Multirate Retry Loss Models Supporting Elastic Traffic of Quasi-Random Input

I. D. Moscholios*, V. G. Vassilakis**, J. S. Vardakas***, and M. D. Logothetis****

*Dept. of Informatics & Telecommunications, University of Peloponnese, 221 00 Tripolis, Greece
** Dept. of Electronic Engineering, University of Surrey, GU2 7XH Guildford, U.K.
*** Iquadrat, 080 09 Barcelona, Spain
****WCL, Dept. of Electrical & Computer Engineering, University of Patras, 265 04 Patras, Greece.
E-mails: idm@uop.gr, v.vasilakis@surrey.ac.uk, jvardakas@iquadrat.com, mlogo@upatras.gr

***Abstract*–We propose a teletraffic loss model of a single link that accommodates multirate service-classes of elastic calls. New calls arrive in the link according to a quasi-random process (calls are generated by a finite number of traffic sources), have elastic bandwidth requirements and an exponentially distributed service time. A new call is accepted in the system with its peak-bandwidth requirement if there is available link bandwidth. If not, the call retries one or more times (single and multi-retry loss model, respectively) to be accepted in the link with reduced bandwidth. If the available link bandwidth is lower than the call's last bandwidth requirement, the call can still try to be connected in the link by compressing its last bandwidth requirement (down to a certain bandwidth) together with the bandwidth of all in-service calls. If, after compression, the call's bandwidth exceeds the available link bandwidth, then the call is blocked and lost. The analysis of the proposed models is based on approximate but recursive formulas, whereby we determine time and call congestion probabilities and link utilization. The accuracy of the proposed models is verified by simulation and is found to be quite satisfactory.**

I. INTRODUCTION

Elastic traffic of multirate service-classes grows rapidly in modern wired or wireless networks, a fact that necessitates the development of analytical tools for the call-level performance analysis of such networks. The term "elastic traffic" refers to in-service calls which have the ability to compress/expand their bandwidth and simultaneously increase/decrease their service time, during their lifetime in a system. Assuming that the call arrival process is Poisson then the calculation of various performance measures such as blocking probabilities and system's utilization can be based on the Erlang Multirate Loss Model (EMLM) [1]-[2] which has been extensively used for the call-level analysis of wired (e.g., [3]-[8]), wireless (e.g., [9]-[13]) and optical networks (e.g., [14]-[16]). If the call arrival process is quasi-random, i.e., calls come from a finite number of users then the Engset Multirate Loss Model (EnMLM) arises [17].

In both the EMLM and the EnMLM, calls compete for the available link bandwidth according to the complete sharing policy (i.e., calls compete for all bandwidth resources) and have fixed bandwidth requirements. The latter means that in-service calls do not compress their bandwidth during their lifetime in the system. A new call is blocked and lost if its required bandwidth is not available. In both models the steady state probabilities have a Product Form Solution (PFS) which leads to an accurate calculation of call blocking probabilities (see e.g., [1], [2] and [17]). In [18], [19] the EMLM and the EnMLM, respectively, have been extended to include retrials. Blocked calls retry one or more times (Single-Retry Model (SRM) and Multi-Retry Model (MRM), respectively) to be accepted in the link by requiring less bandwidth. A retry call is blocked and lost if the available link bandwidth is lower than the call's last bandwidth requirement. Note that the case of single retry in the EnMLM has also been considered in [20]. In [21], an approximate method has been proposed for both single and multi retries in the EnMLM that simplifies the calculation of blocking probabilities. In [22], the authors have extended [18] by incorporating the notion of elastic traffic. Instead of rejecting immediately a retry call, the link may accept this call by compressing its bandwidth and the bandwidth of all in-service calls. Elastic calls increase their service time so that the product *bandwidth* by *service time* remains constant. When a call with compressed bandwidth leaves the system, then the remaining in-service calls expand their bandwidth. A retry call is blocked and lost, if its compressed bandwidth should be less than a minimum proportion of its last bandwidth requirement. In [23], [24], the authors have extended [22] to include adaptive traffic and the bandwidth reservation policy (in this policy a fraction of bandwidth is reserved to benefit calls of certain service-classes), respectively. Adaptive calls compress or expand their bandwidth without altering their service time.

In this paper, we extend [19], [21] to include elastic traffic. Due to the existence of retrials and bandwidth compression, the proposed elastic single-retry and multi-retry loss models for quasi-random input do not have a PFS. However, we propose approximate but recursive formulas for the calculation of the link occupancy distribution and consequently time and call congestion probabilities as well as link utilization. Applications of the proposed models are in the area of wireless networks where calls may come from finite sources (the limited coverage of a cell justifies the case of a finite number of users) and their bandwidth can be compressed [25]-[27].

This paper is organized as follows. In Section II, we propose the elastic SRM for quasi-random input and prove formulas for the calculation of the various performance measures. In Section III, we propose the corresponding elastic MRM for quasi-random input. In Section IV, we provide numerical results whereby the proposed models are compared to existing models and evaluated via simulation. We conclude in Section V.

II. THE ELASTIC SINGLE RETRY LOSS MODEL FOR QUASI-RANDOM INPUT

Consider a link of capacity C bandwidth units (b.u.) that accommodates elastic calls of K service-classes. Calls of service-class k (k=1,…,K) are generated by a finite source population N_k and have a peak-bandwidth requirement of b_k b.u. If this peak-bandwidth is available, a service-class k call remains in the system for an exponentially distributed service-time with mean μ_k^{-1}. Otherwise, the call is blocked and retries to be connected in the system with "retry parameters" (b_{kr}, μ_{kr}^{-1}) where $b_{kr} < b_k$ and $\mu_{kr}^{-1} > \mu_k^{-1}$. The mean call arrival rate of service-class k idle sources is $\lambda_k = (N_k - n_k)v_k$ where v_k is the arrival rate per idle source and n_k is the number of in-service calls. This call arrival process is a quasi-random process [28]. A Poisson process arises from a quasi-random process if $N_k \to \infty$ for k=1,…,K and the total offered traffic-load remains constant. Bandwidth compression is introduced in the model by assuming that the occupied link bandwidth j may exceed C up to a value of T b.u.

