February 1994 COST 242 TD(94)001

ON SOME SIMPLE BLOCKING APPROXIMATIONS FOR THE G/G/c QUEUE

Sándor Molnár
Technical University of Budapest
Department of Telecommunications and Telematics
Stoczek u. 2
H-1111 Budapest, Hungary

Søren Blaabjerg
Ellemtel
Scheelevägen 17
S-223 70 Lund
Sweden

Correspondence with:

Sándor Molnár
Technical University of Budapest
Department of Telecommunications and Telematics
Stoczek u. 2
H-1111 Budapest, Hungary
e-mail: molnar@bme-tel.ttt.bme.hu

Abstract: An appropriate model for the investigation of the call scale traffic in B-ISDN is the G/G/c loss system. The exact evaluation of this system is complicated and even for the subclass of phase-type renewal arrival processes is limited to rather small system sizes. In this paper we consider two simple approximations. Both are based on matching the mean and variance of the occupancy distribution, and it is pointed out how the concept of generalized peakedness makes the evaluation of the variance very simple for a large subclass of the G/G/∞ systems. The first approximation is based on BPP distributions while the second approach is based on the concept of maximizing entropy which turns out to be closely related to approximations with the normal distribution. The results for the infinite server system is finally used to compute approximate blocking probabilities for the finite capacity system.

1 INTRODUCTION

Estimation of blocking probabilities in traditional telephone networks and the related issue of network dimensioning has a long tradition initiated with the pioneering work of Erlang, see [BHJ48]. Many years of experience has shown that the Poisson process is a reasonable model of the process of telephone call requests, and its nice mathematical properties have to a great extend simplified the analysis of such networks resulting in properties such as product form solutions and insensitivity of holding time distributions [Kel86].

In order to analyse overflow traffic in e.g alternative routing, it has been necessary to include less tractable arrival processes like renewal, or Markov Modulated Poisson Processes. For these cases it has normally been assumed that the holding time distribution is exponential, see [Sys60], [Del81]. Alternatively, the class of state dependent Poisson processes has been used for the arrival process.

It is expected that in a future ATM based B-ISDN supporting a large variety of services, the arrival process of new connection request will differ from service to service and will often be quite different from a Poisson process. It is also expected that the holding time in most cases will differ significantly from the exponential distribution. More general queueing models are needed to achieve appropriate call scale models in ATM. These expectations actualize the investigation of networks with G/G/c queues both as loss and delay systems.

In order to be able to dimension such networks, the first step is to study the G/G/c loss system, and for mathematical reasons the G/G/∞ system is also of interest. So far one of the few approaches to obtain the exact steady state distribution of a G/G/c system is by restricting the arrival process to be of the type Markovian Arrival Process, see [Neu79] and to assume the holding time distribution to be of phase type. Then the system can be analysed through the usual Markovian approach [Neu81] and [Han83]. However, the size of the state space very quickly exceeds any in practice manageable size even when the sparseness of the transition matrix is taken into account, thereby limiting the feasible models to those with a rather limited number of phases in the arrival process, a very limited number of phases of the holding time distribution and only a small number of servers. Furthermore any exact solution approach is far to complicated to be used for dimensioning purposes.

The analysis is significantly simplified when the holding time distribution is exponential, and the solution for the GI/M/∞ and GI/M/c has also been known for about fifty years, see [Pal43]. However, deriving the occupancy distribution is numerically challenging and in practice impossible when the offered traffic and number of servers grow. It has been common practice to approximate the occupancy distribution of the GI/M/∞ and GI/M/c systems by using the state dependent Poisson process of the type BPP, see [Del81] and [Ive87]. When the offered traffic increases the BPP distribution converges towards a normal distribution [Fel50], and for high capacity systems it has been suggested to use the normal distribution [Del81].

By restricting the considerations to finding the mean and variance of the distribution of the number of occupied servers in an infinite server group, [Eck83] presents formulas for finding the peakedness factor (defined as the fraction variance to mean) for a general stationary arrival process and general holding time distribution. The fact that it is relatively easy to obtain the mean and variance of the distribution, but difficult to obtain the distribution itself, calls for approximations.

