A Simplified Blocking Probability Calculation in the Retry Loss Models for Finite Sources

Ioannis D. Moscholios, Michael D. Logothetis and George K. Kokkinakis
WCL, Dept. of Electrical & Computer Engineering, University of Patras, 265 00 Patras, Greece
E-mail: {moscholios, m-logo, gkokkin}@wcl.ee.upatras.gr

Abstract—We consider the single and multi-retry loss models for finite sources (f-SRM, f-MRM, respectively) in which calls of K service-classes are accommodated in a link of capacity C and compete for the available bandwidth under the complete sharing (CS) policy. Blocked calls can immediately retry one or more times to be connected in the system with reduced bandwidth and increased service time requirements. In both f-SRM and f-MRM the calculation of the link occupancy distribution can be quite complex since it requires enumeration and processing of the state space. To simplify this calculation and consequently the calculation of blocking probabilities, we present an approximate method based on the corresponding retry loss models of infinite sources. Simulation results verify the method’s accuracy.

Keywords: Loss System; Finite Population; Retrials; Blocking Probability; Recurrent Formula.

I. INTRODUCTION

The wireless multi-service networking environment necessitates the use of teletraffic models with finite source population for the assessment of the call-level QoS [1].

The basic multi-rate loss model for finite sources refers to fixed-rate traffic and we name it Engset Multirate Loss Model (EnMLM), since for a single service-class, it gives the same blocking probability results with the Enset formula for the time congestion probability [2]. According to the EnMLM a link of capacity C accommodates calls of K different service-classes under the CS policy, i.e. calls are accepted in the system if their required bandwidth is available. The offered traffic load of each service-class k (k=1,...,K) comes from a finite number of N_k sources. The model has a product form solution (PFS) and the calculation of the link occupancy distribution G(j), where j is the occupied link bandwidth, is based on an accurate and recursive formula [3].

Extensions of the EnMLM which cope with elastic traffic are: a) the f-SRM [4] and the f-MRM [5], in which blocked calls can immediately retry once or many times to be connected in the system with reduced bandwidth and increased service time requirements and b) the threshold models where calls, prior to being blocked, may adjust their bandwidth and service time requirements according to either a single threshold, or a set of thresholds common to all service-classes or a different set of thresholds for each service-class [5]. An important consideration for multi-service networks is the application of the bandwidth reservation policy in the above mentioned models [1],[6]. In all these extensions the PFS is destroyed and therefore only approximate but recursive G(j)’s formulas are available. Furthermore, in both the EnMLM and its extensions the G(j)’s determination is more complex compared to that in the corresponding infinite source models ([7]-[10]), since the system’s state space needs enumeration and processing prior to the G(j)’s calculation.

In the case of the EnMLM one can overcome the state space enumeration and processing by an approximate method proposed, in [11], by Glabowski and Stasiak (G&S method) which is based on the corresponding infinite source model, the classical Erlang Multirate Loss Model (EMLM) [12]. According to [11], the number of in-service sources, n_k, of service-class k in a state j of an EnMLM system is approximated by the average number, y_k(j), of service-class k calls in the same state j of the corresponding EMLM system. The values of y_k(j) are easily computed via the well-known Kaufman-Roberts formula ([7],[8]) and then replace the corresponding n_k’s in the G(j)’s calculation of the EnMLM.

In this paper we study the applicability of the G&S method in the case of the f-SRM and f-MRM. Its application requires the knowledge of the average number of service-class k calls in each state j from the corresponding single and multi-retry models (SRM and MRM, respectively) for infinite number of sources [9]. The analytical blocking probability results obtained by the G&S method are in accordance with simulation results.

The remainder of the paper is as follows: In Section II we review the f-SRM and the f-MRM. In Section II.A we present the analytical model of the f-SRM while in Section II.B the analytical model of the f-MRM. In Section II.C we present, through a simple example, the method for the state space determination in the case of the f-SRM. In Section III we proceed with the application of the G&S method in both models. Numerical results are presented in Section IV. In Section V we conclude.

