

$$P(Q > x) \approx B \exp \left\{ - \frac{(x + \mu t^*)^2}{2\sigma_{t^*}^2} \right\} = B \exp \left\{ - \frac{\mu^2 \sigma_{t^*}^2}{2\beta_{t^*}^2} \right\}. \tag{20}$$

Then the integral representation (8) leads the formula (13).

3 DISCUSSION WITH NUMERICAL EXAMPLES

3.1 Gaussian fluid queues with long-range dependent input and short-range dependent one

Whether the traffic is long-range dependent or not is determined from the behavior of an autocovariance function $\Lambda(t)$ as $t \rightarrow \infty$, namely, $\{\Lambda(t)\}_{T < t < \infty}$. However, the formula (13)

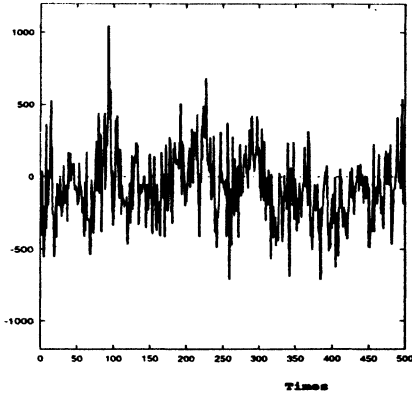


Figure 2 Sample sequence of FGN

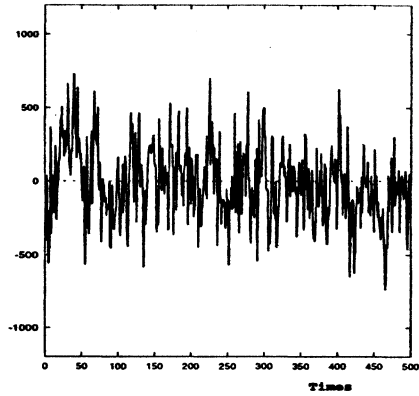


Figure 3 Sample sequence of AR(10)

implies that $\{\Lambda(t)\}_{0 \leq t \leq T}$ is a big factor of the tail probability $P(Q > x)$ for x in a finite region.

Here we generate two sample sequences, one from a fractional Gaussian noise (FGN) model and one from an autoregressive (AR) model, both have almost the same autocovariance function $\Lambda(t)$ within a finite interval $(0, T)$. The FGN model is expected to generate a long-range dependent sequence and the AR model a short-range dependent sequence. If we see the behavior of the tail probability $P(Q > x)$ of Gaussian fluid queues having these sequences as input sequence, we can check whether the approximation formula (13) is valid or not.

We first generate a GFN sequence $\{F_k = Z_k - Z_{k-1}; k = 1, 2, \dots, 100000\}$ via a fractional Brownian motion sequence $\{Z_t\}$ with Hurst parameter $H = 0.8$. The first 500 values of F_k are shown in Figure 2.

Next, by supposing that the sequence $\{F_k\}$ generated above is derived from a 10th-order autoregressive model, we estimate the parameters of the AR model using the method of least squares. Then we generate a sample sequence $\{G_k; k = 1, 2, \dots, 100000\}$ using the AR model. The first 500 values of G_k are shown in Figure 3.

From the graphs of these sequences, it is difficult to state the difference between the two sequences. However, if we compare their autocorrelation functions $\gamma(t) = \Lambda(t)/\sigma_R^2$, they are clearly different, as seen in Figure 4. The autocorrelation functions are almost the same in the range $(0, 10)$, but as the lag increases over 10 the autocorrelation of the AR(10) decays rapidly while that of the FGN decays slowly. The figure clearly shows that the AR(10) is short-range dependent and the FGN is long-range dependent.

Figures 5, 6 and 7 show the covariance β_t , the standard deviation σ_t and the function $\phi(t)$ in (14), respectively. Here we set $\mu = 1.5 * \sigma_R$ where σ_R is the variance of the FGN. From Figures 5 and 6, we find that both β_t 's are almost the same for $t < 15$ and both σ_t 's are concave. From Remark 2, the tail probabilities for the FGN and the AR(10) are expected to be almost the same in the range $(0, \phi(15))$, where $\phi(15) \approx 1500$ from Figure 7.

