Performance Evaluation of the Threshold Call Admission Policy in Multi-rate Loss Systems

Ioannis D. Moscholios, Iskanter-Alexandros Chousainov, Panagiotis I. Panagoulias, Panagiotis G. Sarigiannidis, and Michael D. Logothetis

1 Department of Informatics and Telecommunications, University of the Peloponnese, Tripolis, Greece
2 Department of Electrical and Computer Engineering, University of West Macedonia, Kozani, Greece
3 WCL, Department of Electrical and Computer Engineering, University of Patras, Patra, Greece

Abstract—In this paper we consider a link, characterized by specific capacity, that services multi-rate random or quasi-random traffic. Random traffic is generated by an infinite number of traffic sources, while quasi-random traffic is generated by a finite population of traffic sources. The link is modeled as a multi-rate loss system. Handover and new calls are distinguished. New calls compete for the available bandwidth under a threshold call admission policy. In that policy, a new call of a particular service-class is not allowed to enter the system if the in-service handover and new calls of the same service-class plus the new call, exceed a predefined threshold (which can be different for each service-class). On the other hand, handover calls compete for the available bandwidth based on the complete sharing policy. We show that the steady state probabilities in the proposed models have a product form solution (PFS). The PFS leads to a convolution algorithm for accurate calculation of congestion probabilities and link utilization.

Keywords—blocking, congestion, convolution, product form, quasi-random, random.

1. Introduction

A call admission policy is a necessary quality of service (QoS) mechanism in a link that services multi-rate traffic, since it provides access to the resources needed by arriving calls. The complete sharing (CS) policy, also known as a “full accessibility” policy, is considered in the literature as the simplest call admission policy [1], [2]. However, it is important to mention that the CS policy is not designed with the aim of providing a certain QoS to calls of a particular service-class. This is because the only criterion for the acceptance of a new call in a system that adopts the CS policy is the availability of bandwidth required by that call. In that sense, the CS policy is unfair with regard to calls characterized by high bandwidth requirements, since it results in high call blocking probabilities (CBP), e.g. [3]–[12]).

Focusing on wireless networks, the CS policy cannot treat handover and new calls of the same service-class differently, though it is obvious that handover calls should have higher priority compared to new calls. Based on these factors, various other call admission policies have been analyzed in the literature, including the bandwidth reservation (BR) policy (e.g. [13]–[22]), the multiple fractional channel reservation (MFCR) policy (e.g. [23]–[25]) and the threshold (TH) policy, all attempting to prioritize certain types of calls. The BR policy may introduce a simple prioritization mechanism while dealing with calls of different service-classes, by reserving bandwidth units (b.u.) favoring high-speed calls or handover calls. The MFCR policy generalizes the BR policy by allowing the reservation of a real (not an integer) number of b.u. We consider the TH policy, in which in-service calls of a particular service-class cannot exceed a specific threshold, even if resources are available in the system. This policy cannot be described via a particular selection of BR or MFCR parameters and has been applied not only in wired but also in wireless and satellite networks (e.g. [26]–[40]).

To prioritize handover calls over new calls, a variant of the TH policy may be considered, known as the threshold call admission (TCA) policy. In the TCA policy, handover calls are assigned with higher priority than new calls. More specifically, a new call (of a particular service-class) cannot enter the system if the number of in-service handover and new calls of the same service-class, plus the new call, exceeds a certain threshold (which is predefined and which differs for each particular service-class). On the other hand, the TCA policy is not applied to handover calls, i.e., these calls are blocked only if their required bandwidth is not available at the time of their arrival (CS policy). As a possible application of the TCA policy, consider the case of a cloud radio access network which consists of a large number of distributed base stations where the remote radio heads (RRHs), including radio frequency components and antennas, are separated from the baseband signal processing servers, named baseband units [41]. A possible movement of users from the serving RRH to a new RRH may create handover traffic [42]. In that case, the TCA policy may
be applied in the new RRH in order to prioritize handover traffic over traffic created by new users. 

In this paper, we consider a link under the TCA policy, which services random or quasi-random multi-rate traffic. Random traffic is generated by an infinite number of calls (handover and new) of different service-classes, while quasi-random traffic is generated by a finite number of calls of different service-classes. We model the link as a multi-rate loss system and show that this loss system may be analytically described via a continuous time and reversible Markov chain, both in the case of random and quasi-random traffic. In both cases, the steady state distribution has a product form solution (PFS). Based on the PFS, a convolution algorithm is proposed for the calculation of congestion probabilities and link utilization. In general, convolution algorithms can be efficiently applied only in PFS models. Their main advantage is the fact that they capture information about the number of in-service calls belonging to each service-class. That information is significant when the TH and TCA policies are considered. On the other hand, in non-PFS models (such as models adopting the BR policy), the application of convolution algorithms for CBP determination may be quite complicated [43]–[46].