The objective of the present paper is three fold. First, we shall investigate how the occupancy distribution and peakedness can be obtained in the G/G/∞ system. In the case of computing the exact occupancy distribution we will only consider phase type renewal arrival processes. The peakedness

study is greatly simplified by the work of Eckberg, and we demonstrate how a closed form expression can be derived for the class of Coxian holding time distributions when the only restriction on the arrival process is that it must be stationary and the Laplace transform of the renewal function must be available. Secondly, we study two approximations for the occupancy distributions based on matching the mean and variance. The first one is the classical BPP approximation while the second one is the distribution which is obtained by maximizing the entropy subject to matching of the mean and variance. It is shown how the maximum entropy method enforces a distribution which is a discrete version of a normal density function restricted to the positive line. Finally, approximations for the blocking probability of the finite capacity system is suggested based on the derived occupancy distribution of the infinite system.

In this paper we do not investigate the issue of how the third and higher moments affects the occupancy distribution but refer to the work by [Hol73] in which he derives some bounds based on the Laplace transform of the occupancy distribution. Furthermore, we do not explicitly consider the case where calls have different bandwidth demand. A straightforward extension to the multi rate case exists for the Eckberg peakedness but for the exact occupancy distribution an extension is complicated and will not be treated in this paper.

2 OCCUPANCY DISTRIBUTION IN THE G/G/∞ CASE

In this section the model under study is an infinite server group which is offered traffic from a stationary process S with arrival intensity m. All servers act independently of each other and have the same holding time distribution H which is assumed general but with finite mean $1/\mu$ and variance σ^2.

The main interest is the stationary distribution of the number of occupied servers, the occupancy distribution. This distribution itself is not possible to compute in an exact way for a general stationary arrival process. We restrict the arrival process to be renewal and the interarrival time to be of phase-type. The occupancy distribution for this case can be obtained by applying the results of [RNe80]. The peakedness of the occupancy distribution can be computed for general stationary arrival process, as shown in [Eck83].

2.1 Exact Occupancy Distribution in Case of Phase Type Arrivals

Ramaswami and Neuts have studied the infinite-server queue with phase type input and general holding time distribution [RNe80]. Here we present very briefly the main results, which are applied to obtain the exact occupancy distribution.

A phase type distribution is defined as a distribution function for the time until absorption in an $(n+1)$ state Markov chain with n transient state and one absorbing state. The infinitesimal generator is of the form

$$Q = \begin{bmatrix} T & \boldsymbol{T}^{\circ} \\ \mathbf{0} & 0 \end{bmatrix}$$

where $T = (T_{ij})$ is a non-singular $n \times n$ matrix such that $T_{ii} < 0$ and $T_{ij} \geq 0$ for $i \neq j$, and $\boldsymbol{T}^{\circ} \geq \mathbf{0}$ is an n-vector satisfying $T\boldsymbol{e} + \boldsymbol{T}^{\circ} = \mathbf{0}$, where $\boldsymbol{e}' = (1, \ldots, 1)$. A phase type renewal process can be obtained by resetting the Markov chain Q according to an initial probability vector $(\mathbf{a}, a_{n+1})$ after each transition to the absorbing state. We consider as a model of an arrival process this PH renewal process, which has the infinitesimal generator $Q^* = T + T^{\circ}A^{\circ}$, where $A^{\circ} = \text{diag}\,(a_1, \ldots, a_n)$ and $T^{\circ} = (\boldsymbol{T}^{\circ}, \ldots, \boldsymbol{T}^{\circ})$.

Based on the basic system of differential equations (derived in [RNe80]) we get an infinite system of differential equations for the time dependent occupancy distribution. Truncated at a sufficient large value of the index k they can be solved numerically.

Define $\boldsymbol{g}_k(t)$ in the following way: the i^{th} entry of $\boldsymbol{g}_k(t)$ is given by

$$g_{ki}(t) = P[X(t) = k \mid X(0) = 0,\ J(0) = i], \qquad t \geq 0$$

for $k \geq 0$, $i,j = 1, \ldots, n$, where the number of occupied servers and the phase of the arrival process at time $t+$ are denoted by $X(t)$ and $J(t)$, respectively.

