An Exact Algorithm for Calculating Blocking Probabilities in Multicast Networks

Eeva Nyberg, Jorma Virtamo, and Samuli Aalto
Laboratory of Telecommunications Technology
Helsinki University of Technology
P.O.Box 3000, FIN-02015 HUT, Finland

5th October 1999

Abstract The paper deals with tree-structured point-to-multipoint networks, where users from infinite user populations at the leaf nodes subscribe to a variety of channels, offered by one source. The users joining the network form dynamic multicast connections that share the network resources. An exact algorithm for calculating end-to-end call blocking probabilities for dynamic connections is devised for this multicast model. The algorithm is based on the well-known algorithm for calculating blocking probabilities in hierarchical multiservice access networks, where link occupancy distributions are alternately convolved and truncated. The resource sharing of multicast connections requires the modification of the algorithm by using a new type of convolution, the OR-convolution. The exact algorithm for end-to-end call blocking probabilities enables us to study the accuracy of earlier results based on Reduced Load Approximation. The model is further extended to include background traffic, allowing the analysis of networks carrying mixed traffic e.g. multicast and unicast traffic.

Keywords Multicast, blocking, network, OR-convolution

1 Introduction

A multicast transmission originates at a source and, opposed to a unicast transmission, is replicated at various network nodes to form a tree-and-branch structure. The transmission reaches many different end-users without a separate transmission required for each user. A multicast connection has therefore a bandwidth saving nature. Blocking occurs in a network when, due to limited capacity, at least one link on the route is not able to admit a new call. Traditional mathematical models to calculate blocking probabilities in tree-structured networks exist for unicast traffic. Due to different resource usage, these models cannot directly be used for multicast networks where requests from different users arrive dynamically. Only recently, have mathematical models to calculate blocking probabilities in multicast networks been studied.

The research has mainly been focused on blocking probabilities in multicast capable switches. Kim [7] studied blocking probabilities in a multirate multicast switch. Three stage switches are studied by Yang and Wang [12] and Listanti and Veltri [8]. Stasiak and Zwierzykowski [11] studied blocking in an ATM node with multicast switching nodes carrying different multi-rate traffic (unicast and multicast), using Kaufman-Roberts recursion and Reduced Load Approximation. Admission control algorithms are studied in [10]. Almeroth and Ammar [1] investigated Multicast group behavior in the Multicast Backbone (MBone). From this data, they concluded that interarrival times are exponentially distributed while membership duration times are exponentially distributed for small networks and Zipf distributed for larger networks. The study of intersession data suggests that simultaneous sessions, where a user subscribes to more than one channel, occur, but not frequently.

Chan and Geraniotis [2] have studied blocking due to finite capacity in network links. They formulated a closed form expression for time blocking probabilities in a network transmitting layered video signals. The model is a multipoint-to-multipoint model. The network consists of several video sources, where each source node can also act as a receiver. The video signals are coded into different layers defining the quality of the video signal received by the user. The traffic class is defined by the triplet: physical path, source node, and class of video quality. The behavior of each user is modeled as a two state Markov chain, with unique transition rates defined for each traffic class triplet.

Karvo et al. [4] and [5] studied blocking in a point-to-multipoint network with only one source node. The source is called the service center and it can offer a variety of channels, e.g. TV-channels. The users subscribing to the network may join and leave the channel at any time. The behavior of the user population defines the state probabilities at the links of the tree-structured network. The user population is assumed infinite and the requests to join the network arrive as from a Poisson process. The model studied in [4] considered the model in a simplified case where all but one link in a network have infinite capacity. They derived an exact algorithm to calculate both the channel and call blocking probability in this simplified case. Extending the model to the whole network was done only approximately in [5], where end-to-end blocking probabilities are estimated using the Reduced Load Approximation (RLA) approach.

The present study extends the single link case discussed in [4] and [5] to a mathematical model for a multicast network with any number of finite links. The network model remains as a point-to-multipoint model, allowing the formulation of an exact algorithm, which in the case of multiple sources would be too complex. The exact algorithm is based on the well-known algorithm for calculating blocking probabilities in hierarchical multiservice access networks, where link occupancy distributions are alternately convolved and truncated. The resource sharing of multicast connections requires the modification of the algorithm by using a new type of convolution, the OR-convolution. The exact algorithm for end-to-end call blocking probabilities enables us to study the accuracy of earlier results based on RLA. The model is further extended to include background traffic, allowing the analysis of networks carrying mixed traffic e.g. multicast and unicast traffic.

This paper continues by presenting, in section 2, the single link model of Karvo et al. Some of the expressions given in the paper [4] are formulated in a new way. The main results of this paper are presented in section 3, which is divided into five parts. First, the notation is presented. Secondly, the model for a network with infinite link capacities is presented and thirdly, the OR-convolution used to convolve multicast state distributions in tree networks is introduced. Then, an expression for the call blocking probability in a network with any number of finite links is given and finally, the algorithm to calculate the call blocking probability is introduced. In section 4, comparisons between the exact solution and Reduced Load Approximation are carried through. The network model is extended to include non-multicast traffic as background traffic in section 5. The paper is concluded in section 6.