To demonstrate this, we simulate the behaviors of the tail probabilities in the range $(0, 1800)$ for the FGN and the AR(10). Figures 8 and 9 show the simulation results and

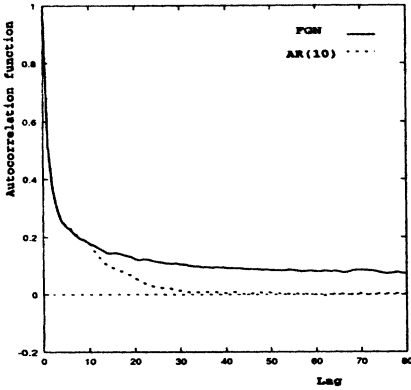


Figure 4 Autocorrelation function of FGN and AR(10)

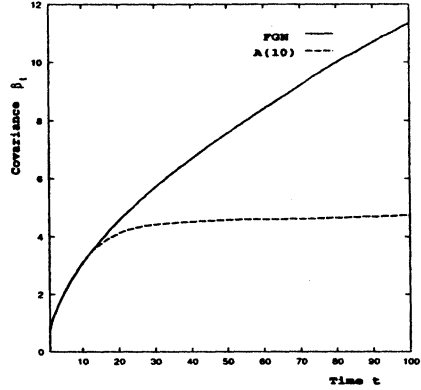


Figure 5 Covariance β_t

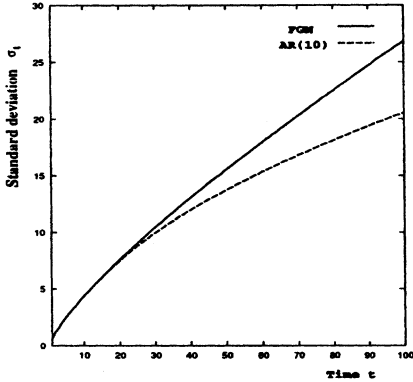


Figure 6 Standard deviation σ_t

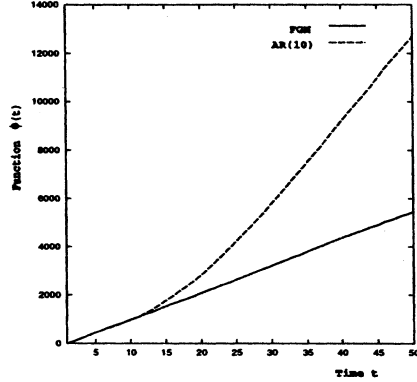
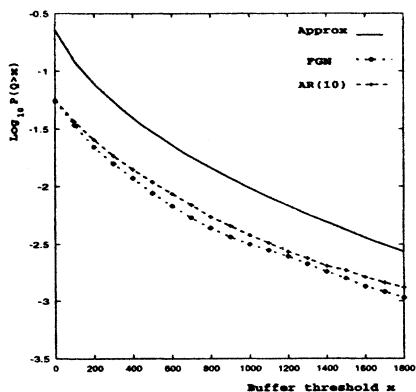
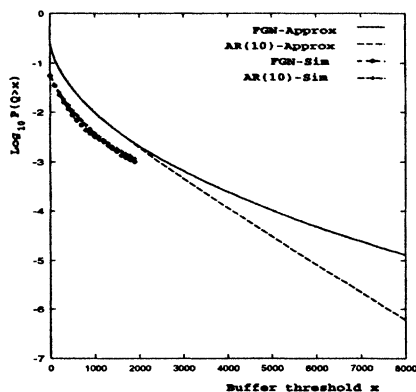


Figure 7 Function $\phi(t)$

the numerical results of the approximation (13) of $P(Q > x)$ with $B = 1$ for $x \in (0, 1800)$ and for $x \in (0, 8000)$. These figures indicate that the tail probabilities for the FGN and the AR(10) are almost the same in the range $(0, 1800)$ as expected and that they are well approximated by the formula (13) except for the multiplicative constant B . For $x > 2000$, the values of the approximation (13) for the two models FGN and AR(10) become different, as seen in Figure 9.

