ITC19/ Performance Challenges for Efficient Next Generation Networks 1559-1570
LIANG X.J. and XIN Z.H.(Editors)
V.B. IVERSEN and KUO G.S.(Editors)
Beijing University of Posts and Telecommunications Press

The Unified Approach for Teletraffic Models to Convert Recursions for Global State Probabilities into Stable Form

Iversen, V.B. [1] and Stepanov, S.N. [2]

[1] COM Center
Building 345v, Technical University of Denmark, DK-2800 Kgs. Lyngby, Denmark
vbi@com.dtu.dk
[2] Sistema Telecom
125047, Moscow, 1st Tverskay-Yamskaya 5, Russia
stepanov@sistel.ru

Abstract. In this paper we describe a unified approach for teletraffic models to convert recursions for global state probabilities into stable form. At each step of recurrence we deal with normalised values of global state probabilities used for estimation of main stationary performance measures. This makes it suited for solving problems of dimensioning. In addition to being numerically stable, the main positive features of the suggested approach are its simplicity and small storage and computational requirements.
Keywords: teletraffic models, global state probabilities, stable recurrence, dimensioning.

1 Introduction

One of the main problems of an operator is to determine the volume of resources that are necessary for serving given input flows of demands with prescribed characteristics of QoS. As the resource is considered number of circuits, value of bandwidth represented by integer number of units, number of radio channels, and so on. The necessary volume is determined by comparison of suitably chosen performance measure with prescribed level of system functioning.

The general scheme of solving the formulated problem is as follows. In the first place it is necessary to construct the mathematical model that gives the required performance measure. Usually this is done in terms of queueing theory. Then based on numerical methods the algorithm is elaborated for estimating the performance measure. After performing the searching the necessary value of telecommunication volume is found. The numerical complexity of an approach is estimated by the numerical complexity for finding the performance measures for given input parameters multiplied by the number of timessearching.

The significance of solving the formulated problem for optimal distribution of telecommunication resources has forced researchers to look for ways to decrease the amount of computational work. In this paper we described qa unified approach for finding the desired solution and formulate the necessary conditions for it's realization. We construct an approach that allows us to find the solution in one single iteration. It is important to notice that the method described gives numerically stable algorithm because at each step it deals with normalised values of state probabilities calculated for subset of states when an arriving request for resource is blocked. The proposed method is a direct generalisation of the well-known recursion for Erlang's B-formula. We revive the corresponding recurrence in Section 2. The general form of solution will be presented in Section 3. The application of the unified approach will be shown for a number of cases in Section 4 (where alternative and very simple solutions will be given for some already known results) and in Section 5 (where new solutions will be presented).

2 Recursions for Erlang's B-formula

Let us consider a single link traffic model, where the link transmission capacity is represented by k basic bandwidth units. In the model we have one Poissonian flow of demands for connections with intensity λ which we for simplicity will name by calls. We assume that each call uses one bandwidth unit for the time of connection having mean value equal to one. To simplify the description of the model dynamics we assume that the duration of connection (service time) is exponentially distributed with parameter equal to one, but most of the obtained results are valid for general service time distributions.

Let $i(t)$ be number of calls being served at time t. The dynamics of the model are described by a Markov process $r(t)$ having one component $r(t) = i(t)$. The process $r(t)$ takes values in the finite set of states $S = \{0, 1, \dots, k\}$ defined in accordance with link capacity. Let us by $P(i)$ denote the unnormalised values of stationary probabilities of $r(t)$. After normalisation, the value $P(i)$ denotes the mean proportion of time when exactly i connections are established or i bandwidth units are occupied. From the system of state equations follows:

$$P(i) = \frac{\lambda}{i}\, P(i-1), \quad i = 1, 2, \dots, k\,. \tag{1}$$

Together with the normalising condition relation (1) gives the Erlang–B loss formula for estimation of the mean proportion of time when all k bandwidth units are occupied:

$$E_k(\lambda) = \frac{\frac{\lambda^k}{k!}}{1 + \lambda + \frac{\lambda^2}{2!} + \cdots + \frac{\lambda^k}{k!}}\,. \tag{2}$$

For dimensioning purposes (2) can be written in the form of recurrence:

$$E_k(\lambda) = \frac{\frac{\lambda}{k} \cdot E_{k-1}(\lambda)}{1 + \frac{\lambda}{k} \cdot E_{k-1}(\lambda)}\,, \quad k = 1, 2, \dots, \quad E_0(\lambda) = 1\,. \tag{3}$$