To prove a recursive formula for the calculation of the link occupancy distribution, $G(j)$, we consider an example of a link that accommodates two service-classes, with the following traffic parameters: $(N_1, v_1, \mu_1^{-1}, b_1)$ for the 1st service-class and $(N_2, v_2, \mu_2^{-1}, \mu_{2r}^{-1}, b_2, b_{2r})$ for the 2nd service-class. Only calls of the 2nd service-class have "retry parameters" with $b_{2r} < b_2$ and $\mu_{2r}^{-1} > \mu_2^{-1}$.

The description of call admission is based on a new service-class k call (k=1, 2) that arrives in the system when the occupied link bandwidth is j b.u. Then:

i) If $j + b_k \leq C$, the call is accepted in the system with b_k b.u. for an exponentially distributed service time with mean μ_k^{-1}.

ii) If $j + b_k > C$ we consider the following subcases:

a) If $T \geq j + b_1 > C$, a 1st service-class call is accepted in the system by compressing b_1, as well as the assigned bandwidth of all in-service calls. The compressed bandwidth of the 1st service-class call is given by $b_1^{'} = rb_1 = (C/j^{'})b_1$ where $r = C/j^{'}$, $j^{'} = j + b_1 = \boldsymbol{nb} + b_1$. Similarly, the bandwidth of all in-service calls will be compressed (by the same factor r) and become $b_k^{'} = (C/j^{'})b_k$ for k=1, 2. After compression has taken place, all calls share the C b.u. in proportion to their bandwidth requirement, while the link operates at its full capacity C. The minimum bandwidth that a 1st service-class call can tolerate is $b_{1,\min}^{'} = r_{\min} b_1 = (C/T)b_1$.

b) If $j+b_1 > T$, the 1st service-class call is blocked and lost.

c) If $j + b_2 > C$, a 2nd service-class call is blocked and retries with $b_{2r} < b_2$. Now, we consider three cases: 1) If $j + b_{2r} \leq C$ the retry call is accepted in the system with b_{2r}. 2) If $j + b_{2r} > T$ the call is blocked and lost. 3) If $C < j + b_{2r} \leq T$ the call is accepted in the system by compressing b_{2r} together with the bandwidth of all in-service calls. The compressed bandwidth of the call is $b_{2r}^{'} = rb_{2r} = (C/j^{'})b_{2r}$ where $j^{'} = j + b_{2r}$. Similarly, the bandwidth of all in-service calls are compressed (by the same factor r) and become $b_k^{'} = (C/j^{'})b_k$ for k=1, 2. The minimum bandwidth that a 2nd service-class call tolerates is $b_{2r,\min}^{'} = (C/T)b_{2r}$.

Although the steady state probabilities in the proposed model do not have a PFS, we assume that local balance exists between the adjacent states of the 1st service-class:

$$(N_1 - n_1 + 1)v_1 P(\boldsymbol{n}_1^-) = n_1 \mu_1 \phi_1(\boldsymbol{n}) P(\boldsymbol{n}) \quad 1 \leq \boldsymbol{nb} \leq T \tag{1}$$

where: $\boldsymbol{n}_1^- = (n_1 - 1, n_2, n_{2r}), \boldsymbol{n} = (n_1, n_2, n_{2r}), \boldsymbol{b} = (b_1, b_2, b_{2r})$, $n_1 \geq 1$, $P(\boldsymbol{n})$ is the probability distribution of state $\boldsymbol{n}$ and

$$\phi_1(\boldsymbol{n}) = \begin{cases} 1 & , \text{ when } \boldsymbol{nb} \leq C \\ x(\boldsymbol{n}_1^-)/x(\boldsymbol{n}), & \text{when } C < \boldsymbol{nb} \leq T \\ 0 & , \text{ otherwise} \end{cases} \tag{2}$$

$$x(\boldsymbol{n}) = \begin{cases} 1 & , \text{ when } \boldsymbol{nb} \leq C \\ \frac{1}{C}\left(\sum_{k=1}^{2} n_k b_k x(\boldsymbol{n}_k^-) + n_{2r} b_{2r} x(\boldsymbol{n}_{2r}^-) \right) & , \text{ when } C < \boldsymbol{nb} \leq T \\ 0 & , \text{ otherwise} \end{cases} \tag{3}$$

Note that $\varphi_k(\boldsymbol{n})$ is a state dependent factor which describes: i) bandwidth compression and ii) the increase factor of service time of service-class k calls in state $\boldsymbol{n}$. In other words, $\varphi_k(\boldsymbol{n})$ has the same role with r but it may be different for each service-class.

By multiplying both sides of (1) with b_1, and based on (2) we have:

$$(N_1 - n_1 + 1)a_1 b_1 x(\boldsymbol{n}) P(\boldsymbol{n}_1^-) = n_1 b_1 x(\boldsymbol{n}_1^-) P(\boldsymbol{n}) \quad 1 \leq \boldsymbol{nb} \leq T \tag{4}$$

where $a_1 = v_1 \mu_1^{-1}$ is the offered traffic-load per idle source of 1st service-class.