2 Single link model

In this section, we review the point-to-multipoint model for a dynamic multicast network with an infinite user population presented in [4]. A single source offers a variety of channels, belonging to the set $\mathcal{I}$. Subscriptions to channel $i \in \mathcal{I}$ arrive according to a Poisson process with intensity λ_i. The channel i is chosen independently according to a preference probability α_i. The offered traffic intensity for a multicast channel i is then $a_i = \lambda_i/\mu_i = \alpha_i\lambda/\mu_i$, where $1/\mu_i$ is the average holding time of channel i and is generally distributed. It is shown in [4] that in a multicast network with all links having infinite capacity, the distribution of the number of users simultaneously connected to channel i is the distribution of the number of customers in an $M/G/\infty$ queue. The probability of having channel i on is the probability that at least one user subscribes to channel i and is therefore,

$$p_i = 1 - e^{-a_i}. \tag{1}$$

When the link capacity is restricted, blocking occurs. To calculate these blocking probabilities, Karvo et al. [4] considered the case where only one link has finite capacity and all other links in the network have infinite capacity. The blocking probability is divided into three types: channel blocking B_i^c, call blocking b_i^c, and time blocking b_i^t. The channel blocking probability is the probability that an attempt to turn channel i on fails due to the capacity restriction of the link. Channel blocking can therefore occur only, if the channel is idle. Call blocking occurs when a user is not able to subscribe to channel i. Call and channel blocking are different as a subscription to a channel is always accepted if the channel is already active. Time blocking probability of channel i is defined as the probability that at an arbitrary instant the link is not able to accept a request for channel i, i.e. the channel is off and at least $C - d_i + 1$ capacity units are occupied. Here d_i denotes the capacity requirement of channel i.

Karvo et al. deduce that when the capacity of the link is restricted, the mean length of an off-period of channel i is

$$T_{i,\text{off}} = \frac{1}{\lambda_i(1 - B_i^c)}$$

and the average time the channel is on is the same as the average on-period for a link with infinite capacity

$$T_{i,\text{on}} = \frac{p_i}{(1-p_i)\lambda_i}.$$

Karvo et al. give an expression for the call blocking in a finite link, which by inserting equation (1) into the original expression given in [4] can be rewritten as

$$b_i^c = \frac{(1-p_i)B_i^c}{1-p_iB_i^c}. \tag{2}$$

The equation confirms that channel blocking probability equals call blocking probability upon condition that the channel is off, as in the finite system the probability that the channel is off is

$$\frac{T_{i,\text{off}}}{T_{i,\text{on}}+T_{i,\text{off}}} = \frac{\frac{1}{\lambda_i(1-B_i^c)}}{\frac{p_i}{(1-p_i)\lambda_i}+\frac{1}{\lambda_i(1-B_i^c)}} = \frac{1-p_i}{1-p_iB_i^c}.$$

Karvo et al. [4] proceed in deriving an expression for the channel blocking probability B_i^c of channel i. They observe that channel blocking in the one finite link case is equal to the call blocking in a certain generalized Engset system, that is a $M/G/C/C/I$ queue. Thus, the channel blocking probability B_i^c is equal to the time blocking probability b_i^t in a system where channel i is removed,

$$B_i^c = \frac{\sum_{j=C-d_i+1}^{C} {\pi_j}^{(i)}}{\sum_{j=0}^{C} {\pi_j}^{(i)}}, \tag{3}$$

where $\pi_j^{(i)}$ is the link occupancy distribution of a system with channel i removed.

An alternative way of calculating the call blocking probability is calculating the time blocking probability of the whole system,

$$b_i^t = \frac{\sum_{j=C-d_i+1}^{C} {\pi_j}^{(x_i=0)}}{\sum_{j=0}^{C} \pi_j}, \tag{4}$$

where π_j is the link occupancy distribution for an infinite system and $\pi_j^{(x_i=0)}$ is the link occupancy distribution restricted to the states with channel i off,

$$\pi_j^{(x_i=0)} = \sum_{\mathbf{x}\cdot\mathbf{d}=j, x_i=0} \pi(\mathbf{x}) = (1-p_i)\pi_j^{(i)}, \tag{5}$$

where $\mathbf{d} = (d_i, i \in \mathcal{I})$. The last expression stems from the product form of the link occupancy distribution π_j. Similarly, the link occupancy distribution π_j in terms of the link occupancy distribution $\pi_j^{(i)}$ is

$$\pi_j = (1-p_i)\pi_j^{(i)} + p_i\pi_{j-d_i}^{(i)}. \tag{6}$$

Because of Poisson arrivals, the time blocking and the call blocking probabilities for the whole system are equal. This can also be verified by investigating equations (2) through