The recursion (3) works with normalised values $E_k(\lambda)$ so its implementation does not suffer from numerical problems such as overflow or loss of precision that very often occur for large values of k. Using (3) consecutively for $k = 1, 2, \dots$, it is easy to find the value of link capacity k that is necessary for serving the given traffic flow with prescribed QoS. It is easy to see that the computational work for solving the problem equals computational work for estimating $E_k(\lambda)$.

criteria (4). If number of service units k does not satisfy (4), then the value of k is increased and we repeat the described actions once more.

The implementation of this procedure has at least two negative aspects. They are clearly observed for a large number of service units k. For practically cases of interest almost all probability mass is located in states with large number of occupied service units. It means that states with a small number of occupied service units have very small probabilities of existence. The state with number of occupied service units equals to zero has the smallest probability different from zero. So when we start to express all $P(i)$ through $P(0)$ we very quickly experience problems of overflow. Another negative aspect is the increase of computational efforts. To solve the formulated dimensioning problem we need to perform a run of (5) for each value of k serving as a candidate for desired solution. Each time we need to start calculation from $i = 0$.

Let us formulate a condition on implementation of function $f(\cdot)$ where we can solve this problem in one run as was done by recurrence (3) for service systems whose performance measures numerically are estimated by means of Erlang's B-formula. we indicate when necessary the number of available service units by lower index for the corresponding set of probabilities. Let us assume that for any $k > 1$ the implementation of recurrence (5) for number of service units $k-1$ and k and the same starting value gives sets of unnormalised probabilities $P_{k-1}(0) = a$, $P_{k-1}(1)$, $\ldots$, $P_{k-1}(k-1)$, and $P_k(0) = a$, $P_k(1)$, $\ldots$, $P_k(k-1)$, $P_k(k)$ satisfying the following relations:

$$P_k(i) = P_{k-1}(i)\,, \quad i = 0, 1, \ldots, k-1. \tag{7}$$

Property (7) allows us to rewrite (5) in a form dealing with normalised probabilities.

Let us assume that we know the normalised set of probabilities $p_{k-1}(0)$, $p_{k-1}(1)$,$\ldots$, $p_{k-1}(k-1)$ for number of service units $k-1$. Let us assume that the number of service units is increased by one to k. Then taking as starting value $P_k(0) = p_{k-1}(0)$ and applying (5) we obtain according to property (7) the following set of unnormalised probabilities:

$$\begin{aligned} P_k(0) &= p_{k-1}(0)\,, \\ P_k(1) &= p_{k-1}(1)\,, \\ \ldots \quad & \quad \ldots \\ P_k(k-1) &= p_{k-1}(k-1)\,, \\ P_k(k) &= f(p_{k-1}(k-1), p_{k-1}(k-2), \ldots, p_{k-1}(k-m))\,. \end{aligned} \tag{8}$$

For normalised probability $P_k(k)$ we have:

$$\begin{aligned} p_k(k) &= \frac{P_k(k)}{P_k(0) + P_k(1) + \cdots + P_k(k)} \\ &= \frac{P_k(k)}{p_{k-1}(0) + p_{k-1}(1) + \cdots + p_{k-1}(k-1) + P_k(k)} \\ &= \frac{P_k(k)}{1 + P_k(k)}\,. \end{aligned} \tag{9}$$

Using (8) we rewrite (9) in the form:

$$p_k(k) = \frac{f(p_{k-1}(k-1), p_{k-1}(k-2), \ldots, p_{k-1}(k-m))}{1 + f(p_{k-1}(k-1), p_{k-1}(k-2), \ldots, p_{k-1}(k-m))}\,. \tag{10}$$