II. THE RETRY MODELS FOR INFINITE SOURCES

A. The f-SRM - the analytical model

We assume a link of capacity C bandwidth units (b.u.) that accommodates calls of K service-classes which compete for the available bandwidth under the CS policy. Each service-class k has a finite source population N_k and requires b_k b.u per call. Service-class k calls arrive to the link according to a quasi-random process [2], with mean arrival rate \( \lambda_k = (N_k - n_k) \mu_k \), where n_k is the number of in-service sources and \( \mu_k \) is the mean arrival rate per idle source. The offered traffic load per idle service-class k source is \( \alpha_k = v_k / \mu_k \) where \( \mu_k^{-1} \) is the mean service time, exponentially distributed. Blocked service-class k calls retry to be connected in the system with parameters \( (b_k, \mu_k^{-1}) \) where \( b_k < b_l \) and \( \mu_k^{-1} > \mu_l^{-1} \). The f-SRM does not have a PFS and therefore the G(j)’s calculation is based on an approximate recursive formula [4]:
where: \( n_{bk} \) is the in-service retry calls of service-class \( k \), 
\( \alpha_{bk} = \gamma_{bk} \mu_{bk} \) (offered traffic load per idle source of retry service-class \( k \) calls), \( \gamma_{bk} = 1 \) when \( j > C(b_{k-1}-b_{k}) \).

The proof of eq.(1) is based on two assumptions [4]: 1) the existence of local balance, which appears only in PFS models and 2) the application of the migration approximation (M.A.) which assumes that the population of retry service-class \( k \) calls are negligible when \( j \leq C(b_{k}-b_{k}) \). The existence of M.A. in eq.(1) is expressed by the variable \( \gamma_{bk} \). Call blocking probabilities (CBP) of service-class \( k \), \( B_{ak} \), is the probability of a call to be blocked with its retry bandwidth and is given by:

\[
B_{ak} = \sum_{j=C-b_{k+1}}^{C} G^{-1}G(j), \quad \text{where} \quad G = \sum_{j=0}^{C} G(j) \tag{2}
\]

**Note:** By the term CBP we mean time congestion probabilities. For quasi-random arrivals, there is a distinction between time and call congestion probabilities. These probabilities coincide in the case of Poisson arrivals (PASTA property [2]). Time congestion probability is determined by the proportion of time the system is congested and can be measured by an outside observer.

When \( N_{k} \) approaches infinity for \( k = 1,...,K \) then the SRM results. In that case the \( G(j) \)'s calculation is based on the following approximate but recursive formula [9]:

\[
G(j) = \begin{cases} 
1 & \text{for } j = 0 \\
\frac{1}{j} \sum_{k=1}^{K} \alpha_{bk} b_{k} G(j - b_{k}) \\
\frac{1}{j} \sum_{k=1}^{K} \alpha_{bk} b_{k} \gamma_{bk} G(j - b_{k}) & \text{for } j = 1,..., C \\
0 & \text{otherwise}
\end{cases} \tag{3}
\]

where: \( \alpha_{bk} = \lambda_{bk} \mu_{bk} \) and \( \gamma_{bk} = 1 \) when \( j > C(b_{k}-b_{k}) \).

The proof of eq. (3) is based on the assumptions of eq. (1), while the CBP of service-class \( k \), \( B_{ak} \), is determined by eq. (2). Based on eq. (3), one can calculate, in state \( j \), the average number of service-class \( k \) calls either with \( b_{k} \) b.u., \( y_{k}(j) \), or with \( b_{k} \) b.u., \( y_{bk}(j) \):

\[
y_{k}(j) = a_{k} G(j - b_{k})/G(j) \tag{4}
\]

where: \( j-b_{k} > 0 \)

\[
y_{bk}(j) = a_{bk} \gamma_{bk}(j) G(j - b_{k})/G(j) \tag{5}
\]

where: \( j-b_{k} > 0 \) and \( \gamma_{bk}(j) = 1 \) when \( j > C(b_{k}-b_{k}) \).

**B. The f-MRM - the analytical model**

In the f-MRM, a blocked service-class \( k \) call may retry many times with “retry parameters” \((b_{kj}, \mu_{kj})\) for \( l=1,...,s(k) \), where \( b_{kj} \leq ... \leq b_{kj} < b_{k} \) and \( \mu_{kj}^{-1} \), \( j \geq s(k) \), \( \mu_{s(k)}^{-1} >... > \mu_{1}^{-1} \).

Similar to the f-SRM, the f-MRM does not have a PFS and therefore the \( G(j) \)'s calculation is based on the following approximate recursive formula [5]:

\[
G(j) = \begin{cases} 
1 & \text{for } j = 0 \\
\frac{1}{j} \sum_{k=1}^{K} \alpha_{bk} b_{k} G(j - b_{k}) \\
\frac{1}{j} \sum_{k=1}^{K} \alpha_{bk} b_{k} \gamma_{bk} G(j - b_{k}) & \text{for } j = 1,..., C \\
0 & \text{otherwise}
\end{cases} \tag{6}
\]

where: \( n_{bk} \) is the in-service retry service-class \( k \) calls with \( b_{kj} = \gamma_{bk} \mu_{bk} \) and \( \gamma_{bk}(j) = 1 \) when \( C \geq j > C(b_{kj} - b_{kj}) \).