(4) and rewriting the expressions for $\pi_j^{(x_i=0)}$ and π_j with the help of equations (5) and (6),

$$\begin{aligned} b_i^t &= \frac{\sum_{j=C-d_i+1}^{C} {\pi_j}^{(x_i=0)}}{\sum_{j=0}^{C} \pi_j} = \frac{(1-p_i)\sum_{j=C-d_i+1}^{C} {\pi_j}^{(i)}}{(1-p_i)\sum_{j=0}^{C} \pi_j^{(i)} + p_i \sum_{j=0}^{C-d_i} \pi_j^{(i)}} \\ &= \frac{(1-p_i)\sum_{j=C-d_i+1}^{C} {\pi_j}^{(i)}}{(1-p_i)\sum_{j=0}^{C} \pi_j^{(i)} + p_i(\sum_{j=0}^{C} \pi_j^{(i)} - \sum_{j=C-d_i+1}^{C} \pi_j^{(i)})} \\ &= \frac{(1-p_i)B_i^c}{1-p_i+p_i(1-B_i^c)} = \frac{(1-p_i)B_i^c}{1-p_iB_i^c} = b_i^c. \end{aligned}$$

Equation (4) is computationally more convenient when calculating blocking probabilities for different traffic classes, as the denominator does not have to be calculated anew for each traffic class.

3 Network model

This section is divided into five parts: the notation used, the model for the network with infinite link capacities, the OR-convolution, the expression for blocking probabilities, and the algorithm to calculate call blocking probabilities in a network.

3.1 Notation

The notation used throughout this paper is as follows. The set of all links is denoted by $\mathcal{J}$. Let $\mathcal{U} \subset \mathcal{J}$ denote the set of leaf links. The leaf link and user population behind the leaf link is denoted by $u \in \mathcal{U} = \{1, ..., U\}$. The set of links on the route from leaf link u to the source is denoted by $\mathcal{R}_u$. The set of links downstream link $j \in \mathcal{J}$ including link j is denoted by $\mathcal{M}_j$, while the set of downstream links terminating at link $j \in \mathcal{J}$ are denoted by $\mathcal{N}_j$. The set of user populations downstream link j is denoted by $\mathcal{U}_j$. The set of channels offered by the source is $\mathcal{I}$, with channel $i = 1, .., I$. Let $\mathbf{d} = \{d_i; i \in \mathcal{I}\}$, where d_i is the capacity requirement of channel i. Here we assume that the capacity requirements depend only on the channel, but link dependencies could also be included into the model. The capacity of the link j is denoted by C_j. The different sets are shown in figure 1.

3.2 Network with infinite link capacities

We first consider a network with all links having infinite capacity. Subscriptions to channel i behind leaf link u arrive from an infinite user population as from a Poisson process with intensity $\lambda_{u,i} = \alpha_i \lambda_u$, where α_i is generated from a preference distribution for channel i and λ_u is the arrival intensity for user population u. The channel holding time is assumed to be generally distributed with mean $1/\mu_i$. In addition we denote the traffic intensity

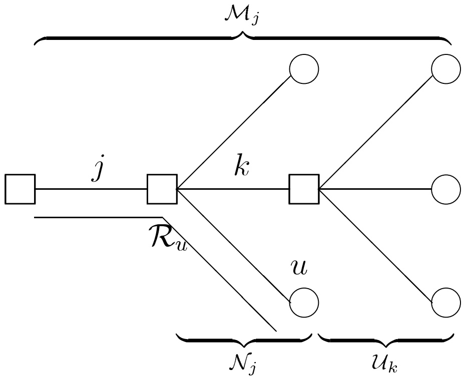

Figure 1: An example network to show the notation used.

$a_{u,i} = \alpha_i \lambda_u / \mu_i$. Let the pair $(u, i) \in \mathcal{U} \times \mathcal{I}$ denote a traffic class also called a connection. The connection state, which may be off or on, is denoted by $X_{u,i} \in \{0, 1\}$. The state probability for a connection, according to the $M/G/\infty$ model, is

$$\pi_{u,i}(x_{u,i}) = P(X_{u,i} = x_{u,i}) = (p_{u,i})^{x_{u,i}}(1 - p_{u,i})^{1-x_{u,i}},$$

where $p_{u,i} = 1 - e^{-a_{u,i}}$.

In the infinite link capacity case, all connections are independent of each other. For leaf link u, the state probability has a product form and is

$$\pi_u(\mathbf{x}_u) = P(\mathbf{X}_u = \mathbf{x}_u) = \prod_{i \in \mathcal{I}} \pi_{u,i}(x_{u,i}),$$

where $\mathbf{X}_u = (X_{u,i}; i \in \mathcal{I}) \in \mathcal{S}$ is the state vector for the leaf link, and $\mathcal{S} = \{0, 1\}^I$ denotes the link state space.

The leaf link states jointly define the network state $\mathbf{X}$,

$$\mathbf{X} = (\mathbf{X}_u; u \in \mathcal{U}) = (X_{u,i}; u \in \mathcal{U}, i \in \mathcal{I}) \in \Omega, \tag{7}$$

where $\Omega = \{0, 1\}^{U \times I}$ denotes the network state space. For the whole network, the state probability is

$$\pi(\mathbf{x}) = P(\mathbf{X} = \mathbf{x}) = \prod_{u \in \mathcal{U}} \pi_u(\mathbf{x}_u),$$

as each user population is independent of each other.