The differential equations are,

$$\frac{d}{dt}\boldsymbol{g}_0(t) = [T + H(t)\, T^{\circ}A^{\circ}]\, \boldsymbol{g}_0(t), \quad \boldsymbol{g}_0(0) = \boldsymbol{e} \tag{2.1}$$

and for $k \geq 1$

$$\frac{d}{dt}\boldsymbol{g}_k(t) = [T + H(t)\, T^{\circ}A^{\circ}]\, \boldsymbol{g}_k(t) + \{1 - H(t)\}\, T^{\circ}A^{\circ}\boldsymbol{g}_{k-1}(t), \quad \boldsymbol{g}_k(0) = \boldsymbol{0} \tag{2.2}$$

The main advantages of this approach that it produces both the transient and the steady state occupancy distributions, moreover, the complexity of the solution is independent on the holding time distribution. However, for steady state analysis, this solution approach is rather complex, and we have experienced difficulties in the numerical calculations for cases in which the steady state probability of an empty system is very small.

2.2 The Peakedness of the Occupancy Distribution

In [Eck83] the peakedness (variance to mean ratio) of the occupancy distribution has been studied assuming only that the arrival process is stationary. [Eck83] presents several different formulations for the peakedness, but here we shall restrict ourselves to two different formulations.

Let $U(x)$ denote the renewal function of the process S that is: $U(x) = E[N(a, a+x)]$ with $N(a,b)$ denoting the number of arrivals in the interval $]a,b]$, when an arrival occurred at time a.

Furthermore, define $H_2^c(x) = \int_{-\infty}^{\infty} (1 - H(u))\,(1 - H(u-x))\, du$

According to formula (3) in [Eck83] then the peakedness Z of the complementary holding time distribution 1-H is

$$Z(H) = 1 + 2\mu \int_{0-}^{\infty} H_2^c(x)\, dU(x) - \frac{m}{\mu} \tag{2.3}$$

If the holding time is exponentially distributed with mean $1/\mu$, the peakedness formula reduces to:

$$Z_{exp}(\mu) = 1 + U^*(\mu) - \frac{m}{\mu} \tag{2.4}$$

in which U^* denotes the Laplace-Stieltjes Transform of U.

Restricting the holding time distributions to be of the class which,

$$1 - H(x) = \int_{-\infty}^{\infty} e^{-xt} a(t)\, dt \quad \text{for } x > 0 \tag{2.5}$$

and a is a generalized function, [Eck83] obtains the following relation between the peakedness of H and the peakedness of an exponential holding time distribution with mean $1/\mu$.

$$Z[H] = 1 + 2\mu \int_{0-}^{\infty} \alpha(y)\, (Z_{exp}(y) - 1)\, dy \tag{2.6}$$

where

$$\alpha(y) = a(y) \int \frac{a(x)}{x+y} dx \tag{2.7}$$

In Appendix A it is shown how expression (2.5), (2.6) and (2.7) are applied when the holding time distribution is Coxian.

3 APPROXIMATIONS FOR THE DISTRIBUTION

Here we shall consider two different approximation procedures both based on matching the mean and variance of the occupancy distribution in the infinite server case.

3.1 The BPP Approximation

The BPP arrival process is a state dependent Poisson process characterized by two parameters α and β such that the Poissonian arrival intensity when k servers are occupied is $\alpha + k\beta$. In the case $\beta = 0$ it reduces to the plain Poisson process, for $\beta < 0$ it represents a process of less variability than Poisson (finite source model), and for $\beta > 0$ it represents a process of higher variability than Poisson. The mean and variance of the occupancy distribution turns out to be:

$$M = \frac{\alpha}{\mu - \beta} \qquad V = \frac{\alpha\mu}{(\mu - \beta)^2} \tag{3.1}$$

where $1/\mu$ is the mean holding time.

With the mean M and variance V given (from e.g the Eckberg approach) we suggest to approximate the occupancy distribution with the BPP distribution with the same mean and variance. The BPP parameters α and β then should be chosen as:

$$\beta = 1 - \frac{1}{Z} \qquad \alpha = M(1 - \beta) \tag{3.2}$$

assuming a mean holding time of 1.