This paper is organized as follows. In Section 2, we review the model of [47] which considers a link under the TCA policy that accommodates multi-rate Poisson traffic (TCA model). In Section 3, the corresponding model for finite sources (f-TCA model) is proposed, the PFS is shown and a convolution algorithm for the calculation of congestion probabilities and link utilization is provided. In Section 4, analytical CBP and link utilization results for the TCA model, the CS model of [48] and the TH model of [26] assuming multi-rate Poisson traffic, are presented. Additionally, the analytical congestion probabilities results for the TCA and the f-TCA models are shown. The paper concludes in Section 5. In the Appendix, a tutorial example is presented in order to clarify the necessary calculations for the determination of congestion probabilities, when the convolution algorithm is applied in the f-TCA model.

2. Review of the TCA Model

2.1. Analytical Model

In the model of [47], we consider a link with a fixed capacity of $C$ b.u. and let $j$ express the b.u. occupied in the link $(j = 0, 1, \ldots, C)$. The link services Poisson traffic generated via $K$ different service-classes. Let the arrival rate of handover and new service-class $k$ $(k = 1, \ldots, K)$ calls be $\lambda_{hk}$ and $\lambda_{nk}$, respectively. The total arrival rate of service-class $k$ calls is $\lambda_k = \lambda_{hk} + \lambda_{nk}$. Each service-class $k$ (new or handover) call requires $b_k$ b.u. in order to be accepted in the link and has a generally distributed service time with a mean $\mu_k^{-1}$. New and handover calls are accepted in the link based on the TCA and the CS policy, respectively.

In the TCA policy, each service-class $k$ has a predefined threshold $n_{k,\text{max}}$ whose role is to allow the TCA policy to prioritize handover calls against new calls. More precisely, a new call of service-class $k$ will be accepted in the system if:

1. $j + b_k \leq C$ (i.e. if there exist available b.u. in the link)
2. the number of in-service handover and new calls of service-class $k$, $n_k$, plus the new call is at most equal to $n_{k,\text{max}}$, i.e. $n_k + 1 \leq n_{k,\text{max}}$.

The second constraint expresses the fact that a new call may not become accepted for service even if the b.u. it requires are available. On the other hand, a handover service-class $k$ call needs to meet the first constraint only in order to be accepted in the link.

By assuming that the system is in a steady state, let $\mathbf{n} = (n_1, \ldots, n_k, \ldots, n_K)$ be the steady state vector and $\pi(\mathbf{n})$ be the corresponding probability distribution, which can be expressed by the following PFS [47]:

$$
\pi(\mathbf{n}) = G^{-1} \left( \prod_{k=1}^{K} \left( \frac{n_k^p}{n_k!} \right) \right),
$$

where:

$$
\frac{n_k^p}{n_k!} = \begin{cases} 
\frac{\alpha_k^p}{\alpha_k} & \text{for } n_k \leq n_{k,\text{max}} \\
\frac{\alpha_{k,\text{max}}^{n_k-n_{k,\text{max}}}}{\alpha_k} & \text{for } n_k > n_{k,\text{max}}
\end{cases}
$$

$G$ refers to a constant used for normalization and is expressed with the following formula:

$$
G = \sum_{\mathbf{n} \in \Omega} \left( \prod_{k=1}^{K} \left( \frac{n_k^p}{n_k!} \right) \right),
$$

$\alpha_k = \frac{\lambda_k}{\mu_k}$, $\alpha_{hk} = \frac{\lambda_{hk}}{\mu_k}$ (in Erl) while $\Omega$ refers to the state space, $\Omega = \{ \mathbf{n} : 0 \leq n_b \leq C, k = 1, \ldots, K \}$ with $n_b = \sum_{k=1}^{K} n_k b_k$ and $b = (b_1, \ldots, b_K)^T$.

Based on Eq. (1), we see that the term $\frac{\alpha_{k,\text{max}}^{n_k-n_{k,\text{max}}}}{\alpha_k}$ expresses the application of the TCA policy in new service-class $k$ calls when $n_k > n_{k,\text{max}}$.

2.2. Convolution Algorithm

According to Eq. (1), the following convolution algorithm can be adopted in the TCA model for the determination of the various performance measures including CBP.

**Step 1.** Assuming that the link of capacity $C$ services only service-class $k$ $(k = 1, \ldots, K)$ calls, determine the occupancy distribution $q_k(j)$, where $j = n_k b_k$, as follows:

$$
q_k(j) = \begin{cases} 
q_k(0) \frac{n_k^p}{n_k!} & \text{for } n_k \leq n_{k,\text{max}} \\
q_k(0) \frac{n_{k,\text{max}}^{n_k-n_{k,\text{max}}}}{n_k^p} & \text{for } n_k > n_{k,\text{max}}
\end{cases}
$$

(2)
Step 2. Based on a successive convolution of all service-classes excluding service-class \( k \), determine the occupancy distribution \( Q_{(-k)} \):

\[
Q_{(-k)} = q_1 \cdots q_{k-1} \cdots q_{k+1} \cdots q_K .
\]  