Based on the call admission mechanism described for 2nd service-class calls the following local balance equations can be derived:

$$\text{a) } (N_2 - n_2 + 1)v_2 P(\boldsymbol{n}_2^-) = n_2 \mu_2 \phi_2(\boldsymbol{n}) P(\boldsymbol{n}) \quad 1 \leq \boldsymbol{nb} \leq C \tag{5}$$

where: $\boldsymbol{n}_2^- = (n_1, n_2 - 1, n_{2r})$, $n_2 \geq 1$ and

$$\phi_2(\boldsymbol{n}) = \begin{cases} 1 & , \text{ when } \boldsymbol{nb} \leq C \\ x(\boldsymbol{n}_2^-)/x(\boldsymbol{n}), & \text{when } C < \boldsymbol{nb} \leq T \\ 0 & , \text{ otherwise} \end{cases} \tag{6}$$

By multiplying both sides of (5) with b_2, and based on (6) we have:

$$(N_2 - n_2 + 1)a_2 b_2 x(\boldsymbol{n}) P(\boldsymbol{n}_2^-) = n_2 b_2 x(\boldsymbol{n}_2^-) P(\boldsymbol{n}) \quad 1 \leq \boldsymbol{nb} \leq C \tag{7}$$

where: $a_2 = v_2 \mu_2^{-1}$ and the values of $x(\boldsymbol{n})$ are given by (3).

$$\text{b) } \begin{array}{l} (N_2 - n_2 - n_{2r} + 1)v_2 P(\boldsymbol{n}_{2r}^-) = n_{2r} \mu_{2r} \phi_{2r}(\boldsymbol{n}) P(\boldsymbol{n}) \\ C - b_2 + b_{2r} < \boldsymbol{nb} \leq T \end{array} \tag{8}$$

where: $P(\boldsymbol{n}_{2r}^-)$ is the probability distribution of state $\boldsymbol{n}_{2r}^- = (n_1, n_2, n_{2r} - 1)$ and

$$\phi_{2r}(\boldsymbol{n}) = \begin{cases} 1 & , \text{ when } \boldsymbol{nb} \leq C \\ x(\boldsymbol{n}_{2r}^-)/x(\boldsymbol{n}), & \text{when } C < \boldsymbol{nb} \leq T \\ 0 & , \text{ otherwise} \end{cases} \tag{9}$$

By multiplying both sides of (8) with b_{2r}, and based on (9) we have:

$$(N_2-n_2-n_{2r}+1)a_{2r}b_{2r}x(\boldsymbol{n})P(\boldsymbol{n}_{2r}^-)=n_{2r}b_{2r}x(\boldsymbol{n}_{2r}^-)P(\boldsymbol{n}) \quad C-b_2+b_{2r}<\boldsymbol{nb}\le T \tag{10}$$

Equations (4), (7) and (10) lead to a system of equations:

$$\begin{aligned}&(N_1-n_1+1)a_1b_1x(\boldsymbol{n})P(\boldsymbol{n}_1^-)+(N_2-n_2+1)a_2b_2x(\boldsymbol{n})P(\boldsymbol{n}_2^-)\\&=(n_1b_1x(\boldsymbol{n}_1^-)+n_2b_2x(\boldsymbol{n}_2^-))P(\boldsymbol{n}),\ 1\le\boldsymbol{nb}\le C-b_2+b_{2r}\end{aligned} \tag{11}$$

$$\begin{aligned}&(N_1-n_1+1)a_1b_1x(\boldsymbol{n})P(\boldsymbol{n}_1^-)+(N_2-n_2+1)a_2b_2x(\boldsymbol{n})P(\boldsymbol{n}_2^-)\\&+(N_2-n_2-n_{2r}+1)a_{2r}b_{2r}x(\boldsymbol{n})P(\boldsymbol{n}_{2r}^-)\\&=(n_1b_1x(\boldsymbol{n}_1^-)+n_2b_2x(\boldsymbol{n}_2^-)+n_{2r}b_{2r}x(\boldsymbol{n}_{2r}^-))P(\boldsymbol{n}),\\&C-b_2+b_{2r}<\boldsymbol{nb}\le C\end{aligned} \tag{12}$$

$$\begin{aligned}&(N_1-n_1+1)a_1b_1x(\boldsymbol{n})P(\boldsymbol{n}_1^-)+\\&(N_2-n_2-n_{2r}+1)a_{2r}b_{2r}x(\boldsymbol{n})P(\boldsymbol{n}_{2r}^-)\\&=(n_1b_1x(\boldsymbol{n}_1^-)+n_{2r}b_{2r}x(\boldsymbol{n}_{2r}^-))P(\boldsymbol{n}),\ C<\boldsymbol{nb}\le T\end{aligned} \tag{13}$$

By assuming that retry calls with b_{2r} are negligible when $1\le\boldsymbol{nb}\le C-b_2+b_{2r}$ and that the population of calls with b_2 is negligible when $C<\boldsymbol{nb}\le T$, we can combine (11), (12) and (13) into the following equation:

$$\begin{aligned}&(N_1-n_1+1)a_1b_1x(\boldsymbol{n})P(\boldsymbol{n}_1^-)+(N_2-n_2+1)\gamma_2(\boldsymbol{nb})a_2b_2x(\boldsymbol{n})P(\boldsymbol{n}_2^-)\\&+(N_2-n_2-n_{2r}+1)\gamma_{2r}(\boldsymbol{nb})a_{2r}b_{2r}x(\boldsymbol{n})P(\boldsymbol{n}_{2r}^-)\\&=(n_1b_1x(\boldsymbol{n}_1^-)+n_2b_2x(\boldsymbol{n}_2^-)+n_{2r}b_{2r}x(\boldsymbol{n}_{2r}^-))P(\boldsymbol{n}),\\&1\le\boldsymbol{nb}\le T\end{aligned} \tag{14}$$

where: $\gamma_2(\boldsymbol{nb})=1$ for $1\le\boldsymbol{nb}\le C$, otherwise $\gamma_2(\boldsymbol{nb})=0$ and $\gamma_{2r}(\boldsymbol{nb})=1$ for $C-b_2+b_{2r}<\boldsymbol{nb}\le T$, otherwise $\gamma_{2r}(\boldsymbol{nb})=0$.