3.3 OR-convolution

The leaf link state distributions jointly define the network state distribution, as was shown in the previous section. In order to calculate the link state distributions in a tree-structured network a convolution operation is needed. The resource sharing characteristic of multicast traffic requires a new type of convolution, the OR-convolution. Consider two downstream links $s, t \in \mathcal{N}_v$ terminating at link v, where $s, t, v \in \mathcal{J}$. Channel i is idle in link v if it is idle in both links s and t and active in all other cases, which is equivalent to the binary OR-operation. In other words, for $\mathbf{y}_s, \mathbf{y}_t \in \mathcal{S}$

$$\mathbf{y}_v = \mathbf{y}_s \oplus \mathbf{y}_t \in \mathcal{S}, \tag{8}$$

where the vector operator $\oplus$ denotes the OR-operation taken componentwise. The OR-convolution, denoted by $\otimes$, is then the operation,

$$[f_s \otimes f_t](\mathbf{y}_v) = \sum_{\mathbf{y}_s \oplus \mathbf{y}_t = \mathbf{y}_v} f_s(\mathbf{y}_s) f_t(\mathbf{y}_t)$$

defined for any distributions f_s and f_t.

In a multicast link, the link state depends on the user states downstream the link. If a channel is idle in all links downstream link j it is off in link j and in all other cases the channel is active. The OR-operation on the network state gives the link state $\mathbf{Y}_j = (Y_{j,i}; i \in \mathcal{I}) \in \mathcal{S}, j \in \mathcal{J}$ as a function of the network state,

$$\mathbf{Y}_j = \mathbf{g}_j(\mathbf{X}) = \bigoplus_{k \in \mathcal{U}_j} \mathbf{X}_k.$$

Similarly, the OR-convolution on the network state distribution gives the link state distribution. Thus, the state probability, denoted by $\sigma_j(\mathbf{y}_j)$, for $\mathbf{y}_j \in \mathcal{S}$, is equal to

$$\sigma_j(\mathbf{y}_j) = P(\mathbf{Y}_j = \mathbf{y}_j) = [\bigotimes_{k \in \mathcal{U}_j} \pi_k](\mathbf{y}_j) = \begin{cases} \pi_j(\mathbf{y}_j) & \text{, if } j \in \mathcal{U} \\ [\bigotimes_{k \in \mathcal{N}_j} \sigma_k](\mathbf{y}_j) & \text{, otherwise .} \end{cases}$$

When $\mathbf{X} = \mathbf{x}$ the occupied capacity on the link j is $\mathbf{d} \cdot \mathbf{g}_j(\mathbf{x})$.

3.4 Blocking probabilities in a network with finite link capacities

When the capacities of one or more links in the network are finite, the state spaces defined above are truncated according to the capacity restrictions. The network state $\mathbf{X}$ defined in equation (7) is replaced by the truncated network state $\tilde{\mathbf{X}} \in \tilde{\Omega}$, where $\tilde{\Omega}$ denotes the truncated state space

$$\tilde{\Omega} = \{\mathbf{x} \in \Omega \mid \mathbf{d} \cdot \mathbf{g}_j(\mathbf{x}) \leq C_j, \forall j \in \mathcal{J}\}.$$

The insensitivity [6] and truncation principles [3] apply for this product form network, and for the truncated system the state probabilities of the network differ only by the normalization constant $G(\tilde{\Omega}) = \sum_{\mathbf{x} \in \tilde{\Omega}} \pi(\mathbf{x})$. The state probabilities of the truncated system are therefore

$$\tilde{\pi}(\mathbf{x}) = P(\tilde{\mathbf{X}} = \mathbf{x}) = \frac{\pi(\mathbf{x})}{G(\tilde{\Omega})}, \text{ for } \mathbf{x} \in \tilde{\Omega}.$$

When the capacity on the links is finite, blocking occurs. Due to Poisson arrivals, the call blocking probability is equal to the time blocking probability of the system. A call in traffic class (u, i) is blocked if there is not enough capacity in the network to set up the connection. Note that, once the channel is active on all links belonging to the route R_u of user population u, no extra capacity is required for a new connection. Let us define another truncated set $\tilde{\Omega}_{u,i} \subset \tilde{\Omega}$ with a tighter capacity restriction for links with channel i idle,

$$\tilde{\Omega}_{u,i} = \{\mathbf{x} \in \Omega \mid \mathbf{d} \cdot \big(\mathbf{g}_j(\mathbf{x}) \oplus (\mathbf{e}_i 1_{j\in\mathcal{R}_u})\big) \leq C_j, \forall j \in \mathcal{J}\},$$

where $\mathbf{e}_i$ is the I-dimensional vector consisting of only zeroes except for a one in the ith component and $1_{j\in\mathcal{R}_u}$ is the indicator function equal to one for $j \in \mathcal{R}_u$ and zero otherwise. This set defines the states where blocking does not occur when user u requests a connection to channel i. The call blocking probability b_i^c for traffic class (u, i) is thus,

$$b_{u,i}^c = 1 - P(\tilde{\mathbf{X}} \in \tilde{\Omega}_{u,i}) = 1 - \frac{G(\tilde{\Omega}_{u,i})}{G(\tilde{\Omega})}. \tag{9}$$

This approach requires calculating two sets of state probabilities: the set of non-blocking states appearing in the numerator and the set of allowed states appearing in the denominator of equation (9).

