Peakedness of Stochastic Models for High-Speed Network Traffic

Brian L. Mark[1]
Dept. Elect. Eng., Princeton University,
Princeton, New Jersey, U.S.A.

David L. Jagerman, G. Ramamurthy
C&C Research Laboratories, NEC USA Inc.,
Princeton, New Jersey, U.S.A.

Abstract — *Peakedness* **was originally developed by teletraffic engineers as a tool for characterizing call arrival processes at a trunk group. We generalize the peakedness theory to include a class of stochastic models used in studies of high-speed networks and apply it to the approximate analysis of a statistical multiplexer.**

I. Introduction

In networks based on the Asynchronous Transfer Mode (ATM), information is transmitted asynchronously over high-speed links in the form of 53-byte units called *cells*. Accurate traffic characterization is a crucial step in performing network resource allocation and dimensioning.

II. Generalized Arrival Process

Define a *rate process* $\{R_t, t > 0\}$ to be a strictly stationary random process with finite, nontrivial first two moment measures. The process $\{R_t, t > 0\}$ is to be understood in the generalized function sense with the interpretation that $R_t dt$ represents the *amount* of work arriving in the infinitesimal interval $[t, t+dt)$. The *generalized arrival process* is then defined by

$$N_t = \int_0^t R_\tau d\tau, \tag{1}$$

where N_t represents the amount of work arriving in the interval $(0, t]$.

The standard arrival process defined as a *stationary point process* is a special case with

$$R_t = \sum_{i=1}^{\infty} b_i \ \delta(t - T_i), \tag{2}$$

where b_i is the number of arrivals at the ith arrival epoch, T_i, and $\delta(t)$ is the Dirac delta function. Another special case is the *discrete-level fluid process* with

$$R_t = \sum_{i=1}^{\infty} f_i \ \mathrm{rect}\left(\frac{t - T_i}{T_{i+1} - T_i}\right), \tag{3}$$

where f_i is the fluid flow rate, T_i is the epoch of the ith transition and $\mathrm{rect}(t) = u(t) - u(t-1)$, where $u(t)$ is the unit step function.

III. Generalized Peakedness

We introduce a concept of peakedness for a general arrival process as defined by (1). The arrival process is offered to an *infinite server system* which is represented by an i.i.d. process, $\{D_t, t > 0\}$, with marginal cdf F. Define

$$S_t = \int_0^t 1_{\{D_u > t-u\}} R_u du, \tag{4}$$

with the following interpretation: In the interval $[u, u + du)$, $R_u du$ units of work are offered to a new server, introduced at time u, which removes this work from the system after a duration D_u. Then S_t represents the amount of work present in the system at time t. The *peakedness functional* with respect to the service time cdf F is defined by

$$z[F] = \lim_{t\to\infty} \frac{\mathrm{Var}[S_t]}{E[S_t]}. \tag{5}$$

For the case of an orderly point process, the definitions (4) and (5) reduce to the standard concept of peakedness.

The following result of Eckberg [1] extends to our generalized notion of peakedness:

$$z[F] = 1 + \frac{\mu}{\lambda} \int_{-\infty}^{\infty} [k(x) - \lambda\delta(x)] F^*(x) dx. \tag{6}$$

Here, F^* is the autocorrelation function of $F^c = 1 - F$, $\mu^{-1} = \int_0^\infty F^c(x)dx$ is the mean service time, $\lambda = E[R_t]$ is the mean arrival rate, and $k(\tau) = \mathrm{Cov}(R_{t+\tau}, R_t)$ is the covariance function of the rate process.

IV. Application

The generalized peakedness can be obtained in closed form via (6) for a large class of stochastic traffic models, including the popular Markov modulated fluid models. In particular, the *peakedness function* of a Markov on-off fluid with peak rate r, mean *on* time β^{-1} and mean *off* time α^{-1} with respect to constant service time distribution is given by

$$z_{const}(\mu) = \frac{2r\beta}{(\alpha+\beta)^3}[\alpha + \beta + \mu(1 - e^{-(\alpha+\beta)/\mu})]. \tag{7}$$

Peakedness can also be estimated empirically through measurements of an actual traffic stream and then used to construct a stochastic traffic model.

Lee and Mark [2] propose a method for approximating a general arrival process with a more computationally tractable superposition of two types of on-off Markov fluid sources by matching central moments of the rate process R_t and an index of dispersion measure. Since the peakedness function contains strictly more information about the arrival process than the index of dispersion, a more accurate traffic characterization can be achieved by using the peakedness function (7) to perform the match. We demonstrate the effectiveness of our approach with an application to the analysis of a statistical multiplexer.

References

[1] A. E. Eckberg, "Generalized Peakedness of Teletraffic Processes," *Proc. 10-th International Teletraffic Congress*, Montreal, Canada, 1983.

[2] H. W. Lee and J. W. Mark, "ATM Network Traffic Characterization Using Two Types of On-Off Sources," *INFOCOM '93*, pp. 152-159, 1993.

[1] The first author has been supported by an NSERC Postgraduate Scholarship.