3.2 The Maximum Entropy Approximation

The concept of entropy has played a significant role in various technical disciplines like thermodynamics, statistical mechanics and information theory. The idea also appears in the mathematical theory of interconnecting networks and in the queueing theory [Ben65], [Kou86]. In queueing theory we only know of its existence in the 1-server case and usually it has been applied in a way where only average quantities like mean queue length and utilization has been matched. Here we shall apply this technique on a many server loss system where also the variance is matched.

Consider a stationary stochastic process $X(t)$ in a discrete state space and let p_i be the probability of being in state i. The entropy of $X(t)$ with stationary distribution $\{p_i\}$ is defined as:

$$H(\boldsymbol{p}) = -\sum_i p_i \ln p_i$$

If a certain value comes out with probability 1, the entropy is 0, and in all other cases it is larger than 0. If the outcome is constrained to a finite set of values, the distribution which maximizes the entropy is the uniform distribution in which all outcome are equally likely. The less we know of the system, the larger the entropy is.

The idea here is, as an approximation for the occupancy distribution, to take the one which maximizes the entropy under the constrains that

- it should be a proper probability distribution i.e $\sum_{i=0}^{\infty} p_i = 1$
- the mean should be correct i.e $\sum_{i=0}^{\infty} i p_i = E(X) = \frac{m}{\mu}$
- and the second moment should be correct i.e. $\sum_{i=0}^{\infty} i^2 p_i = E(X^2) = \frac{m}{\mu}\left(Z(1-H) + \frac{m}{\mu}\right)$

In Chapter 8.4.1 of [Har92] the following theorem is presented and proved

THEOREM: The probability mass function $\{p_i\}$ which maximizes $H(\boldsymbol{p}) = -\sum_i p_i \ln p_i$

subject to $\sum_{i=0}^{\infty} p_i = 1$

and $\sum_i f_j(i)\, p_i = \bar{f}_j$ for $1 \le j \le k$

(where $\{\bar{f}_j \mid (1 \le j \le k)\}$ are prescribed mean values of the functions $\{f_j\}$)

is:

$$p_i = g \prod_{j=1}^{m} x_j^{f_j(i)}$$

where g is the normalization constant.

If this result is applied to a single server queue and the mean queue length is matched, the queue length distribution which maximizes entropy is geometrical thus yielding the exact distribution for the queue length in the M/M/1 case.

When matching the first and second moment in the infinite server case, a straightforward application of the theorem shows that the occupancy distribution which maximizes entropy is:

$$p_i = P\{X = i\} = g x_1^i x_2^{i^2} = g e^{i \ln x_1} e^{i^2 \ln x_2}$$

Thus the maximum entropy approach enforces a distribution which comes from sampling the normal density function at non-negative integer values. The 3 equations needed to match the mean, variance and obtain a proper distribution is:

$$\sum_{i=0}^{\infty} g x_1^i x_2^{i^2} = 1 \qquad \sum_{i=0}^{\infty} g i x_1^i x_2^{i^2} = E\{X\} \qquad \sum_{i=0}^{\infty} g i^2 x_1^i x_2^{i^2} = E\{X^2\}$$

and we have only been able to solve them by heuristic methods. However, applying the corresponding continuous approach fitting a normal density function restricted to the positive line yields a system of equation which can be solved in an exact way as demonstrated in Appendix B. Our numerical results indicates that the difference between the continuous and the discrete results are very small.

4 ESTIMATION OF BLOCKING PROBABILITIES

4.1 Truncation of the occupancy distribution

So far we have only considered the system with *infinitely* many servers. In order to arrive at the occupancy distribution of the finite system with c servers we first assume that the system under consideration can be modelled as a time reversible Markov process. When this is the case we have the following corollary from [Kel79] (corollary 1.10)

Corollary: If a reversible Markov process with state space S and equilibrium distribution $\pi(i)$, $i \in S$ is truncated to a set $A \subset S$ then the resulting Markov process is reversible in equilibrium and has equilibrium distribution

$$\pi_{\text{truncation}}(i) = \frac{\pi(i)}{\sum_{j \in A} \pi(j)} \qquad i \in A$$

In the case where the arrival process is a state dependent Poisson process (this includes the BPP case) this approach is exact.

For all cases where the system is not a reversible Markov process we will approximate the occupancy distribution in the finite case by the same formula.