Note that the approximations of (14) are similar to those introduced in the single-threshold model of [19].

Since the values of $x(\boldsymbol{n})=1$, when $1\le j\le C$, it is proved in [20] that:

$$\begin{aligned}&(N_1-n_1+1)a_1b_1G(j-b_1)+(N_2-n_2+1)a_2b_2G(j-b_2)+\\&(N_2-n_2-n_{2r}+1)a_{2r}b_{2r}\gamma_{2r}(j)G(j-b_{2r})=jG(j),\\&1\le j\le C\end{aligned} \tag{15}$$

where: $G(j)$ is the link occupancy distribution, $\gamma_{2r}(j)=1$ for $C-b_2+b_{2r}<j$, otherwise $\gamma_{2r}(j)=0$.

When $C<j\le T$, we have $\gamma_2(j)=0$ and due to (3), we may write (14) as follows:

$$\begin{aligned}&x(\boldsymbol{n})\left[(N_1-n_1+1)a_1b_1P(\boldsymbol{n}_1^-)+(N_2-n_2-n_{2r}+1)\gamma_{2r}(\boldsymbol{nb})a_{2r}b_{2r}P(\boldsymbol{n}_{2r}^-)\right]\\&=(n_1b_1x(\boldsymbol{n}_1^-)+n_2b_2x(\boldsymbol{n}_2^-)+n_{2r}b_{2r}x(\boldsymbol{n}_{2r}^-))P(\boldsymbol{n})\end{aligned} \tag{16}$$

or

$$\begin{aligned}&(N_1-n_1+1)a_1b_1P(\boldsymbol{n}_1^-)+(N_2-n_2-n_{2r}+1)\gamma_{2r}(\boldsymbol{nb})a_{2r}b_{2r}P(\boldsymbol{n}_{2r}^-)\\&=CP(\boldsymbol{n})\end{aligned} \tag{17}$$

Summing over the set of states $\{\boldsymbol{n}\in\Omega|\boldsymbol{nb}=j\}$, where $\Omega=\{\boldsymbol{n}:0\le\boldsymbol{nb}\le T\}$, we have:

$$\begin{aligned}&(N_1-n_1+1)a_1b_1\sum_{\{\boldsymbol{n}|\boldsymbol{nb}=j\}}P(\boldsymbol{n}_1^-)\\&+(N_2-n_2-n_{2r}+1)\gamma_{2r}(\boldsymbol{nb})a_{2r}b_{2r}\sum_{\{\boldsymbol{n}|\boldsymbol{nb}=j\}}P(\boldsymbol{n}_{2r}^-)\\&=C\sum_{\{\boldsymbol{n}|\boldsymbol{nb}=j\}}P(\boldsymbol{n})\end{aligned} \tag{18}$$

or

$$\begin{aligned}&(N_1-n_1+1)a_1b_1G(j-b_1)+(N_2-n_2-n_{2r}+1)\gamma_{2r}(j)a_{2r}b_{2r}G(j-b_{2r})\\&=CG(j)\end{aligned} \tag{19}$$

since by definition: $G(j)=\sum_{\boldsymbol{n}\in\Omega_j}P(\boldsymbol{n})$.

The combination of (15) and (19) gives the following recursive formula for the $G(j)$'s calculation when K=2 and only calls of the 2nd service-class retry:

$$\begin{aligned}&\min(j,C)G(j)=(N_1-n_1+1)a_1b_1G(j-b_1)+\\&(N_2-n_2+1)a_2b_2\gamma_2(j)G(j-b_2)+\\&(N_2-n_2-n_{2r}+1)a_{2r}b_{2r}\gamma_{2r}(j)G(j-b_{2r}),\\&1\le j\le T\end{aligned} \tag{20}$$

where: $\gamma_2(j)=1$ for $1\le j\le C$, otherwise $\gamma_2(j)=0$ and $\gamma_{2r}(j)=1$ for $C-b_2+b_{2r}<j\le T$, otherwise $\gamma_{2r}(j)=0$.

In the general case of K different service-classes, where all calls may retry, (20) takes the form:

$$G(j)=\left\langle\begin{array}{l}1\ \ for\ j=0\\ \dfrac{1}{\min(j,C)}\displaystyle\sum_{k=1}^{K}(N_k-n_k+1)a_kb_k\gamma_k(j)G(j-b_k)+\\ \dfrac{1}{\min(j,C)}\displaystyle\sum_{k=1}^{K}(N_k-n_k-n_{kr}+1)a_{kr}b_{kr}\gamma_{kr}(j)G(j-b_{kr})\\ for\ j=1,\dots,T\\ 0\ \ otherwise\end{array}\right\rangle \tag{21}$$

where:

$$a_{kr}=v_k\mu_{kr}^{-1},\ \gamma_k(j)=\begin{cases}1 & for\ 1\le j\le C\ and\ b_{kr}>0\\ 1 & for\ 1\le j\le T\ and\ b_{kr}=0\\ 0 & otherwise\end{cases},$$

$$\gamma_{kr}(j)=\begin{cases}1 & for\ C-b_k+b_{kr}<j\le T\ and\ b_{kr}>0\\ 0 & otherwise\end{cases}.$$