A multicast network is a tree-type network, and much of the theory in calculating blocking probabilities in hierarchical multiservice access networks [9] can be used to formulate the end-to-end blocking probability in a multicast network as well.

3.5 The algorithm

As in the case of access networks, the blocking probability can be calculated by recursively convolving the state probabilities of individual links from the leaf links to the origin link. At each step, the state probabilities are truncated according to the capacity restriction of the link.

In order to calculate the denominator of equation (9), let us define a new subset $\tilde{\mathcal{S}}_j$ of set $\mathcal{S}$,

$$\tilde{\mathcal{S}}_j = \{\mathbf{y} \in \mathcal{S} | \mathbf{d} \cdot \mathbf{y} \leq C_j\}, \text{ for } j \in \mathcal{J}.$$

The corresponding truncation operator acting on any distribution f is

$$T_j f(\mathbf{y}) = \begin{cases} f(\mathbf{y}) & , \text{ if } \mathbf{y} \in \tilde{\mathcal{S}}_j \\ 0 & , \text{ otherwise.} \end{cases} \tag{10}$$

Let

$$Q_j(\mathbf{y}_j) = P(\mathbf{Y}_j = \mathbf{y}_j; \mathbf{Y}_k \in \tilde{\mathcal{S}}_k, \forall k \in \mathcal{M}_j), \text{ for } \mathbf{y}_j \in \mathcal{S}. \tag{11}$$

It follows that the $Q_j(\mathbf{y})$'s can be calculated recursively,

$$Q_j(\mathbf{y}) = \begin{cases} T_j \pi_j(\mathbf{y}) & , \text{ if } j \in \mathcal{U} \\ T_j[\bigotimes_{k\in\mathcal{N}_j} Q_k](\mathbf{y}) & , \text{ otherwise.} \end{cases}$$

Note that, if the capacity constraint of link $j \in \mathcal{M}_j$ is relaxed, then the branches terminating at link j are independent, and the jointly requested channel state can be obtained by the OR-convolution. The effect of the finite capacity C_j of link j is then just the truncation of the distribution to the states for which the requested capacity is no more than C_j.

The state sum $G(\tilde{\Omega})$ needed to calculate the blocking probability in equation (9) is equal to

$$G(\tilde{\Omega}) = \sum_{\mathbf{y}\in\mathcal{S}} Q_J(\mathbf{y}),$$

where Q_J is the probability (11) related to the common link $j = J$.

Similarly for the numerator of equation (9), let $\tilde{\mathcal{S}}_j^{u,i} \subset \tilde{\mathcal{S}}_j$ be defined as the set of states on link j that do not prevent user u from connecting to multicast channel i. In other words

$$\tilde{\mathcal{S}}_j^{u,i} = \{\mathbf{y} \in \mathcal{S} \mid \mathbf{d} \cdot \big(\mathbf{y} \oplus (\mathbf{e}_i 1_{j\in\mathcal{R}_u})\big) \leq C_j\}, \text{ for } j \in \mathcal{J}, i \in \mathcal{I}.$$

The truncation operator is then

$$T_j^{u,i} f(\mathbf{y}) = \begin{cases} f(\mathbf{y}) & , \text{ if } \mathbf{y} \in \tilde{\mathcal{S}}_j^{u,i} \\ 0 & , \text{ otherwise.} \end{cases} \quad (12)$$

The non-blocking probability of link j is

$$Q_j^{u,i}(\mathbf{y}_j) = P(\mathbf{Y}_j = \mathbf{y}_j; \mathbf{Y}_k \in \tilde{\mathcal{S}}_j^{u,i}, \forall k \in \mathcal{M}_j), \text{ for } \mathbf{y}_j \in \mathcal{S}. \quad (13)$$

Similarly, as above, it follows that

$$Q_j^{u,i}(\mathbf{y}) = \begin{cases} T_j^{u,i} \pi_j(\mathbf{y}) & , \text{ if } j \in \mathcal{U} \\ T_j^{u,i}[\bigotimes_{k\in\mathcal{N}_j} Q_k^{u,i}](\mathbf{y}) & , \text{ otherwise.} \end{cases}$$

The state sum in the numerator of equation (9) is then

$$G(\tilde{\Omega}_{u,i}) = \sum_{\mathbf{y}\in\mathcal{S}} Q_J^{u,i}(\mathbf{y}),$$

where $Q_J^{u,i}$ is the probability (13) related to the common link $j = J$.

The blocking probability in equation (9) is therefore

$$b_{u,i}^c = 1 - \frac{\sum_{\mathbf{y}\in\mathcal{S}} Q_J^{u,i}(\mathbf{y})}{\sum_{\mathbf{y}\in\mathcal{S}} Q_J(\mathbf{y})}.$$

The single link model by Karvo et al. is a special case of the network model presented, hence the same results can be obtained using the network algorithm.