As blocking probability we use the traffic congestion defined as:

$$\text{traffic congestion} = \frac{\text{offered traffic} - \text{carried traffic}}{\text{offered traffic}}$$

where the offered traffic is the mean number of occupied servers in the infinite system (mean holding time = 1), and the carried traffic is the mean of the truncated distribution.

4.2 When occupancy distribution is a Normal Distribution

When the occupancy distribution is a truncated normal distribution, we are able to find a closed form expression for the traffic congestion.

We have the infinite server occupancy distribution $f(x) = \begin{cases} ge^{-\alpha x}e^{-\beta x^2} & \text{for } x > 0 \\ 0 & \text{for } x < 0 \end{cases}$ implying that

$$\text{offered traffic} = g\sqrt{\frac{\pi}{\beta}}e^{\frac{\alpha^2}{4\beta}}\left\{\left(-\frac{\alpha}{2\beta}\right)\Phi\left(\frac{-\alpha}{\sqrt{2\beta}}\right)+\frac{1}{\sqrt{2\beta}}\varphi\left(\frac{-\alpha}{\sqrt{2\beta}}\right)\right\} \tag{4.1}$$

where Φ is the standard normal distribution and φ is the corresponding density function. Thereby

$$\text{carried traffic} = \frac{\int_0^c xge^{-(\beta x^2+\alpha x)}dx}{\int_0^c ge^{-(\beta x^2+\alpha x)}dx}$$

$$= \frac{\frac{-\alpha}{2\beta}\left(\Phi\left(\frac{-\alpha}{\sqrt{2\beta}}\right)-\Phi\left(-\sqrt{2\beta}c-\frac{\alpha}{\sqrt{2\beta}}\right)\right)+\frac{1}{\sqrt{2\beta}}\left(\varphi\left(\frac{\alpha}{\sqrt{2\beta}}\right)-\varphi\left(\sqrt{2\beta}c+\frac{\alpha}{\sqrt{2\beta}}\right)\right)}{\Phi\left(\sqrt{2\beta}c+\frac{\alpha}{\sqrt{2\beta}}\right)-\Phi\left(\frac{\alpha}{\sqrt{2\beta}}\right)} \tag{4.2}$$

thereby yielding a closed form expression for the blocking probability.

4.3 In case of BPP arrivals

For the BPP approximation no closed form expression for the carried traffic is available. Instead, it is possible to use the call congestion (= probability of an arbitrary call is blocked) of the truncated system which by the arrival theorem can be found directly from the time congestion (= the probability of a full

system at an arbitrary point in time) of the system with one fewer customer. However, the call congestion can differ significantly from the traffic congestion since the truncation immediately effects the arrival process (states above the truncation level disappear).

4.4 In case of renewal input and exponential holding times

For the GI/M/c system Theorem 4 chapter 4 in [Tak62] provides the following relation between the occupancy distribution $p_i^{(t)}$ at an arbitrary time and the occupancy distribution $p_i^{(a)}$ just prior to an arrival

$$p_i^{(t)} = E\{X\}\frac{p_{i-1}^{(a)}}{i} \quad \text{for} \quad i \in \{1,..,c\} \qquad \text{and} \quad p_0^{(t)} = 1 - E\{X\}\sum_{i=1}^{c}\frac{p_{i-1}^{(a)}}{i} \tag{4.3}$$

$E\{X\}$ is the mean of the number of occupied servers in the corresponding infinite server system. Using (4.3) the call congestion can be written as

$$p_c^{(a)} = 1 - \sum_{k=0}^{c-1} p_k^{(a)} = 1 - \sum_{k=0}^{c-1}\frac{p_{k+1}^{(t)}(k+1)}{E\{X\}} = 1 - \frac{\sum_{k=1}^{c} kp_k^{(t)}}{E\{X\}} = \frac{E\{X\} - E\{X_{tr}\}}{E\{X\}}. \tag{4.4}$$

Here $E\{X_{tr}\}$ is the mean of the number of occupied servers in the finite server system. This small calculation shows that the call congestion equals to the traffic congestion.

For the GI/M/c system the occupancy distribution is computed by the method of Takács (Chapter 3 and 4 in [Tak62]).