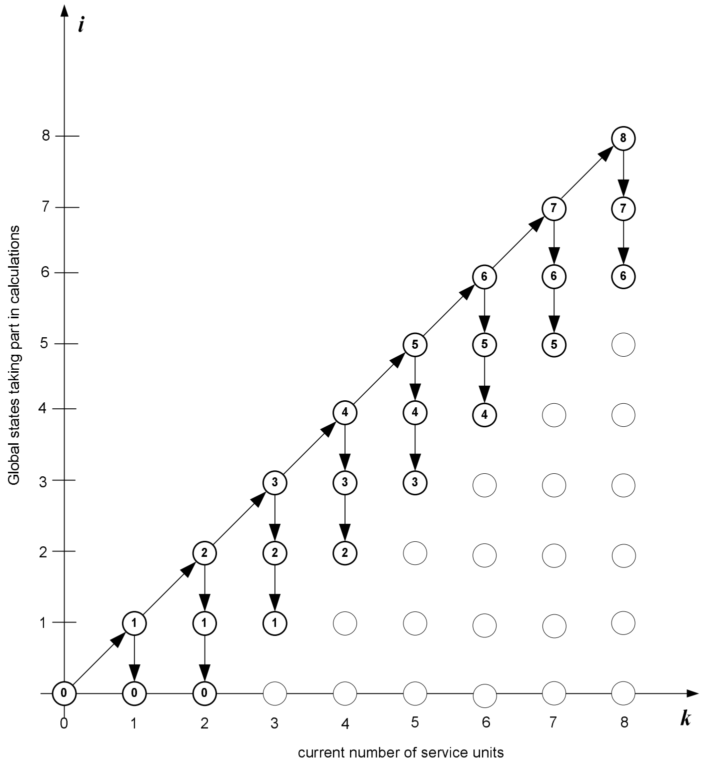

Fig. 1. The order of calculation of the normalised values of stationary probabilities in accordance with suggested approach ($j = 2$)

some intermediate value chosen according to availability of some additional information concerning model functioning. The first modification gives exact results and the second gives approximate values, but with very good convergence to exact results.

Let us consider the first modification of the basic algorithm. It is clear that we can start stable recursions from some intermediate value of $k = k_0$ if in some way (for example by means of (5)) we know normalised values of $p_{k_0}(k_0), p_{k_0}(k_0-1), \ldots, p_{k_0}(k_0-j)$. This follows directly from discussions previous stated algorithm (7)–(12) and Fig. 1. Let us for simplicity assume that $k_0 \geq j$. Modified in such way the algorithm looks as follows
Step 1. Using (5) find normalised values $p_{k_0}(k_0), p_{k_0}(k_0-1), \ldots, p_{k_0}(k_0-j)$.
Step 2. Let $j = \max(b-1, m)$. For fixed $k = k_0+1, k_0+2, \ldots,$ find normalised value of $p_k(i)$, $i = k, k-1, \ldots, k-j$, by using relations (10),(12):

$$p_k(k) = \frac{f(p_{k-1}(k-1), p_{k-1}(k-2), \ldots, p_{k-1}(k-j))}{1 + f(p_{k-1}(k-1), p_{k-1}(k-2), \ldots, p_{k-1}(k-j))}\,; \tag{14}$$

$$p_k(i) = \frac{p_{k-1}(i)}{1 + f(p_{k-1}(k-1), p_{k-1}(k-2), \ldots, p_{k-1}(k-j))}\,, \quad i = k-1,\, k-2,\, \ldots, k-j\,.$$

Step 3. The same as formulated for basic version of algorithm.

The computational efforts are in accordance with modified versions of basic algorithm is estimated by $O\{(N_f + (N_B + j + 1)(1 - \frac{k_0}{k}))\,k\}$. Storage requirements are estimated by $O\{j+1\}$.

4 One-link multi-service models with product form solutions

Let us consider a single link traffic model, where the link transmission capacity is represented by k basic bandwidth units, and let us suppose that we have n incoming Poisson flows of calls with intensities λ_s, $s = 1, \ldots, n$. A call of s'th flow uses b_s bandwidth units for the time of connection. Without loss of generality we shall assume that the holding times all are exponentially distributed with the same mean value chosen to one, but it is known that the model considered is insensitive to the distribution of the holding time, and each flow may furthermore have individual mean holding times.