The complexity of the algorithm increases exponentially with the number of channels, as the number of states in the distributions to be convolved is 2^I. Therefore the use of RLA as a computationally simpler method is studied.

4 Comparisons between the exact model and RLA

The calculation of end-to-end blocking probabilities for multicast networks was done approximately in [5] using the RLA-algorithm. The fixed point equation used treats each link as an independent link with finite capacity. The fixed point equation is based on equation (4) derived for the single link case,

$$L_j^i = b_i^c(\mathbf{s}_j, \mathbf{d}, C_j), \text{ for } i \in \mathcal{I} \text{ and } j \in \mathcal{J}, \tag{14}$$

where $\mathbf{s}_j$ is the thinned traffic intensity vector $\mathbf{s}_j = (s_{j,i}, i \in \mathcal{I})$,

$$s_{j,i} = \sum_{u \in \mathcal{U}_j} a_{u,i} \prod_{k \in \mathcal{R}_u - \{j\}} (1 - L_k^i). \tag{15}$$

Writing equation (14) in matrix form gives a fixed point equation $T(\mathbf{L}) = \mathbf{L}$ for $\mathbf{L} = (L_j^i)$. For a multiservice network, the solution is not always unique see [9]. Solving the fixed point equation gives the call blocking probability for each individual link in the network. Under the independence assumption, the blocking of the traffic class (u, i) is

$$b_{u,i}^c = 1 - \prod_{k \in \mathcal{R}_u} (1 - L_k^i).$$

The exact algorithm derived in the previous section allows us to study the accuracy of RLA. To this end, we consider the example network depicted in figure 2. The figure

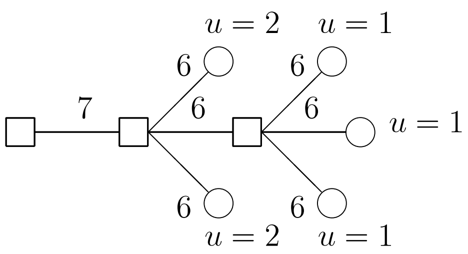

Figure 2: The example network used to compare the exact result with the result given by the RLA-algorithm.

shows the two possible routes in the network and the capacities of the links. The links are numbered in the following way. The leaf links of user one have $u, j = 1$, the leaf links of user two have $u, j = 2$, the middle link has $j = 3$, and the common link has $j = 4$. Comparisons were made between the exact solution and the RLA-algorithm. The number of channels offered is eight. Each channel requires one unit of capacity. The common link in the network has a capacity of seven units. All other links have a capacity of six units. The blocking probabilities are calculated for the least used channel using a truncated geometric distribution for the channel preference

$$\alpha_i = \frac{p(1-p)^{i-1}}{1-(1-p)^I},$$

with parameter $p = 0.2$. The mean holding time is the same for all channels, $1/\mu = 1$. In addition, the arrival intensity is the same for both user populations, $\lambda_u = \lambda$ and consequently, the traffic intensity $a = \lambda/\mu$ is the same for both user populations.

The results are given in table 1. The comparison was also done for multiservice traffic, where the capacity requirement is one for odd channels and two for even channel numbers. The capacity of the common link was eleven units and those of the other links were nine units. The results are given in table 2.

Table 1: Call blocking probabilities for the network in figure 2.

	Route1 ($u = 1$)			Route2 ($u = 2$)		
a	Exact	RLA	error	Exact	RLA	error
1.0	0.0056	0.0064	14 %	0.0027	0.0028	4 %
1.1	0.0084	0.0098	17 %	0.0041	0.0044	7 %
1.2	0.0121	0.0141	17 %	0.0060	0.0064	8 %
1.3	0.0166	0.0195	17 %	0.0083	0.0090	8 %
1.4	0.0220	0.0260	18 %	0.0112	0.0121	8 %
1.5	0.0282	0.0336	19 %	0.0146	0.0157	8 %
2.0	0.0715	0.0868	21 %	0.0382	0.0416	9 %

Table 2: Call blocking probabilities for the network in figure 2, with capacity requirements $c_{odd} = 1$ and $c_{even} = 2$.

		Route1 ($u = 1$)			Route2 ($u = 2$)		
Channel	a	Exact	RLA	error	Exact	RLA	error
7	1.0	0.0051	0.0058	14 %	0.0019	0.0022	16 %
8	1.0	0.0127	0.0139	9 %	0.0028	0.0029	4 %
7	1.1	0.0074	0.0086	16 %	0.0029	0.0034	17 %
8	1.1	0.0179	0.0198	11 %	0.0042	0.0045	7 %
7	1.2	0.0103	0.0120	17 %	0.0042	0.0049	17 %
8	1.2	0.0243	0.0270	11 %	0.0062	0.0066	6 %
7	1.3	0.0138	0.0162	17 %	0.0058	0.0068	17 %
8	1.3	0.0318	0.0355	12 %	0.0086	0.0092	7 %
7	1.4	0.0179	0.0211	18 %	0.0077	0.0091	18 %
8	1.4	0.0403	0.0454	13 %	0.0116	0.0124	7 %
7	1.5	0.0226	0.0268	19 %	0.0100	0.0118	18 %
8	1.5	0.0499	0.0566	13 %	0.0151	0.0162	7 %
7	2.0	0.0536	0.0645	20 %	0.0255	0.0299	17 %
8	2.0	0.1101	0.1276	16 %	0.0400	0.0431	8 %

The results confirm the comparisons made in [5]. RLA-algorithm yields blocking probabilities of the same magnitude as the exact method. As a rule, RLA gives larger blocking values for both routes. For route 2, RLA gives good results. This is because the route is very short, and the assumption of independence between the links is not violated severely.