5 NUMERICAL RESULTS

We have found the exact blocking probability of the GI/M/c case for four different interarrival time distributions, the deterministic, the Erlang-4, a hyperexponential with squared coefficient of variation $c_a^2 = 4$ and a hyperexponential with $c_a^2 = 20$. In all cases the offered traffic is kept equal to 10 and the number of servers is 15. Then we have evaluated four different approximative blocking probabilities.

In the column named "truncated", we have used the occupancy distribution in the infinite system at arbitrary time, and found the traffic congestion by truncation and renormalizing it.

In the column "BPP-app", we have used the BPP occupancy distribution, normalized it and used the relation $P_{\text{call cong}} = P_{\text{time cong}} \left(1 + \frac{c}{M}(Z-1)\right)$ from formula 16 in [Del81] to get the blocking probability from the time congestion.

In the column "tr. normal" we have used the traffic congestion given from (4.1) and (4.2).

Finally in column "Hayward" we have applied the classical Hayward approximation in which the blocking probability in a system with c servers and an offered traffic A with peakedness Z is approximated by the blocking probability in a system with c/Z servers and offered Poisson traffic A/Z.

Table 1: Approximate and exact blocking probabilities (offered traffic = 10, c = 15)

	exact	truncated	BPP-app	tr. normal	Hayward
$c_a^2 = 0$	0.67%	0.60%	0.40%	0.78%	0.90%
$c_a^2 = \frac{1}{4}$	1.31%	1.24%	1.03%	1.42%	1.55%
$c_a^2 = 4$	13.5%	15.0%	12.1%	15.2%	11.6%
$c_a^2 = 20$	25.1%	44.4%	32.2%	42.5%	33.9%

Except for the peaky arrival process with $c_a^2 = 20$ all the approximations performs reasonable well. The BPP approximation underestimates in the cases with a regular arrival process while the truncated normal distribution overestimates for peaky arrival processes.

6 CONCLUSION

From the investigation we have done and from our numerical results we conclude that it is possible to obtain reasonable approximations for the occupancy distribution in the G/G/∞ by approximating it with either a BPP distribution or a normal distribution restricted to the positive line with the same mean and variance. It is also to be concluded that the concept of generalized peakedness offers a very simple way of obtaining the variance of the exact occupancy distribution for a large subset of the G/G/∞. Approximate formulas for the blocking probability in the corresponding finite capacity system has been presented and evaluated with satisfactory result in the case with exponential holding time.

Since our approximation procedure is as easy to apply in a heterogeneous environment in which connections with different bandwidth requirements and different holding time distributions are offered to the same network, an obvious future activity is to evaluate the accuracy in these cases. Due to the large dimension of the state space involved this is a challenging task.

How to apply the approximation on the network dimensioning problem is another interesting subject for future research.

APPENDIX A: DERIVING THE PEAKEDNESS FOR COXIAN HOLDING TIME

In this section we derive the peakedness formula for cases with Coxian holding time distributions.

Consider a Coxian distribution represented by a weighted sum of generalized Erlang distributions as shown in Fig. A1. For simplicity reasons we assume that $\lambda_i \neq \lambda_j$ for $i \neq j$ but the general case can be included with only minor changes in the expressions.

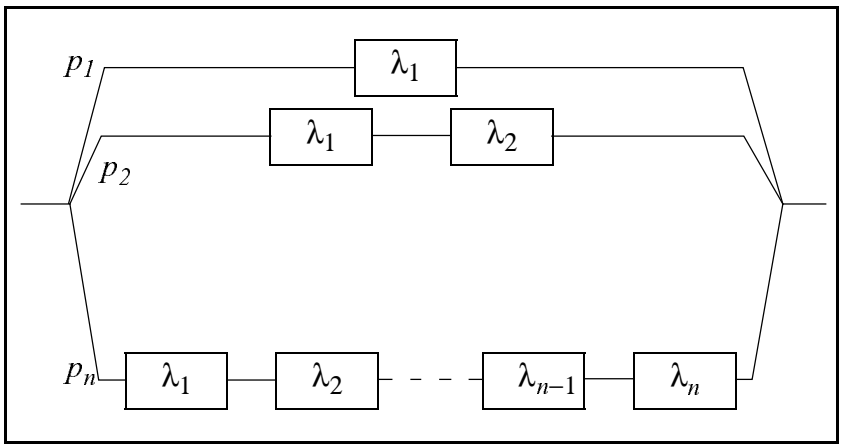

Fig. A1: Coxian distribution represented as a weighted sum of generalized Erlang distributions