This result implies that as long as we discuss $P(Q > x)$ in a finite region of x , it is not important whether an input rate process is long-range dependent or not. It is more important for practical applications whether $\Lambda(t)$ in a sufficient range can be estimated from measured traffic data.

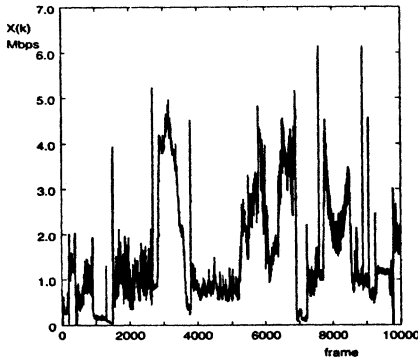
Figure 8 $P(Q > x)$ for $x \in (0, 1800)$ Figure 9 $P(Q > x)$ for $x \in (0, 8000)$

3.2 Performance evaluation of an ATM multiplexer with VBR video traffic

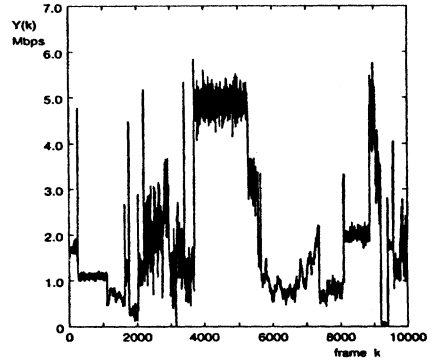
In this section, for the purpose of evaluating performance of an ATM multiplexer with VBR video traffic input, we check whether $\Lambda(t)$ in a sufficient range can be estimated or not from measured traffic data and whether the formula (13) provides a good approximation or not except for the constant B .

When the formula (13) is used in practical applications, there may be various kinds of errors, such as an estimation error of $\Lambda(t)$, the difference between an actual input traffic and a Gaussian process, an approximation error indwelled in the formula (13) and so on. Nevertheless, the simulation results show that the formula (13) provides a good approximation except for B . It is also shown that if $\Lambda(t)$ is estimated only in the range $(0, 38)$ of frames, we can evaluate $P(Q > x)$ for x in the practical range. Figure 10(a) shows a sequence $\{X(k)\}$ of real VBR traffic from a video source. In the video source the video data is compressed according to a modified H.261 and the frame interval is $1/30$ second. As seen in the figure, the traffic sequence indicates complicated behavior and it is clearly different from a Gaussian input rate process. To reproduce such video traffic by simulation, the authors have developed a simulation model which consists of three submodels, a Markovian scene transition model, a spike model for spikes and an AR model for bit rates following spikes [Koba 96, Kura 96].

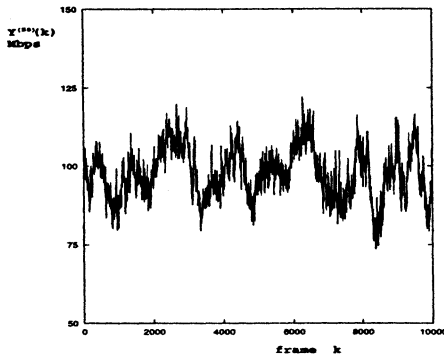
Figures 10(b) and 10(c) show sequences generated by using the simulation model. Figures 10(b) indicates a traffic sequence $\{Y(k)\}$ from one source and Figures 10(c) indicates an aggregated traffic sequence $\{Y^{(50)}(k)\}$ from 50 independent, identical sources. It is reported in [Koba 96, Kura 96] that the simulated traffic resembles the original one very well in four aspects; marginal distribution, autocorrelation function, R/S function and spectral density function. In particular, the simulated sequence, like the original sequence, exhibits long-range dependence in R/S analysis though the model consists only of Markovian-type submodels. The aggregated traffic differs from a Gaussian input rate



(a) Real VBR video traffic



(b) Simulated VBR video traffic



(c) Aggregated traffic

Figure 10 Sample sequences of VBR video traffic

process in the rigorous sense, but from Figure 10(c) we find that it is close to a Gaussian input rate process.