The CBP of a service-class \( k \), denoted as \( B_{ak} \), is the probability of a call to be blocked with its last bandwidth requirement and is calculated by:

\[
B_{ak} = \sum_{j=C-b_{k+1}}^{C} G^{-1}(j), \quad \text{where} \quad G = \sum_{j=0}^{C} G(j) \tag{7}
\]

When \( N_{k} \) approaches infinity for \( k = 1,...,K \) then the MRM results and the \( G(j) \)'s calculation is based on [9]:

\[
G(j) = \begin{cases} 
1 & \text{for } j = 0 \\
\frac{1}{j} \sum_{k=1}^{K} a_{bk} b_{k} G(j - b_{k}) \\
\frac{1}{j} \sum_{k=1}^{K} a_{bk} b_{k} \gamma_{bk}(j) G(j - b_{k}) & \text{for } j = 1,..., C \\
0 & \text{otherwise}
\end{cases} \tag{8}
\]

where: \( a_{bk} = \lambda_{bk} \mu_{bk} \) and \( \gamma_{bk}(j) = 1 \) when \( C \geq j > C(b_{kj} - b_{kj}) \), otherwise \( \gamma_{bk}(j) = 0 \).

Having determined the values of \( G(j) \)'s from eq. (8), then CBP of service-class \( k \) calls, \( B_{ak} \), are given by eq. (7), the average number of service-class \( k \) calls, in state \( j \), with \( b_{kj} \) b.u., \( y_{kj}(j) \), by eq. (4) and the average number of service-class \( k \) calls with \( b_{kj} \) b.u., \( y_{bk}(j) \) by the equation:

\[
y_{jk}(j) = a_{jk} \gamma_{jk}(j) G(j - b_{kj})/G(j) \tag{9}
\]

where: \( j-b_{kj} > 0 \) and \( \gamma_{jk}(j) = 1 \) when \( j > C(b_{kj} - b_{kj}) \), otherwise \( \gamma_{jk}(j) = 0 \).

**C. State space determination and \( G(j) \)'s calculation in the f-SRM**

Consider a link of capacity \( C = 5 \) b.u. and two service-classes whose calls require \( b_{1}=3 \) and \( b_{2}=2 \) b.u., respectively. The offered traffic load per idle source is \( \alpha_{1} = \alpha_{2} = 0.01 \) erl, while the number of sources is \( N_{1}=N_{2}=6 \). Blocked 1st service-class calls retry with reduced bandwidth, \( b_{1}=1 \) b.u. and increased offered traffic \( \alpha_{1r}=0.03 \) erl, so that the total offered traffic load per idle source remains the same (\( \alpha_{1r}=\alpha_{2r}=6 \)). Due to the M.A. retry calls of the 1st service-class are assumed to be negligible when the occupied link bandwidth \( j \leq C(b_{1}-b_{1r}) \) \( \Rightarrow j \leq 3 \). Taking into account the M.A. assumption, the state space consists of 8 states (\( n_{1}, n_{2}, n_{ir} \)), presented in
Table I together with the respective occupied link bandwidth \( j \) and the blocking states \((B_1, B_2, B_{kr})\), where \( B_k \) (\( k=1,2 \)) expresses the CBP of service-class \( k \) calls with \( b_k \) b.u. calculated by eq. (2) if one replaces \( b_k \) with \( b_1 \).

According to Table I the values of \( j = 4 \), 5 appear more than once, and therefore it is impossible to use directly eq. (1) for the calculation of \( G(j)'s \) (e.g. the value \( j = 4 \), corresponds to both \((n_1, n_2, n_r) = (0,2,0) \) and \((1,0,1)\)). To overcome this problem, the authors proposed a method whereby an equivalent stochastic system is determined with the following three characteristics [5]:

a) The states \((n_1, n_2, n_r)\) of the equivalent system are the same with the initial one.

b) Each state \((n_1, n_2, n_r)\) of the equivalent system has a unique value for the occupied link bandwidth.

c) The chosen values of \( b_1 \), \( b_2 \) and \( b_{kr} \) of the equivalent system keep constant (as much as possible) the initial ratios of \( b_1:b_2:b_{kr}:C = 3:2:1:5 \).