For the Coxian distributions *a(t)* in (2.5) can be written as a series of delta functions[1]:

$$a(t) = \sum_{i=1}^{n} \left(\prod_{j=1}^{i-1} \frac{\lambda_j}{\lambda_j - \lambda_i} \right) \left(\sum_{j=i}^{n} p_j \prod_{k=i+1}^{j} \frac{\lambda_k}{\lambda_k - \lambda_i} \right) \delta_{\lambda_i} \tag{A.1}$$

where λ_i ($\lambda_i \neq \lambda_j$ for $i \neq j$) and p_i, $1 \leq i \leq n$, are the intensities and branch probabilities of an n-branch Coxian distribution, respectively. (See Fig. A1.) By applying (2.7) the result of the integral is

$$b(y) = \sum_{i=1}^{n} \left(\prod_{j=1}^{i-1} \frac{\lambda_j}{\lambda_j - \lambda_i} \right) \left(\frac{1}{y + \lambda_i} \right) \left(\sum_{j=i}^{n} p_j \prod_{k=i+1}^{j} \frac{\lambda_k}{\lambda_k - \lambda_i} \right) \tag{A.2}$$

Multiplying by *a(y)* and substituting the obtained $\alpha(y)$ into (2.6) we get the peakedness:

$$Z = 1 + 2\mu \sum_{l=1}^{n} p_l \sum_{i=1}^{l} \left(\prod_{j=1}^{i-1} \frac{\lambda_j}{\lambda_j - \lambda_i} \right) \left(\prod_{k=i+1}^{l} \frac{\lambda_k}{\lambda_k - \lambda_i} \right) b(\lambda_l) \, (Z_{exp}(\lambda_l) - 1) \tag{A.3}$$

---

1. When $\lambda_i = \lambda_j$ for some i and j, derivatives of delta functions appear in the expression for $a(t)$.

where Z_{exp} is given in (2.4).

The importance of (A.3) should be realized. For any stationary arrival process for which the Laplace transform of the renewal function is available, which includes e.g. renewal processes, Markov renewal processes and doubly stochastic processes, insertion of formula (2.4) in (A.3) provides a simple closed form expression for the peakedness of any Coxian holding time distribution.

APPENDIX B: MATCHING MEAN AND VARIANCE

The truncated normal density function is given as:

$$f(x) = \begin{cases} ce^{-\alpha x}e^{-\beta x^2} & \text{for} \quad x > 0 \\ 0 & \text{for} \quad x < 0 \end{cases} \tag{B.1}$$

and has moment generating function given by

$$M(s) = \int_0^\infty e^{sx} ce^{-\alpha x}e^{-\beta x^2} dx = c\sqrt{\frac{\pi}{\beta}} e^{\frac{(s-\alpha)^2}{4\beta}} \Phi\left(\frac{s-\alpha}{\sqrt{2\beta}}\right) \tag{B.2}$$

Thereby the system of equations which needs to be solved to match mean and variance and still have a proper probability distribution is:

$$c\sqrt{\frac{\pi}{\beta}} e^{\frac{\alpha^2}{4\beta}} \Phi\left(\frac{-\alpha}{\sqrt{2\beta}}\right) = 1$$

$$c\sqrt{\frac{\pi}{\beta}} e^{\frac{\alpha^2}{4\beta}} \left\{\left(-\frac{\alpha}{2\beta}\right)\Phi\left(\frac{-\alpha}{\sqrt{2\beta}}\right) + \frac{1}{\sqrt{2\beta}}\varphi\left(\frac{-\alpha}{\sqrt{2\beta}}\right)\right\} = E\{X\}$$

$$c\sqrt{\frac{\pi}{\beta}} e^{\frac{\alpha^2}{4\beta}} \left\{\left(\frac{\alpha^2}{4\beta^2} + \frac{1}{2\beta}\right)\Phi\left(\frac{-\alpha}{\sqrt{2\beta}}\right) - \frac{\alpha}{(2\beta)^{\frac{3}{2}}}\varphi\left(\frac{-\alpha}{\sqrt{2\beta}}\right)\right\} = E\{X^2\} \tag{B.3}$$