5 Including background traffic

The networks considered until now were assumed to transfer only multicast traffic. The model can, however, be extended to cover networks with mixed traffic. In this case, the network transfers, in addition to multicast traffic, non-multicast traffic that is assumed independent on each link. The distribution does not depend on the multicast traffic in the link and the traffic in the other links. The non-multicast traffic in link j is assumed to be Poisson with a traffic intensity A_j. The capacity requirement is equal to one unit of capacity. The link occupancy distribution of the non-multicast traffic in a link with infinite capacity is thus,

$$q_j(z) = \frac{(A_j)^z}{z!} e^{-A_j}. \tag{16}$$

The inclusion of non-multicast traffic affects only the truncation step of the algorithm presented in section 3.5. The state probabilities are defined as in section 3. The state probabilities of the link states that require more capacity than available on the link are set to zero as before. However, the state probabilities of the states that satisfy the capacity restriction of the link are altered, as the available capacity on the link depends on the amount of non-multicast traffic on the link. Another way of describing the relationship between the two different types of traffic, is to consider them as two traffic classes in a two dimensional link occupancy state space as is shown in figure 3. The traffic classes are independent of each other. The capacity of the link is the linear constraint of this state space. We notice that the marginal distribution of the capacity occupancy of the

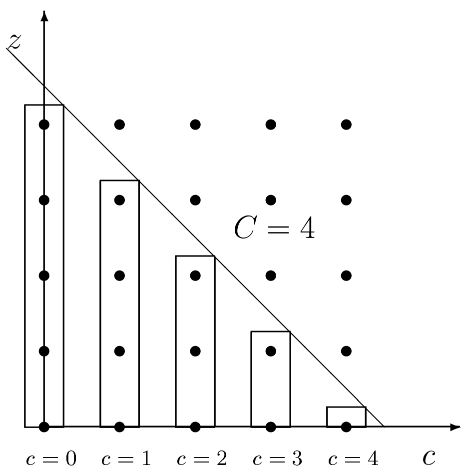

Figure 3: Shaping of the marginal distribution of the capacity occupancy when background traffic is included in the model.

multicast traffic is weighted by the sums over the columns of the occupancy probabilities of the background traffic. If the multicast traffic occupies $c = \mathbf{d} \cdot \mathbf{y}_j$ units of capacity, and the link capacity is C_j, then possible non-multicast traffic states on the link are those with $z \leq C_j - c$, where z is the number of non-multicast calls. Therefore, the truncation functions presented in equations (10) and (12) must be replaced by the operators

$$\hat{T}_j f(\mathbf{y}) = \begin{cases} \sum_{z=0}^{C_j - \mathbf{d}\cdot\mathbf{y}} q_j(z) f(\mathbf{y}) & , \text{ if } \mathbf{y} \in \tilde{\mathcal{S}}_j \\ 0 & , \text{ otherwise} \end{cases}$$

$$\hat{T}_j^{u,i} f(\mathbf{y}) = \begin{cases} \sum_{z=0}^{C_j - \mathbf{d}\cdot(\mathbf{y}\oplus(\mathbf{e}_i 1_{j\in\mathcal{R}_u}))} q_j(z) f(\mathbf{y}) & , \text{ if } \mathbf{y} \in \tilde{\mathcal{S}}_j^{u,i} \\ 0 & , \text{ otherwise.} \end{cases}$$

The algorithm differs therefore only by the truncation function used,

$$\hat{Q}_j(\mathbf{y}) = \begin{cases} \hat{T}_j \pi_j(\mathbf{y}) & , \text{ if } j \in \mathcal{U} \\ \hat{T}_j [\bigotimes_{k\in\mathcal{N}_j} \hat{Q}_k](\mathbf{y}) & , \text{ otherwise.} \end{cases}$$

Similarly,

$$\hat{Q}_j^{u,i}(\mathbf{y}) = \begin{cases} \hat{T}_j^{u,i} \pi_j(\mathbf{y}) & , \text{ if } j \in \mathcal{U} \\ \hat{T}_j^{u,i} [\bigotimes_{k\in\mathcal{N}_j} \hat{Q}_k^{u,i}](\mathbf{y}) & , \text{ otherwise.} \end{cases}$$

The blocking probability in equation (9) is again obtained by two series of convolutions and truncations from the leaf links to the common link J. The call blocking probability of the network is

$$\hat{b}_{u,i}^c = 1 - \frac{\sum_{\mathbf{y}\in\mathcal{S}} \hat{Q}_J^{u,i}(\mathbf{y})}{\sum_{\mathbf{y}\in\mathcal{S}} \hat{Q}_J(\mathbf{y})}.$$

5.1 Numerical results

The end-to-end call blocking probability was calculated using the same network as in section 4, figure 2. The intensity of the non-multicast traffic was set to $A_j = 0.1$ for all links. Table 3 shows the end-to-end call blocking probability for a network with only multicast traffic and for a network transferring multicast and non-multicast traffic. Table 4 shows the end-to-end call blocking probabilities when the multicast traffic requires double the capacity compared to the non-multicast traffic.