Let $i_s(t)$ denote the number of calls of the s'th flow served at time t. The model is described by an n-dimensional Markovian process of the type $r(t) = \{i_1(t), i_2(t), \ldots, i_n(t)\}$ with state space S consisting of vectors $(i_1, \ldots, i_n)$, where i_s is the number of calls of the s'th flow being served by the link under stationary conditions. The state space S is defined as follows: $(i_1, \ldots, i_n) \in S,\ i_s \geq 0\,,\ s = 1, \ldots, n\,,\ \sum_{s=1}^{n} i_s b_s \leq k$. Let us by $P(i_1, \ldots, i_n)$ denote the unnormalised values of stationary probabilities of $r(t)$. After normalisation, $p(i_1, \ldots, i_n)$ denotes the mean proportion of time when exactly $\{i_1, \ldots, i_n\}$ connections are established. Assume that for state $(i_1, \ldots, i_n)$ the value i denotes the total number of occupied bandwidth units $i = i_1 b_1 + \cdots + i_n b_n$.

The process of transmission of s'th flow is described by blocking probability π_s, $s = 1, \ldots, n$. Their formal definition through values of state probabilities are as follows (here and further, summations are for all states $(i_1, \ldots, i_n) \in S$ satisfying formulated condition, and by small characters we denote the normalised values of state probabilities):

$$\pi_s = \sum_{i+b_s>k} p(i_1, \ldots, i_n). \tag{17}$$

There are many algorithms for estimation of π_s. All of them are based on product form relations valid for $P(i_1, \ldots, i_n)$:

$$P(i_1, \ldots, i_n) = P(0, \ldots, 0) \cdot \frac{\lambda_1^{i_1}}{i_1!} \cdot \frac{\lambda_2^{i_2}}{i_2!} \cdot \ldots \cdot \frac{\lambda_n^{i_n}}{i_n!}\,, \quad (i_1, \ldots, i_n) \in S.$$

The most efficient calculation scheme for the model introduced is the recurrence algorithm first obtained in [6] and later also derived in [7],[8]. This algorithm exploits the fact that the performance measures (17) can be found if we for process $r(t)$ know probabilities $p(i)$ of being in the state where exactly i bandwidth units are occupied:

$$p(i) = \sum_{i_1 b_1 + \cdots + i_n b_n = i} p(i_1, \ldots, i_n)\,.$$

The corresponding formulas are as follows:

$$\pi_s = \sum_{i=k-b_s+1}^{k} p(i), \quad s = 1, 2, \ldots, n. \tag{18}$$

The unnormalised values of $P(i)$ are found by the recurrence:

$$P(i) = \begin{cases} 0 & \text{if } i < 0 \\ a & \text{if } i = 0 \\ \frac{1}{i} \sum\limits_{s=1}^{n} \lambda_s\, b_s\, P(i - b_s) & \text{if } i = 1, 2, \ldots, k \end{cases} \tag{19}$$

In the same way as for the initial model we can prove that for the auxiliary model a product form solution is valid for stationary probabilities and corresponding recursion scheme for estimation of unnormalised value of $P^*(i)$

$$P^*(i) = \begin{cases} 0 & \text{if } i < \ell \\ a & \text{if } i = \ell \\ \frac{1}{i} \sum\limits_{s=1}^{n} \lambda_s \, b_s \, P^*(i - b_s) & \text{if } i = \ell + 1, \ell + 2, \ldots, k \end{cases} \tag{21}$$

As result in this particular case we prove the possibility to use the modification of recurrence (5) in the form suggested by relation (15) when dealing with deriving of stable recursion with truncation of global state space.

Using general framework formulated in Section 3 we obtain a one–run algorithm that gives at each step the normalised values of the model's characteristics that are necessary for solving the dimensioning problem based on performance measures (18).

Step 1. Let $p^*_\ell(\ell) = 1$.

Step 2. Let $b = \max\limits_{0 \le s \le n}(b_s)$. For fixed $k = \ell + 1, \ell + 2, \ldots$, find normalised value of $p^*_k(i)$, $i = k;\ k-1, \ldots, \max(k - b, \ell)$ by using relations (21)

$$p^*_k(k) = \frac{\frac{1}{k} \sum\limits_{s=1}^{n} \lambda_s \, b_s \, p^*_{k-1}(k-b_s)}{1 + \frac{1}{k} \sum\limits_{s=1}^{n} \lambda_s \, b_s \, p^*_{k-1}(k-b_s)}; \tag{22}$$

$$p^*_k(i) = \frac{p^*_{k-1}(i)}{1 + \frac{1}{k} \sum\limits_{s=1}^{n} \lambda_s \, b_s \, p^*_{k-1}(k - b_s)}, \quad i = k-1,\ k-2, \ldots, \max(k - b, \ell)$$

Step 3. The same as formulated for basic version.