(3)

More precisely, the corresponding convolution operation of service-classes \( k \) and \( r \), is given by:

\[
q_k \cdot q_r = \begin{cases} 
q_k(0)q_r(0), & \sum_{x=0}^{C} q_k(x)q_r(1-x), \\
\cdots & \cdots \\
\sum_{x=0}^{C} q_k(x)q_r(C-x) 
\end{cases}.
\]  

(4)

Step 3. Compute the convolution of \( Q_{(-k)} \) and \( q_k \), from:

\[
Q_{(-k)} \cdot q_k = \begin{cases} 
Q_{(-k)}(0)q_k(0), & \sum_{x=0}^{C} Q_{(-k)}(x)q_k(1-x), \\
\cdots & \cdots \\
\sum_{x=0}^{C} Q_{(-k)}(x)q_k(C-x) 
\end{cases}.
\]  

(5)

Via Eq. (5), we can also determine the link occupancy distribution \( q(j) \), according to the following formulas:

\[
q(0) = \frac{Q_{(-k)}(0)q_k(0)}{G},
\]

\[
q(j) = \frac{\sum_{x=0}^{C} Q_{(-k)}(x)q_k(j-x)}{G}, \quad j = 1 \ldots C .
\]  

(6)

Based on Eq. (6), we can determine the CBP of service-class \( k \) calls (new or handover), as well as the link utilization (in b.u.) via Eqs. (7), (8) and (9), respectively:

\[
B_{k}=\sum_{j=C-b_k+1}^{C} q(j) + G^{-1}\sum_{t=n_k \text{,max}}^{C-b_k} \sum_{y=t}^{C} Q_{(-k)}(C-b_k-y) ,
\]  

(7)

\[
B_{kh} = \sum_{j=C-b_k+1}^{C} q(j),
\]  

(8)

\[
U = \sum_{j=1}^{C} j q(j).
\]  

(9)

On the right hand side of Eq. (7), the first term refers to call blocking due to the unavailability of b.u., while the second term refers to the TCA policy. Indeed, in states \( t = n_k \text{,max} \) \( b_k \), call blocking occurs (due to the threshold \( n_k \text{,max} \)) although there exist available b.u. for a new service-class \( k \) call.

3. f-TCA Model

3.1. Analytical Model

In the proposed f-TCA model, service-class \( k \) calls arrive in the link according to a quasi-random process with mean arrival rate \( \lambda_{k, \text{fin}} = (N_k - n_k) v_k \), where \( N_k \) refers to the finite number of service-class \( k \) sources that are capable of generating traffic, \( n_k \) is the number of in-service calls (new and handover), while \( v_k \) is the mean call arrival rate per idle service-class \( k \) source. Similarly, let \( v_{k,b} \) and \( v_{k,m} \) be the mean call arrival rate per idle source of handover and new service-class \( k \), respectively. Then, \( v_k = v_{k,b} + v_{k,m} \). The traffic-load per idle source offered is \( q_k \text{,idle} = \frac{v_k}{\mu_k} \).

Assuming that \( N_k \rightarrow \infty \), for \( k = 1, \ldots, K \), and that the total traffic load offered is constant, then the call arrival process becomes Poisson and we have the TCA model of [47]. Each service-class \( k \) call requires \( b_k \) b.u. in order to be accepted in the link and has a generally distributed service time with mean \( \mu_k \). New and handover calls are accepted in the link based on the TCA and the CS policy, respectively.

![State transition diagram](image)

**Fig. 1.** State transition diagram, for service-class \( k \) calls, between states \( n_k \), \( n \) and \( n_k^+ \).

Similarly to the TCA model, we assume that the system is in a steady state and use \( \pi_f(n) \) to denote the probability distribution of \( n = (n_1, \ldots, n_K) \). Figure 1 presents the state transition diagram of the proposed model for service-class \( k \) calls. Regarding the global balance equation for state \( n \), it can be expressed as \( \text{rate into state } n = \text{rate out of state } n \):

\[
\sum_{k=1}^{K} (N_k - n_k + 1)v_k(n_k^-) \pi_f(n^-) + \sum_{k=1}^{K} (n_k + 1) \mu_k \pi_f(n_k^+) = \sum_{k=1}^{K} (N_k - n_k)v_k(n) \pi_f(n) + \sum_{k=1}^{K} n_k \mu_k \pi_f(n) ,
\]  

(10)

where: \( v_k(n) = \begin{cases} 
v_k & \text{for } n_k \leq n_k \text{,max} , \\
v_{kh} & \text{for } n_k > n_k \text{,max} 
\end{cases} \), \( \Omega \) is the state space

\[
\Omega = \{ n : nb \leq C, k = 1, \ldots, K \},
\]

\[
\left\{ n_k^- = \{ n_1, \ldots, n_k - 1, \ldots, n_K \} \right\} \quad \text{and} \quad \left\{ n_k^+ = \{ n_1, \ldots, n_k + 1, \ldots, n_K \} \right\} .
\]