Figure 11 depicts the autocorrelation functions $\gamma(t)$ of $\{X(k)\}$, $\{Y(k)\}$ and $\{Y^{(50)}(k)\}$ estimated from the sequences of length 79686. In this figure $\sigma_\epsilon(t)$ in “ $1.96 \times \sigma_\epsilon(t)$ ” represents the *large sample standard error* and the curve of $1.96 \times \sigma_\epsilon(t)$ indicates the approximate 95% large-sample confidence limits [Vand 83]. In the region $(0, 900)$ where $\gamma(t) > 1.96 \times \sigma_\epsilon(t)$, the null hypothesis that $\gamma(t) = 0$ is rejected with 5% significance. In this sense, we consider that $\gamma(t)$ is valid in the region $(0, 900)$.

We also examine a region in which the estimation error of $\gamma(t)$ is negligible and check whether σ_t is concave or not. Figures 12 and 13 show β_t and σ_t for $\{X(k)\}$, $\{Y(k)\}$ and

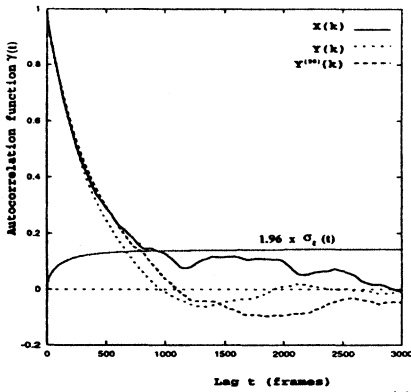


Figure 11 Autocorrelation function $\gamma(t)$

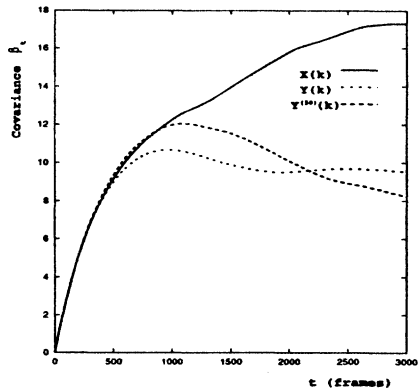


Figure 12 Covariance β_t

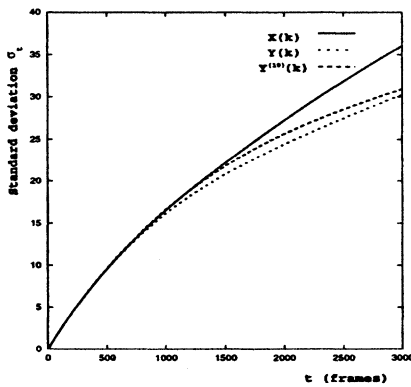


Figure 13 Standard deviation σ_t

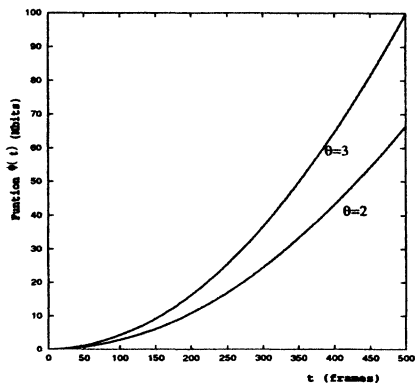


Figure 14 Function $\phi(t)$

$\{Y^{(50)}(k)\}$ with the variances of these processes being normalized so that $\sigma_R^2 = 1$. Since the two sequences $\{Y(k)\}$ and $\{Y^{(50)}(k)\}$ theoretically have the same autocorrelation functions, we examine a region in which these β_t 's are very close with each other. From Figure 12 we find that the two estimates of β_t are almost the same in the region $t < 300$. So we expect that the estimates are good for $t < 300$. From Figure 13 we find that all σ_t 's are concave in the range $(0, 3000)$.

Figure 14 shows the curve of the function (14). In the figure, the parameter θ represents the value $\theta = (c - m)/\sigma_R$. That is, the output rate is given by $c = m + \theta \sigma_R$. Since the estimates of σ_t and β_t are available for $t < 300$, Figure 14 implies that $P(Q < x)$ can be evaluated in the region $x < 20$ Mbits when $\theta \geq 2$.