To illustrate the computational savings obtained when using stable recurrence with or without truncations of global state space we consider numerical example. In Table 1 relative efforts for three schemes of calculation are compared to reach value of $k = 100, 500, 1000, 5000, 10000$. Other model parameters are $n = 4$, $b_1 = 1$, $b_2 = 2$, $b_3 = 3$, $b_4 = 5$, $\lambda_s = \frac{k}{nb_s}$, $s = 1, 2, 3, 4$. In Table 1 the results presented in the column "Full search" are based on step-by-step usage of recurrence (19), the results presented in the column "Stable recurrence" are based on usage of stable recurrence in the form defined by (20), the results presented in the column "Stable recurrence & truncation" are based on usage of stable recurrence with truncation in the form defined by (22), the results presented in the column "Truncation level" give the value of ℓ that is chosen to provide relative error of one percent in estimating of weighted blocking P defined for studied model as $P = \sum\limits_{s=1}^{n} \lambda_s b_s \pi_s / \sum\limits_{s=1}^{n} \lambda_s b_s$. For fixed value of k computational efforts are normalised by efforts needed for usage of stable recurrence in the form defined by (20). The results show considerable savings of computational efforts.

5.2 One-link model with batch arrival demands

Let us consider a single link traffic model, where the link transmission capacity is represented by k basic bandwidth units. In the model we have one Poissonian flow of batches

6 Conclusions

In this paper we describe the general framework for constructing numerically stable versions of recursive algorithms designed for estimation of performance measures of different teletraffic models. The results obtained are based on ideas from [1],[2] and generalise for multi-rate models the well known recurrence for the Erlang–B formula for k servers with Poissonian input of intensity λ. The fact that we at each step of recurrence are dealing with normalised values of state probabilities used for estimation of main stationary performance measures is very well suited for solving problems of dimensioning. In addition to being numerically stable, the main positive features of the suggested approach is its simplicity. All we need is just to take recurrence for global state probabilities in the form (5) and put it into relations (13) or it's modifications (14), (16). Another advantage is the possibility to increase the efficiency of recursive algorithm by truncation of the state space used.

References

1. S.N. Stepanov & V.B. Iversen & V.O. Kostrov. Optimised Dimensioning of Large Bandwidth Resources Under MPLS Technology. Proc. of St. Petersburg Regional International Teletraffic Seminar on Telecommunication Network and Teletraffic Theory. St. Petersburg, Russia. January 2002, pp. 50–63.
2. V.B. Iversen. Teletraffic Engineering. Chapter 10: Multi–dimensional Loss Systems. COM Center, Technical University of Denmark. 2005. 336 pp.
3. A. Nilsson & M. Perry. Multi-Rate Blocking Probabilities: numerically stable computations. ITC 15, Fifteenth International Teletraffic Congress, Washington, DC, USA, Elsevier, 1997, pp. 1359-1368.
4. A. Nilsson & M. Perry & A. Gersht & V.B. Iversen. On Multi–Rate Erlang-B Computations. ITC 16, Sixteenth International Teletraffic Congress, Edinburgh, UK, June 1999. Elsevier 1999. pp. 1051–1060.
5. S.A. Berezner & A.E. Krzesinski. An Efficient Stable Recursion to Compute Multiservice Blocking Probabilities. Performance Evaluation. Vol. 43 (2001) : 151—164.
6. R. Fortet & Ch. Grandjean. Congestion in a Loss System when Some Calls want Several Devices Simultaneously. Electrical Communications, Vol. 39 (1964) : 4, 513–526. Paper presented at ITC–4, Fourth International Teletraffic Congress, London, UK, 15–21 July 1964.
7. J.S. Kaufman. Blocking in a Shared Resource Environment. IEEE Transactions on Communications, Vol. COM–29 (1981) : 10, 1474–1481.
8. J.W. Roberts. A Service System with Heterogeneous User Requirements – Applications to Multi–Service Telecommunication Systems. Pages 423–431 in Performance of Data Communication Systems and their Applications. G. Pujolle (editor), North–Holland Publ. Co. 1981.
9. L.E.N. Delbrouck. On the Steady-State Distribution in a Service Facility carrying Mixtures of Traffic with Different Peakedness Factor and Capacity Requirements. IEEE Transactions on Communications, Vol. COM–31 (1983) : 11, 1209–1211.