Dividing the normalization equation up into the two moment equations, and multiplying the results with $\sqrt{2\beta}$ and 2β respectively, yields:

$$G(-x) - x = \sqrt{2\beta}E\{X\} \quad \text{and} \quad x^2 + 1 - xG(-x) = 2\beta E\{X^2\}$$

where $x = \frac{\alpha}{\sqrt{2\beta}}$ and where $G(x) = \frac{\varphi(x)}{\Phi(x)}$. Squaring the left hand equation and dividing the result into the right hand equation yields the following equation with x as the only unknown.

$$\frac{x^2+1-xG(-x)}{(G(-x)-x)^2} = \frac{E\{X^2\}}{E\{X\}^2}$$

The function $x \to \frac{x^2+1-xG(-x)}{(G(-x)-x)^2}$ is increasing with range]1,2[which can be seen by applying the inequality $\frac{1}{x}-\frac{1}{x^3}+\frac{3}{x^5}-\frac{15}{x^7} \le \frac{\Phi(-x)}{\varphi(x)} \le \frac{1}{x}-\frac{1}{x^3}+\frac{3}{x^5}$ valid for $x>0$.

This implies that for occupancy distributions with $0 < Z < E\{X\}$ a unique truncated normal distribution with same mean and variance exists.

REFERENCES

[Ben65] V.E. Benes, "Mathematical Theory of Interconnecting Networks", Academic Press, 1965.

[BHJ48] E. Brockmeyer, H.L. Hallstrom, A. Jensen, "The Life and Works of A.K. Erlang", Academy of Technical Sciences, Copenhagen, 1948.

[Del81] L.E.N. Delbrouck, "A Unified Approximate Evaluation of Congestion Functions for Smooth and Peaky Traffic", IEEE Trans. on Comm. Vol 29, Feb.1981, pp. 85-91.

[Eck83] A.E.Eckberg, "Generalized Peakedness of Teletraffic Processes", ITC-10 Session 4.4b, Montreal, 1983.

[Fel50] W. Feller, "An Introduction to Probability Theory and Its Applications", Wiley 1950

[Han83] S. Hansen, "Phase Type Distributions in Queueing Theory", Ph.D Thesis no. 41, IMSOR, Denmark, 1983.

[Har92] P.G. Harrison, "Performance Modelling of Communications Networks and Computer Architectures", Addison Wesley 1992.

[Hol73] J.M. Holtzman, "The Accuracy of the Equivalent Random Method with Renewal Input", Bell System Technical Journal, Vol. 52, pp. 1673-1679, 1973.

[Ive87] V.B. Iversen, "The Exact Evaluation of Multi Service Loss Systems with Access Control", Teleteknik (English Edition) Vol 31, 1987.

[Kel79] F.P. Kelly, "Reversibility and Stochastic Networks", Wiley 1979.

[Kel86] F.P. Kelly, "Blocking Probabilities in Large Circuit Switched Networks", Adv. Appl. Prob., **18**, pp. 473-505, (1986)

[Kou86] D.D. Kourvatsos, "Maximum Entropy and the G/G/1/N queue", Acta Informatica, Vol. 23, pp. 545-465, 1986.

[Neu79] M.F. Neuts, "A Versatile Markovian Point Process", J. Appl. Prob. **16**, 1979, pp. 764-779

[Neu81] M.F. Neuts, "Matrix Geometric Solutions in Stochastic Models", John Hopkins University Press, 1981.

[Pal43] C. Palm, "Intensitätsschwankungen im Fernsprechverkehr", Ericsson Technics, 44, pp. 1-189, 1943.

[RNe80] V. Ramaswami, M.F. Neuts, "Some Explicit Formulas and Computational Methods for the Infinite-Server Queues with Phase-Type Arrivals", J. Appl. Prob. **17**, 498-514, 1980.

[Sys60] R. Syski, "Introduction to Congestion Theory in Telephone Systems", Oliver and Boyd 1960.

[Tak62] L. Takács, "Introduction to the Theory of Queues", Oxford University Press, 1962