In our example, the values \( b_1 = 3000 \), \( b_2 = 2000 \), \( b_{kr} = 1001 \) and \( C = 5002 \), is an approximate solution to the initial system; for these values we present in the last column of Table I the unique values of the equivalent occupied link bandwidth, \( f_{eq} \). The resultant CBP are: \( B_1 = 5.98\% \), \( B_2 = 0.66\% \) and \( B_{kr} = 0.33\% \). Note that every system that is a multiple of \( b_1 = 3000 \), \( b_2 = 2000 \), \( b_{kr} = 1001 \) and \( C = 5002 \) (e.g. \( b_1 = 6000 \), \( b_2 = 4000 \), \( b_{kr} = 2002 \) and \( C = 10004 \)) gives exactly the same CBP.

The determination of an equivalent system, and thus CBP, can become difficult when calls retry many times. To circumvent this problem we examine, in the next sections, the applicability of the G&S method in the f-SRM and f-MRM.

### TABLE I

<table>
<thead>
<tr>
<th>( n_1 )</th>
<th>( n_2 )</th>
<th>( n_{kr} )</th>
<th>( j )</th>
<th>( B_1 )</th>
<th>( B_2 )</th>
<th>( B_{kr} )</th>
<th>( f_{eq} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>2000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>2</td>
<td>0</td>
<td>4</td>
<td>*</td>
<td>4000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>2</td>
<td>1</td>
<td>5</td>
<td>*</td>
<td>*</td>
<td>5001</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>*</td>
<td>3000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>1</td>
<td>4</td>
<td>*</td>
<td>*</td>
<td>4001</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>0</td>
<td>5</td>
<td>*</td>
<td>*</td>
<td>5002</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>5</td>
<td>*</td>
<td>*</td>
<td>5000</td>
<td></td>
</tr>
</tbody>
</table>

### III. APPLICATION OF THE G&S METHOD IN THE RETRY MODELS FOR FINITE SOURCES

According to the G&S method, initially proposed in the case of the EnMLM [11], one approximates the number of in-service sources, \( n_k \), of service-class \( k \) in a state \( j \) of an EnMLM system with the average number, \( y_i(j) \), of service-class \( k \) calls with \( b_k \) b.u. in the same state \( j \) of an EMLM system.

Similar to the above, we consider again the simple f-SRM example of the previous section and approximate, for each state \( j = 1, \ldots, 5 \), the number of in-service sources \( n_1 \), \( n_2 \), \( n_{kr} \) with the average number \( y_i(j) \), \( y_s(j) \), \( y_r(j) \), respectively, of the corresponding SRM system. Note that in the SRM system: \( \alpha_1 = \alpha_2 = 0.06 \text{ erl} \) and \( \alpha_r = 0.18 \text{ erl} \).

The values of \( y_i(j) \), \( y_s(j) \), \( y_r(j) \) are given by eq. (4) while those of \( y_{i,r}(j) \) by eq. (5). The results are summarized in Table II.

### TABLE II

<table>
<thead>
<tr>
<th>( y_i(j) )</th>
<th>( y_s(j) )</th>
<th>( y_r(j) )</th>
<th>( j )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>8</td>
<td>4</td>
</tr>
<tr>
<td>0.9569</td>
<td>0.9569</td>
<td>0.2153</td>
<td>5</td>
</tr>
</tbody>
</table>

Having calculated \( y(j)'s \) we continue by determining the link occupancy distribution \( G(j) \) of the f-SRM example, based on the following equation:

\[
G(j)=\begin{cases} 
1 & \text{for } j=0 \\
\sum_{k=1}^{K} (N_k - y_k(j-b_h))\alpha_{tk} b_h G(j-b_h) + \sum_{k=1}^{K} (N_k - y_k(j-b_{kr}))\alpha_{tk} b_{kr} G(j-b_{kr}) & \text{for } j=1, \ldots, C \\
0 & \text{otherwise}
\end{cases}
\]

where: \( \alpha_{tk} = v_{tk} H_{kr}^{-1} \gamma_{tk}(j) = 1 \text{ when } j \geq C-(b_{kr}-b_h) \) and \( \gamma_{tk}(j) = 1 \text{ when } j \geq C-(b_{kr}-b_h) \).

### IV. NUMERICAL EXAMPLE

In this section we compare the analytical CBP results obtained by the application of the G&S method in an f-MRM example with simulation results, which are mean values of 7 runs with 95% confidence interval.