The Markov chain of the proposed loss model retains reversibility and therefore local balance exists between the adjacent states: (1) \( n_k^- \) and \( n \) and (2) \( n \) and \( n_k^+ \). To this end, we can write the following local balance equations (\( \text{rate up} = \text{rate down} \)), for \( k = 1, \ldots, K \) and \( n \in \Omega \):

\[
(N_k - n_k + 1)v_k(n_k^-) \pi_f(n^-) = n_k \mu_k \pi_f(n) ,
\]  

(11)

\[
(N_k - n_k)v_k(n) \pi_f(n) = (n_k + 1) \mu_k \pi_f(n_k^+) .
\]  

(12)

The PFS that satisfies Eqs. (10)–(12) is:

\[
\pi_f(n) = G^{-1} \left[ \prod_{k=1}^{K} \frac{N_k}{n_k} p_{n_k}^{n} \right] ,
\]  

(13)

where:

\[
p_{n_k}^{n} = \begin{cases} 
\alpha_{k, \text{idle}}^{n} & \text{for } n_k \leq n_k \text{,max} , \\
\alpha_{k, \text{idle}}^{n_k \text{,max}} \alpha_{k, \text{idle}}^{n-n_k \text{,max}} & \text{for } n_k > n_k \text{,max} .
\end{cases}
\]


Based on Eq. (13), we see that the term \( \kappa_{k,\text{idle}}^{n_k} e_{k,\text{idle}}^{n_k - n_{k,\text{max}}} \) expresses the application of the TCA policy in new service-class \( k \) calls when \( n_k > n_{k,\text{max}} \).

### 3.2. Convolution Algorithm

Based on (13), we can modify the convolution algorithm from Section 2 as follows:

**Step 1.** Assuming that the link of capacity \( C \) services only service-class \( k \) \((k = 1, \ldots, K)\) calls, determine the occupancy distribution \( q_{f,k}(j) \), where \( j = n_k b_k \), as:

\[
q_{f,k}(j) = \begin{cases} 
q_{f,k}(0) \kappa_{k,\text{idle}}^{n_k} e_{k,\text{idle}}^{n_k - n_{k,\text{max}}} & \text{for } n_k \leq n_{k,\text{max}} \\
q_{f,k}(0) \kappa_{k,\text{idle}}^{n_{k,\text{max}}} e_{k,\text{idle}}^{n_{k,\text{max}} - n_{k,\text{max}}} & \text{for } n_k > n_{k,\text{max}} 
\end{cases}
\]  
(14)

**Step 2.** Based on a successive convolution of all service classes excluding service-class \( k \), determine the occupancy distribution \( Q_{f,(-k)} \):

\[
Q_{f,(-k)} = q_{f,1} \cdot \cdots \cdot q_{f,k-1} \cdot \cdots \cdot q_{f,k+1} \cdot \cdots \cdot q_{f,K}.
\]  
(15)

**Step 3.** Compute the convolution operation of \( Q_{f,(-k)} \) and \( q_{f,k} \) as:

\[
Q_{f,(-k)} q_{f,k} = \begin{bmatrix}
Q_{f,(-k)} q_{f,k}(0) \\
\vdots \\
Q_{f,(-k)} q_{f,K}(C-x) 
\end{bmatrix} \cdot 
\begin{bmatrix}
q_{f,1}(x) \\
\vdots \\
q_{f,K}(x)
\end{bmatrix}.
\]  
(16)

Via Eq. (16), we can determine the link occupancy distribution \( q_f(j) \), according to:

\[
q_f(0) = \frac{Q_{f,(-k)} q_{f,k}(0)}{G}, \quad q_f(j) = \frac{\sum_{x=0}^{j} Q_{f,(-k)}(x) q_{f,k}(j-x)}{G}, \quad j = 1, \ldots, C.
\]  
(17)

Based on Eq. (17), we can determine the time congestion probabilities of service-class \( k \) calls (new or handover) as well as the link utilization (in b.u.) via Eqs. (7), (8) and (9), respectively, where \( q_f(j) \) and \( Q_{f,(-k)} \) should be replaced by \( q_f(j) \) and \( Q_{f,(-k)} \). To calculate call congestion probabilities for service-class \( k \) calls, we should run the convolution algorithm for \( N_k - 1 \) sources.

### 4. Evaluation

In this section, we present analytical CBP results for an application example. Due to the fact that the TCA and the f-TCA loss models have a PFS, we do not present simulation results.

Consider a link of a fixed capacity of \( C = 80 \) b.u. that accommodates Poisson arriving calls of \( K = 3 \) service-classes, with the traffic characteristics shown in Table 1. Based on Table 1, the average traffic-load offered per b.u. equals \( \frac{a_1 + a_2 + a_3}{C} = 0.16875 \).