The calculation of $G(j)$'s in (21) requires the values of n_k and n_{kr} which are unknown. In other finite multirate loss models (e.g., [6], [17], [19]) there exist methods for the determination of these values through an equivalent stochastic system, with the same traffic description parameters and set of states. However, the state space determination of the equivalent system is complex, especially for large systems that serve many service-classes. Thus, we avoid such methods and approximate n_k and n_{kr} in state j, $n_k(j)$, $n_{kr}(j)$, as the mean number of service-class k calls in state j, $y_k(j)$ and $y_{kr}(j)$, respectively, when Poisson arrivals are considered. Such approximations are common in the literature (e.g., [21], [29], [30]). In that case, (21) takes the form:

$$G(j)=\begin{cases}1 \ \ for\ j=0\\ \frac{1}{\min(j,C)}[\sum_{k=1}^{K}(N_k-y_k(j-b_k))a_k b_k \gamma_k(j)G(j-b_k)+\\ \sum_{k=1}^{K}(N_k-(y_k(j-b_{kr})+y_{kr}(j-b_{kr}))a_{kr}b_{kr}\gamma_{kr}(j)G(j-b_{kr})]\\ for\ j=1,...,T\\ 0 \ \ otherwise\end{cases} \quad (22)$$

where the values of $y_k(j)$ and $y_{kr}(j)$ are given by:

$$y_k(j)=a_{k,\inf}\gamma_k(j)G_{\inf}(j-b_k)/G_{\inf}(j) \quad (23)$$

$$y_{kr}(j)=a_{kr,\inf}\gamma_{kr}(j)G_{\inf}(j-b_{kr})/G_{\inf}(j) \quad (24)$$

where: $a_{k,\inf}$ and $G_{\inf}(j)$ are the offered traffic-load (in erl) of service-class k and the link occupancy distribution, respectively, of the corresponding infinite model [22]:

$$G_{\inf}(j)=\begin{cases}1 \ \ for\ j=0\\ \frac{1}{\min(j,C)}\sum_{k=1}^{K}a_{k,\inf}b_k\gamma_k(j)G_{\inf}(j-b_k)+\\ \frac{1}{\min(j,C)}\sum_{k=1}^{K}a_{kr,\inf}b_{kr}\gamma_{kr}(j)G_{\inf}(j-b_{kr})\\ for\ j=1,...,T\\ 0 \ \ otherwise\end{cases} \quad (25)$$

Having determined $G(j)$'s according to (22), we calculate the following performance measures:

1) The Time Congestion (TC) probabilities of service-class k, denoted as P_{b_k}, which is the probability that at least T-b_{kr}+1 b.u. are occupied:

$$P_{b_k}=\sum_{j=T-b_{kr}+1}^{T}G^{-1}G(j) \quad (26)$$

where: $G=\sum_{j=0}^{T}G(j)$ is a normalization constant.

2) The Call Congestion (CC) probabilities of service-class k, denoted as C_{b_k}, which is the probability that a service-class k call is blocked with b_{kr}:

$$C_{b_k}=\sum_{j=T-b_{kr}+1}^{T}G^{-1}G(j) \quad (27)$$

where $G(j)$'s are determined for a system with N_k - 1 traffic sources.

3) The link utilization, denoted as U:

$$U=\sum_{j=1}^{C}jG^{-1}G(j)+\sum_{j=C+1}^{T}CG^{-1}G(j) \quad (28)$$

III. The Elastic Multi Retry Loss Model For Quasi-Random Input

Similar to the single-retry model, the multi-retry model does not have a PFS and therefore the $G(j)$'s calculation is based on an approximate but recursive formula. In the multi-retry model, a blocked service-class k call retries $s(k)$ times with parameters: $(b_{kr_s},\mu_{kr_s}^{-1})$ for s=1,…,$s(k)$ where $b_{kr_{s(k)}}$<…<b_{kr_1}<b_k and $\mu_{kr_{s(k)}}^{-1}>...>\mu_{kr_1}^{-1}>\mu_k^{-1}$. The determination of $G(j)$'s is based on (29) whose proof is similar to that of (21) and therefore is not presented:

$$G(j)=\begin{cases}1 \ \ for\ j=0\\ \frac{1}{\min(j,C)}[\sum_{k=1}^{K}(N_k-n_k+1)a_k b_k\gamma_k(j)G(j-b_k)+\\ \sum_{k=1}^{K}\sum_{s=1}^{s(k)}(N_k-(n_k+n_{kr_1}+...+n_{kr_{s(k)}})+1)a_{kr_s}b_{kr_s}\gamma_{kr_s}(j)G(j-b_{kr_s})]\\ for\ j=1,...,T\\ 0 \ \ otherwise\end{cases} \quad (29)$$

where: $a_{kr_S}=v_k\mu_{kr_S}^{-1}$, $\gamma_k(j)=\begin{cases}1 & for\ 1\le j\le C\ and\ b_{kr_s}>0\\ 1 & for\ 1\le j\le T\ and\ b_{kr_s}=0\\ 0 & otherwise\end{cases}$,

$$\gamma_{kr_s}(j)=\begin{cases}1 & for\ \ C-b_{kr_{s-1}}+b_{kr_s}<j\le C\ \ if\ \ s\ne s(k)\\ 1 & for\ \ C-b_{kr_{s-1}}+b_{kr_s}<j\le T\ \ if\ \ s=s(k)\\ 0 & otherwise\end{cases}.$$

As in the SRM, we approximate $n_k(j)$ and $n_{kr_s}(j)$ for s=1,…,$s(k)$ with the corresponding values of the infinite model [22]. In that case, (29) takes the form:

$$G(j)=\begin{cases}1 \ \ for\ j=0\\ \frac{1}{\min(j,C)}[\sum_{k=1}^{K}(N_k-y_k(j-b_k))a_k b_k\gamma_k(j)G(j-b_k)+\\ \sum_{k=1}^{K}\sum_{s=1}^{s(k)}(N_k-Y_k(j-b_{kr_S}))a_{kr_s}b_{kr_s}\gamma_{kr_s}(j)G(j-b_{kr_s})]\\ for\ j=1,...,T\\ 0 \ \ otherwise\end{cases} \quad (30)$$

where:

$Y_k(j-b_{kr_s})=y_k(j-b_{kr_s})+y_{kr_1}(j-b_{kr_s})+...+y_{kr_{s(k)}}(j-b_{kr_s})$

and the values of $y_k(j)$ and $y_{kr_s}(j)$ are given by:

$$y_k(j)=a_{k,\inf}\gamma_k(j)G_{\inf}(j-b_k)/G_{\inf}(j) \quad (31)$$

$$y_{kr_s}(j)=a_{kr,\inf}\gamma_{kr_s}(j)G_{\inf}(j-b_{kr_s})/G_{\inf}(j) \quad (32)$$

where $G_{\inf}(j)$ refers to the link occupancy distribution of the corresponding infinite model [22]:

$$G_{\inf}(j)=\begin{cases}1 \ \ for\ j=0\\ \frac{1}{\min(j,C)}\sum_{k=1}^{K}a_{k,\inf}b_k\gamma_k(j)G_{\inf}(j-b_k)+\\ \frac{1}{\min(j,C)}\sum_{k=1}^{K}\sum_{s=1}^{s(k)}a_{kr_s,\inf}b_{kr_s}\gamma_{kr_s}(j)G_{\inf}(j-b_{kr_s})\\ for\ j=1,...,T\\ 0 \ \ otherwise\end{cases} \quad (33)$$

A similar procedure in the case of retrials where calls have fixed bandwidth requirements (bandwidth compression is not permitted) has been proposed in [21].

Having determined $G(j)$'s according to (30), we calculate TC and CC probabilities of service-class k, according to (26) and (27), respectively, by replacing b_{kr} with $b_{kr_{s(k)}}$. Link utilization is calculated by (28) where the values of $G(j)$'s are given by (30).

IV. NUMERICAL EXAMPLES - EVALUATION

We consider an application example and compare the analytical TC probabilities with those obtained by simulation. The latter is based on SIMSCRIPT III [31]. For comparison, we show the corresponding analytical results assuming Poisson arrivals [22]. Simulation results are mean values of 7 runs with 95% confidence interval. The resultant reliability ranges of the simulation measurements are very small and, therefore, we present only mean values.

Consider a link of capacity C=80 b.u. that accommodates three service-classes of elastic calls. All calls arrive in the system according to a quasi-random process. The traffic characteristics of each service-class are the following:

1st service-class: N_1=100, $v_1 = 0.20$, $b_1 = 1$ b.u.

2nd service-class: N_2=100, $v_2 = 0.06$, $b_2 = 2$ b.u.

3rd service-class: N_3=100, $v_3 = 0.02$, $b_3 = 6$ b.u.

The call holding time is exponentially distributed with mean value $\mu_1^{-1}=\mu_2^{-1}=\mu_3^{-1}=1$. Calls of the 3rd service-class may retry two times with reduced bandwidth requirement: $b_{3_{r1}}$ = 5 b.u. and $b_{3_{r2}}$ = 4 b.u. and increased service time so that $a_3 b_3 = a_{3_{r1}} b_{3_{r1}} = a_{3_{r2}} b_{3_{r2}}$, where $a_k = v_k \mu_k^{-1}$, k=1,2,3. The corresponding Poisson traffic-loads are: $a_{1,\inf} = 20$, $a_{2,\inf} = 6$, $a_{3,\inf} = 2$ erl. In the x-axis of all figures, we assume that v_3 remains constant while v_1, v_2 increase in steps of 0.01 and 0.005, respectively. The last value of $v_1 = 0.28$ while that of $v_2 = 0.10$. The corresponding last values of the Poisson traffic-loads are: $a_{1,\inf} = 28$, $a_{2,\inf} = 10$, $a_{3,\inf} = 2$ erl.

Two different values of T are considered: a) $T = C = 80$ b.u. where no bandwidth compression takes place. In that case, the proposed model gives exactly the same results with the model of [21], b) T=82 b.u. where bandwidth compression takes place and $r_{\min} = C/T$=80/82.

In Fig. 1, we present the analytical and simulation TC probabilities results of the 1st service-class for all values of T. Similar results are presented in Fig. 2, for the 2nd service-class and in Fig. 3 for the 3rd service-class (TC probabilities of calls with $b_{3_{r2}}$). All figures presented herein show that: i) the model's accuracy is absolutely satisfactory compared to simulation, ii) the increase of T above C results in the decrease of TC probabilities due to the existence of the compression mechanism and iii) the results obtained by the infinite model [22] fail to approximate the results of the proposed finite model.

V. CONCLUSION

We propose multirate retry loss models that support elastic traffic assuming that calls arrive in the link according to a quasi-random process and have an exponentially distributed service time. Blocked calls have the ability to retry to be connected in the system one or more times with reduced bandwidth and increased service time requirements. Furthermore, if a retry call is blocked with its last bandwidth requirement, it can still be accepted in the system by compressing its bandwidth together with the bandwidth of all in-service calls. The proposed models do not have a PFS. However, we propose approximate but recursive formulas for the calculation of the link occupancy distribution and consequently time and call congestion probabilities as well as link utilization. Simulation results verify the analytical results.