Figures 15 and 16 show the simulation results and the numerical results for the tail probability $P(Q > x)$, in the region $x < 20$ Mbits, of a fluid queue with $\theta = 2$ and $\theta = 3$ when the number of sources is 50. In the figures the upper broken curve represents the approximation formula (13) with $B = 1$ and the lower broken curve indicates the

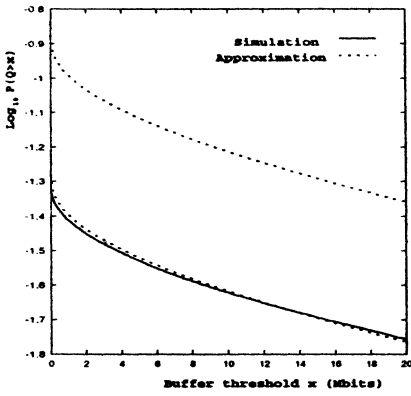


Figure 15 Tail probability for a fluid queue with $\theta = 2$

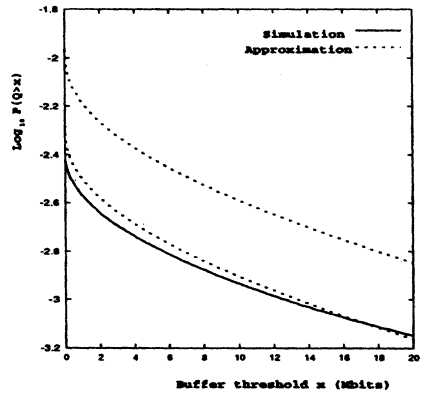


Figure 16 Tail probability for a fluid queue with $\theta = 3$

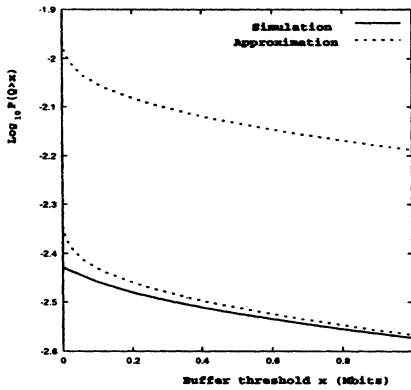


Figure 17 Tail probability of a fluid queue with $\theta = 3$ for $x \in (0, 1)$ Mbits

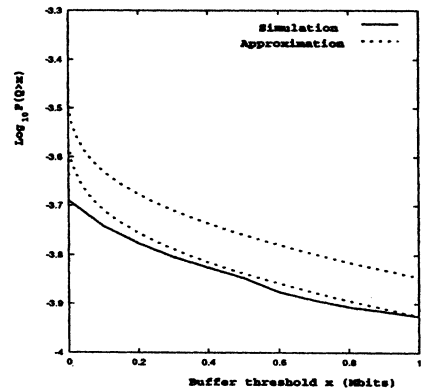


Figure 18 Tail probability of a fluid queue with $\theta = 4$ for $x \in (0, 1)$ Mbits

formula (13) with B set suitably to make the comparison of the decay rates easy. From the figures we find that the tail behavior of the formula (13) almost coincides with the simulation results though various kinds of errors may exist. The covariances β_i 's for the real traffic and the simulated traffic are almost the same in the range $(0, 300)$ as seen in Figure 12. Thus if we simulated the behavior of the tail probability using such real traffic, we would obtain a similar result.

The buffer size of 20 Mbits corresponds to 47000 cells since one cell is 424 bits in ATM networks. In actual ATM multiplexers the buffer size may be thousands of cells, so we also check whether the formula (13) provides a good approximation or not in the range $(0, 1\text{Mbits})$, or equivalently $(0, 2358\text{cells})$.

Figures 17 and 18 show the results for the tail probability $P(Q > x)$ in the region

Table 1 Necessary maximum lag T and $P(Q > 1 \text{ Mbits})$

$\theta = (c - m)/\sigma_R$	3	4	5	6	7
T in frames	45	38	34	30	28
$P(Q > 1 \text{ Mbits})$	2.7×10^{-3}	1.2×10^{-4}	–	–	–

$x < 1$ Mbits of a fluid queue with $\theta = 3$ and $\theta = 4$. We also see that the formula (13) provides a good approximation except for B .