<table>
<thead>
<tr>
<th>Service class</th>
<th>Traffic load (Erl)</th>
<th>Bandwidth per call (b.u.)</th>
<th>Threshold (no. of calls)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st</td>
<td>( a_1 = 9.0 )</td>
<td>( b_1 = 1 )</td>
<td>( n_{1,\text{max}} = 48 )</td>
</tr>
<tr>
<td>2nd</td>
<td>( a_2 = 3.0 )</td>
<td>( b_2 = 5 )</td>
<td>( n_{2,\text{max}} = 13 )</td>
</tr>
<tr>
<td>3rd</td>
<td>( a_3 = 1.5 )</td>
<td>( b_3 = 12 )</td>
<td>( n_{3,\text{max}} = 4 )</td>
</tr>
</tbody>
</table>

The provided analytical CBP and link utilization results for the TCA model and the TH model of [26], are based on two different values of \( n_{3,\text{max}} = 4 \) and 3 calls. As a reference, the analytical CBP and link utilization results for the CS model of [48] also are given. Regarding the f-TCA model, we provide analytical time congestion probabilities results when \( n_{3,\text{max}} = 4 \) calls and consider two different cases for the traffic sources cases: (1) \( N_1 = N_2 = N_3 = 100 \) sources and (2) \( N_1 = N_2 = N_3 = 200 \) sources. In the \( x \)-axis of Figs. 2–5, the offered traffic-load of the 1st, 2nd and 3rd service-class increases in steps of 1.0, 0.5 and 0.25 Erl, respectively. So, point 1 refers to: \( (\alpha_1, \alpha_2, \alpha_3) = (9.0, 3.0, 1.5) \), while point 7 to: \( (\alpha_1, \alpha_2, \alpha_3) = (15.0, 6.0, 3.0) \). The values of the offered traffic-load of service-class \( k \) handover calls, required in the case of the TCA model, are given by the expression \( \alpha_{kh} = \frac{a_k}{2} \). In the \( x \)-axis of Figs. 6–8, which refer to the

![Fig. 2. CBP – first service-class.](image-url)
Performance Evaluation of the Threshold Call Admission Policy in Multi-rate Loss Systems

\[ a_{k,\text{idle}} = \frac{\alpha_k}{N_k} \]

where \( \alpha_k \), idle = \( \alpha_k \), idle.

Based on Figs. 2–4, we observe that:

- The TCA model affects the CBP of all service-classes. Decreasing the value of \( n_{k,\text{max}} \) from 4 to 3 calls, results in a decrease in the CBP of 1st and 2nd service-class calls, and an increase in the CBP of 3rd service-class calls. This result is expected, since fewer in-service 3rd service-class calls are permitted in the system and, therefore, the available link bandwidth may be occupied by more calls of the other service-classes.

- The existing policies (CS and TH) cannot approximate the CBP results obtained via the TCA model, unless \( n_{k,\text{max}} \) becomes quite high. In that case, the TCA policy does not actually affect CBP and behaves similarly to the CS policy.

Based on Fig. 5, we observe that the CS policy achieves a higher link utilization compared to the TCA and the TH policies. The TCA policy achieves a slightly better link utilization compared to the corresponding TH policy due to the fact that it is applied only to new calls of a given service-class.

In Figs. 6–8, we present analytical time congestion probabilities for the proposed models (TCA and f-TCA) and for each service-class, respectively. We observe that the f-TCA model results in lower time congestion probabilities compared to the TCA model, especially when the number of traffic sources is low (\( N_1 = N_2 = N_3 = 100 \) sources). As the number of traffic sources increases, the results of the f-TCA model become similar to those of the TCA model.

To summarize the results, the following facts need to be emphasized:

- The TCA policy is applied only to new calls of a given service-class, while the TH policy is applied to both new and handover calls. On the other hand, the CS policy does not introduce restrictions to calls (apart from the link capacity). In that sense, the TCA policy cannot be approximated by the TH or the CS.

Based on Fig. 5, we observe that the CS policy achieves a higher link utilization compared to the TCA and the TH policies. The TCA policy achieves a slightly better link utilization compared to the corresponding TH policy due to the fact that it is applied only to new calls of a given service-class.

In Figs. 6–8, we present analytical time congestion probabilities for the proposed models (TCA and f-TCA) and for each service-class, respectively. We observe that the f-TCA model results in lower time congestion probabilities compared to the TCA model, especially when the number of traffic sources is low (\( N_1 = N_2 = N_3 = 100 \) sources). As the number of traffic sources increases, the results of the f-TCA model become similar to those of the TCA model.

To summarize the results, the following facts need to be emphasized:

- The TCA policy is applied only to new calls of a given service-class, while the TH policy is applied to both new and handover calls. On the other hand, the CS policy does not introduce restrictions to calls (apart from the link capacity). In that sense, the TCA policy cannot be approximated by the TH or the CS.

Based on Fig. 5, we observe that the CS policy achieves a higher link utilization compared to the TCA and the TH policies. The TCA policy achieves a slightly better link utilization compared to the corresponding TH policy due to the fact that it is applied only to new calls of a given service-class.