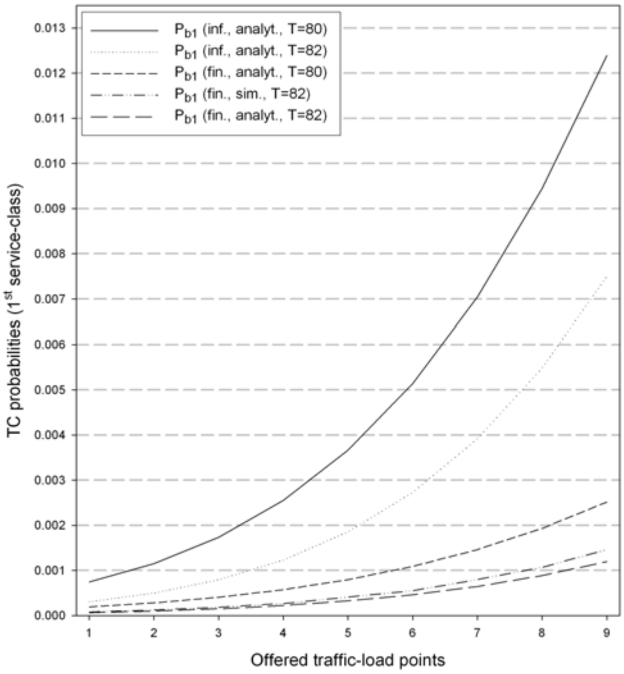

Figure 1. TC probabilities – 1st service-class.

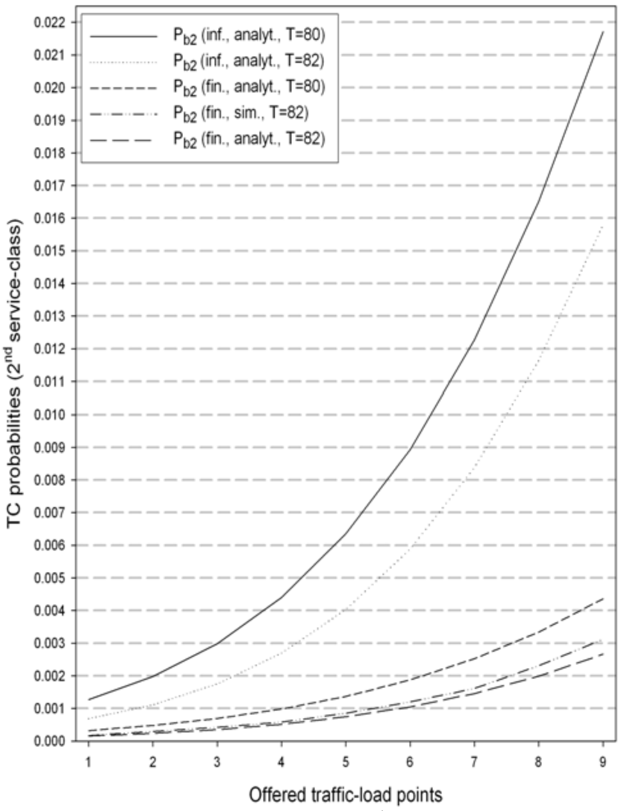

Figure 2. TC probabilities – 2nd service-class.

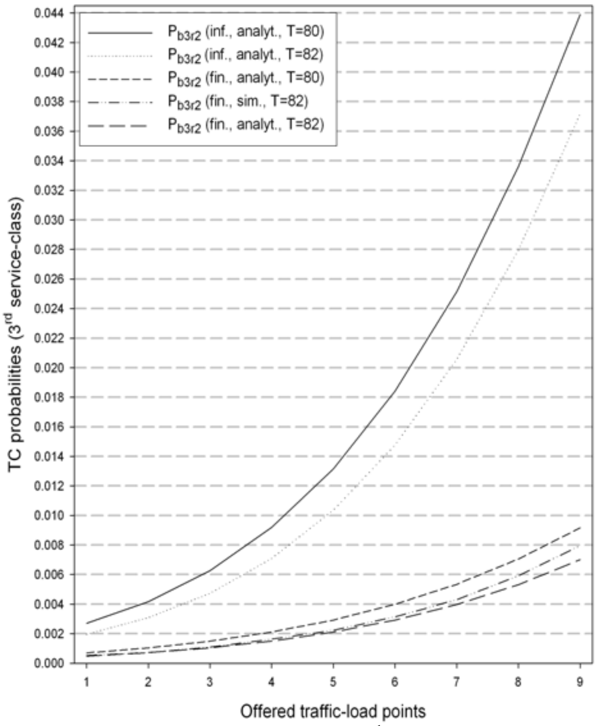

Figure 3. TC probabilities – 3rd service-class.