Table 1 indicates the maximum lag T needed for evaluating $P(Q > x)$ in the region $x < 1$ Mbits and the simulation values of $P(Q > 1 \text{ Mbits})$ at $\theta = 3$ and 4. To guarantee that buffer overflow probabilities are less than, say, 10^{-7} or 10^{-9} , the value of the parameter θ must take more than 4. As θ becomes larger, T becomes smaller. Therefore, if $\Lambda(t)$ can be estimated in the interval $(0, 38)$ of frames, namely, $(0, 1.27)$ seconds, we can evaluate $P(Q > x)$ for x in the practical range.

APPENDIX

Here we derive the approximation (9) from Duffied's result [Duff 95].

Let $A_t^{(l)}$ be an input process from source l in the interval $[-t, 0)$ and $W_t^L = \sum_{l=1}^L A_t^{(l)} - Lc$. The tail probability $P(Q^L > x)$ of a fluid queue with output rate Lc is represented by

$$P(Q^L > x) = P(\sup_{t>0} W_t^L > x). \quad (21)$$

Assume that there exist functions a_t, v_t that increase to infinity, such that for each ϑ , the limits

$$\lambda_t(\vartheta) = \lim_{L \rightarrow \infty} \lambda_t^L(\vartheta) \quad \text{and} \quad \lambda(\vartheta) = \lim_{t \rightarrow \infty} \lambda_t(\vartheta) \quad (22)$$

exist as extended real numbers, where

$$\lambda_t^L(\vartheta) = (Lv_t)^{-1} \log E[\exp(\vartheta W_t^L v_t/a_t)]. \quad (23)$$

Duffied proved in [Duff 95] that under mild conditions

$$\lim_{L \rightarrow \infty} L^{-1} \log P(\sup_{t>0} W_t^L > Lb) = -I(b) \quad (24)$$

where

$$I(b) = \inf_{t>0} v_t \lambda_t^*(b/a_t) \quad \text{and} \quad \lambda_t^*(x) = \sup_{\vartheta} \{x\vartheta - \lambda_t(\vartheta)\}. \quad (25)$$

From (24), the approximation of $P(Q^L > x)$ is written as

$$P(Q^L > x) \approx B e^{-L I(x; L)}, \quad (26)$$

where B is a constant.

We will apply the approximation (26) to a Gaussian fluid queue. Let $W_t^{(l)}$, $l = 1, \dots, L$, be independent Gaussian processes with mean $-\bar{\mu}t$ and variance $\bar{\sigma}_t^2$, and set $v_t = t^2/\bar{\sigma}_t^2$ and $a_t = t$. Then the equation (23) is written as $\lambda_t^L(\vartheta) = -\bar{\mu}\vartheta + \vartheta^2/2$. It is independent of both L and t , and hence $\lambda_t(\vartheta)$ and $\lambda(\vartheta)$ in (22) exist. From (25) it follows that $\lambda_t^*(x) = (x + \bar{\mu})^2/2$ and

$$I(b) = \inf_{t>0} \frac{(b + \bar{\mu}t)^2}{2\bar{\sigma}_t^2}.$$

Thus we obtain

$$P(Q^L > x) \approx B \exp \left\{ - \inf_{t>0} \frac{(x + L\bar{\mu}t)^2}{2L\bar{\sigma}_t^2} \right\}. \quad (27)$$

In the case of a single Gaussian process W_t with mean $-\mu t$ and covariance σ_t^2 , we may regard it as a superposition of L independent Gaussian processes with mean $-\bar{\mu}t = -\mu t/L$ and covariance $\bar{\sigma}_t^2 = \sigma_t^2/L$ and use the formula (27). Thus we obtain

$$P(Q > x) \approx B \exp \left\{ - \inf_{t>0} \frac{(x + \mu t)^2}{2\sigma_t^2} \right\}. \quad (28)$$

Acknowledgments

The authors acknowledge Prof. Naoki Makimoto of Tokyo Institute of Technology for his valuable comments and suggestions.