In Figs. 6–8, we present analytical time congestion probabilities for the proposed models (TCA and f-TCA) and for each service-class, respectively. We observe that the f-TCA model results in lower time congestion probabilities compared to the TCA model, especially when the number of traffic sources is low (\( N_1 = N_2 = N_3 = 100 \) sources). As the number of traffic sources increases, the results of the f-TCA model become similar to those of the TCA model.

To summarize the results, the following facts need to be emphasized:

- The TCA policy is applied only to new calls of a given service-class, while the TH policy is applied to both new and handover calls. On the other hand, the CS policy does not introduce restrictions to calls (apart from the link capacity). In that sense, the TCA policy cannot be approximated by the TH or the CS.

Based on Fig. 5, we observe that the CS policy achieves a higher link utilization compared to the TCA and the TH policies. The TCA policy achieves a slightly better link utilization compared to the corresponding TH policy due to the fact that it is applied only to new calls of a given service-class.

In Figs. 6–8, we present analytical time congestion probabilities for the proposed models (TCA and f-TCA) and for each service-class, respectively. We observe that the f-TCA model results in lower time congestion probabilities compared to the TCA model, especially when the number of traffic sources is low (\( N_1 = N_2 = N_3 = 100 \) sources). As the number of traffic sources increases, the results of the f-TCA model become similar to those of the TCA model.

To summarize the results, the following facts need to be emphasized:

- The TCA policy is applied only to new calls of a given service-class, while the TH policy is applied to both new and handover calls. On the other hand, the CS policy does not introduce restrictions to calls (apart from the link capacity). In that sense, the TCA policy cannot be approximated by the TH or the CS.
policy (see Figs. 2–4) unless \( n_{1, \text{max}} \) becomes quite high. In that case the threshold does not affect the call admission control and the TCA policy behaves as the CS policy.

The consideration of quasi-random traffic results in lower congestion probabilities compared to Poisson traffic, since the number of sources that generate traffic is limited (see Figs. 6–8).

5. Conclusion

We analyze, with the aid of reversible Markov chains, a multi-rate loss system of a single link that services random or quasi-random traffic under the TCA policy. The latter may be adopted when handover calls should be prioritized over new calls of a particular service-class. The TCA policy leads to a PFS for the steady state distribution. Based on the PFS, we can determine time and call congestion probabilities, as well as link utilization via convolution algorithms. As a future work, we intend to study the TCA policy in a multi-rate loss system that carries Pascal traffic. The latter is used to describe incoming traffic that is more bursty than classic random (Poisson) traffic.

Appendix

Tutorial

We consider calls of two different service-classes, accommodated in a link with a fixed capacity of \( C = 4 \) b.u. under the TCA policy. The arriving calls of both service-classes follow a quasi-random process with the following arrival rates: \( v_1 = 0.1 \), \( v_{1h} = 0.05 \), \( v_2 = 0.05 \) and \( v_{2h} = 0.025 \). Also let the number of traffic sources be \( N_1 = N_2 = 10 \). As far
as the bandwidth and the service time requirements of both service-classes are concerned, we have: \( b_1 = 1, b_2 = 2, \mu_1 = \mu_2 = 1.0 \). The corresponding arrival rates in the case of the Poisson process are: \( \lambda_1 = 1.0, \lambda_{1b} = 0.5, \lambda_2 = 0.5, \lambda_{2b} = 0.25 \).

Regarding the TCA policy, let \( n_{1,\text{max}} = n_{2,\text{max}} = 1 \). The state space \( \Omega \) of this example consists of 9 states presented in Table 2, in the form of \( n = (n_1, n_2) \), together with the corresponding values of \( j \) and the blocking states of both service-classes. More precisely, the symbol \( "\#" \) expresses a blocking event due to unavailability of b.u., while the symbol \( "\ast" \) expresses a blocking event due to the TCA policy (applied to new calls only).

In what follows, we present congestion probability calculations by applying the convolution algorithm from Section 3. In the parenthesis, we show the corresponding results assuming Poisson traffic.

**Step 1:**

- Determine \( q_{f1}(j), j = 1, \ldots, C \).

Let \( q_{f1}(0) = 1 \) and determine \( q_{f1}(j) \)'s based on \( q_{f1}(0) \):

\[
\begin{align*}
    j &= 1 \rightarrow q_{f1}(1) = q_{f1}(0) \times \binom{10}{1} \alpha_{1,\text{idle}} \\
    &\implies q_{f1}(1) = 1, \quad (q_{1}(1) = 1),
\end{align*}
\]

\[
\begin{align*}
    j &= 2 \rightarrow q_{f1}(2) = q_{f1}(0) \times \binom{10}{2} \alpha_{1,\text{idle}} \times \alpha_{1b,\text{idle}} \\
    &\implies q_{f1}(2) = 0.225, \quad (q_{1}(2) = 0.25),
\end{align*}
\]