REFERENCES

1. J. Kaufman, "Blocking in a shared resource environment", IEEE Trans. Commun. vol. 29, no. 10, pp. 1474-1481, October 1981.
2. J. Roberts, "A service system with heterogeneous user requirements", in: G. Pujolle (Ed.), Performance of Data Communications systems and their applications, North Holland, Amsterdam, pp.423-431, 1981.
3. M. Stasiak and M. Glabowski, "A simple approximation of the link model with reservation by a one-dimensional Markov chain", Performance Evaluation, vol. 41, July 2000, pp.195–208.
4. S. Rácz, B. Gerő, and G. Fodor, "Flow level performance analysis of a multi-service system supporting elastic and adaptive services", Performance Evaluation, vol. 49, Sept. 2002, pp. 451-469.
5. I. Moscholios, P. Nikolaropoulos and M. Logothetis, "Call level blocking of ON-OFF traffic sources with retrials under the complete sharing policy", Proc. of 18th ITC, Berlin, Germany, 31 August – 5 September 2003, pp. 811-820.
6. I. Moscholios, M. Logothetis, and M. Koukias, "An ON-OFF Multi-Rate Loss Model of Finite Sources", IEICE Trans. Commun., vol. E90-B, no. 7, pp.1608-1619, July 2007.
7. Q. Huang, King-Tim Ko and V. Iversen, "Approximation of loss calculation for hierarchical networks with multiservice overflows", IEEE Trans. Commun., Vol. 56, Issue 3, pp. 466-473, March 2008.
8. I. Moscholios and M. Logothetis, "The Erlang multirate loss model with Batched Poisson arrival processes under the bandwidth reservation policy", Computer Communications, Vol. 33, Supplement 1, pp. S167-S179, Nov. 2010.
9. D. Staehle and A. Mäder, "An Analytic Approximation of the Uplink Capacity in a UMTS Network with Heterogeneous Traffic", Proc. 18th ITC, Berlin, pp. 81-90, Sept. 2003.
10. V. Iversen, V. Benetis, N. Ha, S. Stepanov, "Evaluation of Multi-service CDMA Networks with Soft Blocking", Proc. of ITC Specialist Seminar, Antwerp, Belgium, pp. 223-227, August/September 2004.
11. V. Vassilakis, G. Kallos, I. Moscholios and M. Logothetis, "Call-Level Analysis of W-CDMA Networks Supporting Elastic Services of Finite Population", IEEE ICC 2008, Beijing, China, 19-23 May 2008.
12. M. Logothetis, V. Vassilakis and I. Moscholios, "Call-level Performance Modeling and QoS Assessment of W-CDMA Networks", Nova Science Publishers, New York, USA, 2009, pp. 57-90.
13. M. Glabowski, M. Stasiak, A. Wisniewski, and P. Zwierzykowski, "Blocking Probability Calculation for Cellular Systems with WCDMA Radio Interface Servicing PCT1 and PCT2 Multirate Traffic", IEICE Trans. Commun., vol.E92-B, pp.1156-1165, April 2009.
14. K. Kuppuswamy, D. Lee, "An analytic approach to efficiently computing call blocking probabilities for multiclass WDM networks", IEEE/ACM Trans. Netw., vol. 17, issue 2, pp. 658-670, April 2009.
15. J. Vardakas, I. Moscholios, M. Logothetis and V. Stylianakis, "An Analytical Approach for Dynamic Wavelength Allocation in WDM-TDMA PONs Servicing ON-OFF Traffic", IEEE/OSA Journal of Optical Commun. Netw., vol. 3, no. 4, pp. 347-358, April 2011.
16. J. Vardakas, I. Moscholios, M. Logothetis, and V. Stylianakis, "Performance Analysis of OCDMA PONs Supporting Multi-Rate Bursty Traffic", IEEE Trans. Commun., vol. 61, no. 8, pp. 3374-3384, August 2013.
17. G. Stamatelos and J. Hayes, "Admission control techniques with application to broadband networks", Comput. Commun., vol. 17, no. 9, pp. 663-673, 1994.
18. J. Kaufman, "Blocking with retrials in a completely shared resource environment", Performance Evaluation, vol. 15, issue 2, pp. 99-113, June 1992.
19. I. Moscholios, M. Logothetis and P. Nikolaropoulos, "Engset Multi-Rate State-Dependent Loss Models", Performance Evaluation, vol. 59, issues 2-3, February 2005, pp. 247-277.
20. G. Stamatelos and V. Koukoulidis, "Reservation-Based Bandwidth Allocation in a Radio ATM Network", IEEE/ACM Trans. Netw., vol. 5, no. 3, pp. 420-428, June 1997.
21. I. Moscholios, M. Logothetis and G. Kokkinakis, "A Simplified Blocking Probability Calculation in the Retry Loss Models for Finite Sources", Proc. of Communication Systems, Networks and Digital Signal Processing, Patras, Greece, 19-21 July 2006.
22. I. Moscholios, V. Vassilakis, J. Vardakas and M. Logothetis, "Retry loss models supporting elastic traffic", Advances in Electronics and Telecommunications, Poznan University of Technology, Poland, Vol. 2, No. 3, September 2011, pp. 8-13.
23. I. Moscholios, V. Vassilakis, J. Vardakas and M. Logothetis, "Call Blocking Probabilities of Elastic and Adaptive Traffic with Retrials", Proc. of AICT 2012, Stuttgart, Germany, 27 May-1 June 2012.
24. I. Moscholios, V. Vassilakis, M. Logothetis and M. Koukias, "QoS Equalization in a Multirate Loss Model of Elastic and Adaptive Traffic with Retrials", Proc. of EMERGING 2013, Porto, Portugal, 29 September – 4 October 2013.
25. G. Kallos, V. Vassilakis, I. Moscholios, and M. Logothetis, "Performance Modelling of W-CDMA Networks Supporting Elastic and Adaptive Traffic", Proc. 4th IFIP HET-NETs, Ilkley, U.K., pp. 09/1-09/10, 11-13 Sept. 2006.
26. G. Fodor and M. Telek, "Bounding the Blocking Probabilities in Multirate CDMA Networks Supporting Elastic Services", IEEE/ACM Trans. Netw., vol. 15, Aug. 2007, pp. 944-956.
27. V. Vassilakis, G. Kallos, I. Moscholios and M. Logothetis, "An Analytical Model for Elastic Service-classes in W-CDMA Networks", Performance Modelling and Analysis of Heterogeneous Networks, River Publishers, 2009, pp. 277-299.
28. H. Akimaru and K. Kawashima, Teletraffic – Theory and Applications, 2nd edition, Springer-Verlag, Berlin, 1999.
29. M. Glabowski and M. Stasiak, "An approximate model of the full-availability group with multi-rate traffic and a finite source population", in Proc. of 12th MMB&PGTS, Dresden, Germany, pp. 195-204, Sept. 2004.
30. I. Moscholios, M. Logothetis, V. Stylianakis and J. Vardakas, "The Priority Wavelength Release Protocol for Dynamic Wavelength Allocation in WDM-TDMA PONs Supporting Random and Quasi-Random Bursty Traffic", Proc. of 18th NOC 2013, Graz, Austria, 10-12 July 2013.
31. Simscript III, http://www.simscript.com