REFERENCES

- [Addi 94] Addie, R.G. and Zukerman, M. (1994) An approximation for performance evaluation of stationary single server queues. *IEEE Trans. on Commun.*, **42** (12), 3150–3160.
- [Bene 63] Benes, V.E. (1963) *General Stochastic Processes in the Theory of Queues*, Addison-Wesley, Reading.
- [Bera 95] Beran, J. (1995) *Statistics for Long-Memory Processes*, Chapman & Hall.
- [Beras 95] Beran, J., Sherman, R., Taqqu, M.S. and Willinger, W. (1995) Long-range dependence in variable-bit-rate video traffic. *IEEE Trans. Commun.*, **43** (2/3/4), 1566–1579.
- [Debi 95] Debicki, K. and Rolski, T. (1995) A Gaussian fluid model. *Queueing Systems*, **20**, 433–452.
- [Duff 95] Duffield, N.G. (1995) Economies of scale in queues with sources having power-law large deviation scalings. *J. Appl. Prob.*, **33**, 840–857.
- [Duffs 95] Duffield, N.G. and O’Connell, N. (1995) Large deviations and overflow probabilities for the general single-server queue with applications. *Math. Proc. Camb. Phil. Soc.*, **118**, 363–374.

- [Gros 96] Grossglauser, M. and Bolot, J.C. (1996) On the relevance of long-range dependence in network traffic. *Proceedings of SIGCOMM'96*, 15–34.
- [Heym 92] Heyman, D.P., Tabatabai, A. and Lakshman, T.V. (1992) Statistical analysis and simulation study of video teleconference traffic in ATM networks. *IEEE Trans. on Circuits and Systems for Video Tech.*, 2 (1), 49–59.
- [Kou 97] Kouchi, M., Makimoto, N. and Kobayashi, K. (1997) Bounds for the Stational Tail Probabilities in Queues with Gaussian Inputs. *Research Reports on Mathematical and Computing Sciences*. Tokyo Institute of Technology, B-333.
- [Koba 96] Kobayashi, K. (1996) *Performance Modeling and Analysis of ATM Multiplexers*, Ph.D.thesis, Tokyo Institute of Technology.
- [Kobas 95] Kobayashi, K. and Takahashi, Y. (1995) Steady-state analysis of ATM multiplexer with variable input rate through diffusion approximation. *Performance Evaluation*, 23 (2), 163–184.
- [Kura 96] Kurasugi, T., Kobayashi, K. and Takahashi, Y. (1996) Data analysis and modeling of ATM coded video traffic with scene changes. *Proceedings of the Symposium on Performance Models for Information Communication networks*, Kyoto, 371–382.
- [Lela 94] Leland, W.E., Taqqu, M.S., Willinger, W. and Wilson, D.V. (1994) On the self-similar nature of Ethernet traffic. *IEEE/ACM Transactions on Networking*, 2, 1–15.
- [Mand 68] Mandelbrot, B.B. and Van Ness, J.W. (1968) Fractional Brownian motions, fractional noises and applications. *SIAM Review*, 10 (4), 422–437.
- [Norr 94] Norros, I. (1994) A storage model with self-similar input. *Queueing Systems*, 16, 387–396.
- [Ryu 96] Ryu, B.K. and Elwalid, A. (1996) The importance of long-range dependence of VBR video traffic in ATM traffic engineering: myths and realities. *Proceedings of SIGCOMM'96*, 3–14.
- [Ster 91] Stern, T.E. and Elwalid, A.I. (1991) Analysis of separable Markov-modulated rate models for information-handling system. *Adv. Appl. Prob.*, 23, 105–139.
- [Vand 83] Vandaele, W. (1983) *Applied Time Series and Box-Jenkins Models*, Academic Press.

Kazutomo Kobayashi received M.E. from Kyushu University and Ph.D in Science from Tokyo Institute of Technology. He has been with NEC Corporation since 1977. He is Senior Principal Researcher of C&C Media Research Laboratories, NEC Corporation. He is mainly interested in performance evaluation of computer systems and communication networks.

Yukio Takahashi received M.Sc. and Ph.D. in Science from Tokyo Institute of Technology. He is Professor of Operations Research in the Department of Mathematical and Computing Sciences, Tokyo Institute of Technology. His main research interest is in the theory and applications of stochastic models including queues and Markov chains.