\[
\begin{align*}
    j &= 3 \rightarrow q_{f1}(3) = q_{f1}(0) \times \binom{10}{3} \alpha_{1,\text{idle}} \times \alpha_{1b,\text{idle}} \\
    &\implies q_{f1}(3) = 0.03, \quad (q_{1}(3) = 0.25/6),
\end{align*}
\]

\[
\begin{align*}
    j &= 4 \rightarrow q_{f1}(4) = q_{f1}(0) \times \binom{10}{4} \alpha_{1,\text{idle}} \times \alpha_{1b,\text{idle}}^2 \\
    &\implies q_{f1}(4) = 0.002625, \quad (q_{1}(4) = 0.125/24).
\end{align*}
\]

Since the normalization constant is \( G = 2.257625 \), the normalized values of \( q_{f1}(j) \)'s are:

\[
\begin{align*}
    q_{f1}(0) &= 0.44294335, \\
    q_{f1}(1) &= 0.09966225, \\
    q_{f1}(2) &= 0.0132883, \\
    q_{f1}(3) &= 0.001162726.
\end{align*}
\]

The corresponding values assuming Poisson traffic (and since \( G = 2.296875 \)) are as follows:

\[
\begin{align*}
    q_1(0) &= 0.435374, \\
    q_1(2) &= 0.108844, \\
    q_1(3) &= 0.0181405, \\
    q_1(4) &= 0.0022675.
\end{align*}
\]

**Step 2:**

Determine the values of \( Q_{f1}(-i) \). Since we have two service-classes this step is not required.

**Step 3:**

Determine \( q_{f2}(j) \)'s and consequently the time congestion probabilities (or the CBP in the case of Poisson traffic) of all service-classes.

\[
\begin{align*}
    j &= 0 \rightarrow q_{f2}(0) = q_{f1}(0)q_{f2}(0) = 0.2846222, \\
    (q(0) = 0.2786394),
\end{align*}
\]

\[
\begin{align*}
    j &= 1 \rightarrow q_{f2}(1) = \sum_{x=0}^{1} q_{f1}(x)q_{f2}(1-x) = 0.2846222, \\
    (q(1) = 0.2786394),
\end{align*}
\]

\[
\begin{align*}
    j &= 2 \rightarrow q_{f2}(2) = \sum_{x=0}^{2} q_{f1}(x)q_{f2}(2-x) = 0.2063511, \\
    (q(2) = 0.2089795),
\end{align*}
\]

\[
\begin{align*}
    j &= 3 \rightarrow q_{f2}(3) = \sum_{x=0}^{3} q_{f1}(x)q_{f2}(3-x) = 0.1508498, \\
    (q(3) = 0.1509297),
\end{align*}
\]

\[
\begin{align*}
    j &= 4 \rightarrow q_{f2}(4) = \sum_{x=0}^{4} q_{f1}(x)q_{f2}(4-x) = 0.0487771, \\
    (q(4) = 0.0536961).
\end{align*}
\]

The normalized values of \( q_{f2}(j) \)'s are (since \( G = 0.975222 \)):

\[
\begin{align*}
    q_{f2}(0) &= 0.291854, \\
    q_{f2}(2) &= 0.211594, \\
    q_{f2}(3) &= 0.154682, \\
    q_{f2}(4) &= 0.050016.
\end{align*}
\]

---

### Table 2

<table>
<thead>
<tr>
<th>( n_1 )</th>
<th>( n_2 )</th>
<th>( j )</th>
<th>( B_{1n} )</th>
<th>( B_{1h} )</th>
<th>( B_{2n} )</th>
<th>( B_{2h} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>#</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>2</td>
<td>4</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>1</td>
<td>#</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>3</td>
<td>#</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>2</td>
<td>#</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>4</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>3</td>
<td>#</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>4</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
</tbody>
</table>

---

Performance Evaluation of the Threshold Call Admission Policy in Multi-rate Loss Systems
The corresponding normalized values of $q(j)$’s in the case of Poisson traffic are (since $G = 0.970884$):

\[
q(0) = q(1) = 0.286996, \\
q(2) = 0.215246, \\
q(3) = 0.155456, \\
q(4) = 0.055306.
\]

The time congestion probabilities are:

\[
B_{1a} = \sum_{j=-b_1+1}^{C} q(j) + \frac{C-b_1}{C} \sum_{j=-b_1+1}^{C} q(j)(C-b_1-y) \Rightarrow \\
B_{1a} = q(4) + q(3)(G + q(3)) = 0.975222
\]

\[
B_{1b} = \sum_{j=-b_1+1}^{C} q(j) = 0.56222
\]

\[
B_{2a} = \sum_{j=-b_2+1}^{C} q(j) + \frac{C-b_2}{C} \sum_{j=-b_2+1}^{C} q(j)(C-b_2-y) \Rightarrow \\
B_{2a} = q(4) + q(3)(G + q(3)) = 0.975222
\]

\[
B_{2b} = \sum_{j=-b_2+1}^{C} q(j) = 0.350626
\]

The corresponding values in the case of Poisson traffic are slightly higher: $B_{1a} = 0.56951, B_{1b} = 0.055306, B_{2a} = 0.35426, B_{2b} = 0.21076$.