Consider a link of capacity \( C=50 \text{ b.u.} \) and two service-classes which require \( b_1 = 10 \) and \( b_2 = 7 \text{ b.u.} \) respectively. The offered traffic load (per idle source) of each service-class is \( \alpha_1 = 0.06 \text{ erl} \) and \( \alpha_2 = 0.2 \text{ erl} \) respectively. Blocked calls of the first service-class reduce their bandwidth requirement two times, from 10 to 8 and to 6 b.u. In the first case the offered traffic load is \( \alpha_{1,r} = \alpha_1 b_1/b_{1,r} = 0.075 \text{ erl} \) while in the second \( \alpha_{1,r} = \alpha_1 b_1/b_{1,r} = 0.1 \text{ erl} \). Blocked calls of the second service-class reduce their bandwidth requirement once, from 7 to 4 b.u. The retry offered traffic load of
second service-class calls is $\alpha_1=\alpha_2 b_2 / b_2=0.35$ erl. The number of sources for both service-classes is $N_1=N_2=12$.

The equivalent stochastic system used for the CBP calculation is the following [5]: $C=50006$, $b_1=10000$, $b_2=7001$, $b_1=6000$, $b_2=4000$. Table III shows the state space which consists of 30 states.

The corresponding MRM system used for the $Y(\cdot)$’s and CBP calculation, according to the G&S method, is: $C=50$ b.u., $b_1=10$ b.u., $\alpha_1=0.72$ erl, $b_1=8$ b.u., $\alpha_1=0.9$ erl, $b_2=6$ b.u., $\alpha_1=1.2$ erl and $b_2=7$ b.u., $\alpha_2=2.4$ erl, $b_2=4$ b.u., $\alpha_2=4.2$ erl.

In Table IV, we present the analytical CBP results obtained from the equivalent stochastic system and the G&S method together with the corresponding simulation results. At each point (P) of Table IV (P=1,…,6), the values of $\alpha_1$, $\alpha_1$, $\alpha_1$, $\alpha_1$, $\alpha_2$ are constant (i.e. $\alpha_1=0.06$, $\alpha_1=0.075$ and $\alpha_1=0.1$ erl) while those of $\alpha_2$, $\alpha_2$, are increased by 0.4/12 and 0.7/12 respectively, (for P=1: $\alpha_1=0.2$ and $\alpha_2=0.35$ erl).

According to the results of Table IV, one observes that the G&S method gives almost the same results with the equivalent stochastic system’s method. This behaviour has been observed in many examples. Further study in the case of more complicated examples or in the case of the threshold models for finite sources is needed in order to verify the superiority of the G&S method.

V. CONCLUSION

We review the f-SRM and the f-MRM loss models where blocked calls try to be connected in the system with reduced bandwidth and increased service-time requirements. The calculation of the link occupancy distribution in these models and therefore of CBP is based on the enumeration and processing of the state space which can be complex in systems where the number of retrials is large. To simplify the CBP calculation we examine the applicability of the G&S method, already proposed in the case of the EnMLM, in both models. Results show that the G&S performs quite well.

REFERENCES


TABLE III

<table>
<thead>
<tr>
<th>$n_1$</th>
<th>$n_2$</th>
<th>$n_1$</th>
<th>$n_2$</th>
<th>$n_1$</th>
<th>$n_2$</th>
<th>$n_1$</th>
<th>$n_2$</th>
<th>$n_1$</th>
<th>$n_2$</th>
<th>$n_1$</th>
<th>$n_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

TABLE IV

<table>
<thead>
<tr>
<th>P</th>
<th>Equivalent Stochastic System</th>
<th>G&amp;S method</th>
<th>Simulation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$B_{1}(%)$</td>
<td>$B_{2}(%)$</td>
<td>$B_{1}(%)$</td>
</tr>
<tr>
<td>1</td>
<td>3.05</td>
<td>2.03</td>
<td>3.05</td>
</tr>
<tr>
<td>2</td>
<td>4.32</td>
<td>2.72</td>
<td>4.32</td>
</tr>
<tr>
<td>3</td>
<td>5.78</td>
<td>3.50</td>
<td>5.78</td>
</tr>
<tr>
<td>4</td>
<td>7.39</td>
<td>4.36</td>
<td>7.39</td>
</tr>
<tr>
<td>5</td>
<td>9.12</td>
<td>5.30</td>
<td>9.13</td>
</tr>
<tr>
<td>6</td>
<td>10.94</td>
<td>6.31</td>
<td>10.94</td>
</tr>
</tbody>
</table>