### References


Performance Evaluation of the Threshold Call Admission Policy in Multi-rate Loss Systems

I. Moscholios, V. Vassilakis, and M. Logothetis, “Call blocking probabilities for Poisson traffic under the multiple fractional channel reservation policy”, in Proc. 10th Int. Symp. on Comm. Syst., Netw. and Digi. Sig. Process. CNSDSP 2016, Prague, Czech Republic, 2016 (DOI: 10.1109/CNSDSP2016.7573910).


Ioannis D. Moscholios received his Dipl.Eng. in Electrical and Computer Engineering from the University of Patras, Patras, Greece, in 1999, M.Sc. in Spacecraft Technology and Satellite Communications from the University College London, UK, in 2000, and Ph.D. in Electrical and Computer Engineering from the University of Patras in 2005. From 2005 to 2009, he was a Research Associate at the Wire Communications Laboratory, Department of Electrical and Computer Engineering, University of Patras. From 2009 to 2013, he was a Lecturer at the Department of Telecommunications Science and Technology, University of the Peloponnese, Tripolis, Greece. From 2013 to 2018, he was an Assistant Professor at the Department of Informatics and Telecommunications, University of the Peloponnese, Tripolis, Greece. Currently, he is an Associate Professor at the Department of Informatics and Telecommunications, University of the Peloponnese, Tripolis, Greece. His research interests include teletraffic engineering, as well as simulation and performance analysis of communication networks.
Ioannis Moscholios, Iskanter-Alexandros Chousainov, Panagiotis Panagoulias, Panagiotis Sarigiannidis, and Michael Logothetis

Iskanter-Alexandros Chousainov received his B.Sc. degree from the Department of Mathematics of the University of the Aegean in 2002, the M.Sc. degree in Information Systems from Hellenic Open University in 2017. He is currently a Ph.D. student at the Department of Informatics and Telecommunications, University of the Peloponnese, Tripolis, Greece. His research interests include teletraffic engineering and performance evaluation of 5G networks.

Panagiotis I. Panagoulias received his Dipl.Eng. in Civil Engineering from the National Technical University of Athens, Greece, in 2006, M.Sc. in Earthquake Engineering and Seismic Resistant Structures from the Hellenic Open University, Patras, Greece, in 2011 and M.Sc. in Information Systems from the Hellenic Open University as well, in 2016. Currently, he is a Ph.D. student at the Department of Informatics and Telecommunications, University of the Peloponnese, Tripolis, Greece. His current research interests include teletraffic engineering, as well as simulation and performance analysis of 4G/5G networks.

Panagiotis G. Sarigiannidis has been an Assistant Professor at the Department of Electrical and Computer Engineering, University of Western Macedonia, Kozani, Greece, since 2016. He received his B.Sc. and Ph.D. degrees in Computer Science from the Aristotle University of Thessaloniki, Thessaloniki, Greece, in 2001 and 2007, respectively. His research interests include telecommunication networks, Internet of Things and network security. He is an IEEE member and participates in the Editorial Boards of various journals.

Michael D. Logothetis received his Dipl.Eng. and Ph.D. degrees in Electrical Engineering from the University of Patras, Patra, Greece, in 1981 and 1990, respectively. From 1982 to 1990, he was a Teaching and Research Assistant with the Laboratory of Wire Communications, University of Patras. He has participated in many national and EU research programs, dealing with telecommunication networks and office automation. From 1991 to 1992, he was a Research Associate with NTT’s Telecommunication Networks Laboratories, Tokyo, Japan. He was a Lecturer at the Electrical and Computer Engineering Department, University of Patras, and in 2009 was elected a (Full) Professor at the same department. His current research interests include teletraffic theory and engineering, traffic/network control, simulation and performance optimization of telecommunications networks.

E-mail: idm@uop.gr
Department of Informatics and Telecommunications
University of the Peloponnese
221 00 Tripolis, Greece

https://orcid.org/0000-0003-3656-277X

E-mail: ichousain@uop.gr
Department of Informatics and Telecommunications
University of the Peloponnese
221 00 Tripolis, Greece

https://orcid.org/0000-0003-0244-3385

E-mail: panagoulias@uop.gr
Department of Informatics and Telecommunications
University of the Peloponnese
221 00 Tripolis, Greece

https://orcid.org/0000-0001-6041-0355
E-mail: psarigiannidis@uowm.gr
Department of Electrical and Computer Engineering
University of West Macedonia
Kozani, Greece

https://orcid.org/0000-0001-6315-5382
E-mail: mlogo@upatras.gr
WCL, Department of Electrical and Computer Engineering
University of Patras
265 04 Patra, Greece