The intensity of non-multicast traffic stays the same, as the intensity of the multicast traffic increases. Clearly, the blocking probabilities are affected less, as the intensity of the multicast traffic increases. This can also be seen by studying the relative change in blocking probabilities shown in tables 3 and 4. The effect of the non-multicast traffic to the blocking probability is of the same magnitude on both routes. From table 3 we see that an inclusion of unicast traffic with one tenth the intensity $a = 1.0$ of the multicast traffic almost doubles the blocking probability. From table 1 the blocking probability increases by a factor of 1.5, when the traffic intensity a is increased from 1.0 to 1.1. These two cases are not equivalent as the background traffic is assumed independent of the multicast traffic, but give a good reference to the effect background traffic has on end-to-end blocking probabilities.

Table 3: Blocking probabilities for the network in figure 2 with background traffic and multicast traffic.

	Route1 ($u = 1$)			Route2 ($u = 2$)		
a	Multicast	Background	Rel. change	Multicast	Background	Rel. change
1.0	0.0056	0.0109	1.95	0.0027	0.0053	1.96
1.2	0.0121	0.0206	1.70	0.0060	0.0105	1.75
1.4	0.0220	0.0341	1.55	0.0112	0.0177	1.58
2.0	0.0715	0.0927	1.30	0.0382	0.0501	1.31

Table 4: Blocking probabilities for the network in figure 2 with background traffic requiring one unit and multicast traffic requiring two units of capacity.

	Route1 ($u = 1$)			Route2 ($u = 2$)		
a	Multicast	Background	Rel. change	Multicast	Background	Rel. change
1.0	0.0056	0.01	1.79	0.0027	0.0049	1.81
1.2	0.0121	0.0195	1.61	0.0060	0.0099	1.65
1.4	0.0220	0.0328	1.49	0.0112	0.0171	1.53
2.0	0.0715	0.0914	1.28	0.0382	0.0495	1.30

6 Conclusions

The paper presented a new algorithm for exactly calculating blocking probabilities in tree-structured multicast networks. The algorithm is based on the well-known algorithm for calculating blocking probabilities in hierarchical multiservice access networks. The multicast traffic characteristics were taken into account in the convolution step of the algorithm, using the new OR-convolution. Calculating the exact solution for the end-to-end call blocking probability, however, becomes infeasible as the number of channels increases. In contrast to ordinary access networks the aggregate one dimensional link occupancy description is not sufficient, since in the multicast network it is essential to do all calculations in the link state space, with 2^I states. This is due to the resource sharing property of multicast traffic, namely the capacity in use on a link increases only if a channel is requested when the channel is idle. The use of RLA was studied, as the complexity of the RLA-algorithm does not depend critically on the number of channels in the network. RLA method used in [5], however, gives larger blocking probabilities. Even for small networks, the errors are around 15 %. The network model and the algorithm for calculating call blocking probabilities were further broadened to include background traffic in addition to multicast traffic.

We leave for further research the study of new approximation methods for calculating blocking probabilities. The acceleration of the exact algorithm presented should also be investigated. At present, the model also assumes an infinite user population behind each leaf link. The model can be generalized to allow a finite user population behind a leaf link and is a subject for further study.

References

[1] Almeroth K.C. and Ammar M.H., Multicast Group Behavior in the Internet's Multicast Backbone (MBone), IEEE Communications Magazine 35(6), June 1997, pp. 124 -129.

[2] Chan W.C., Geraniotis E., Tradeoff Between Blocking and Dropping in Multicasting Networks, IEEE International Conference on Communications, Conference Record, Converging Technologies for Tomorrow's Applications pp. 1030 - 1034, 1996.

[3] Hui J., "Switching and traffic theory for integrated broadband networks", Kluwer Academic Publishers, Boston, 1990.

[4] Karvo J., Virtamo J., Aalto S., Martikainen O., Blocking of dynamic multicast connections in a single link, Proc. of Broadband Communications '98, ed. P. Kühn, R. Ulrich, IFIP, pp. 473-483, 1998.

[5] Karvo J., Virtamo J., Aalto S., Martikainen O., Blocking of Dynamic Multicast Connections, In 4th INFORMS Telecommunications Conference, Boca Raton, Florida, March 1998.

[6] Kelly F.P., "Reversibility and Stochastic Networks", John Wiley & Sons, 1979.

[7] Kim C.K., Blocking Probability of Heterogeneous Traffic in a Multirate Multicast Switch, IEEE Journal on Selected Areas in Communications 14(2), 1996, pp. 374-385.

[8] Listanti M., Veltri L., Blocking Probability of Three-Stage Multicast Switches, IEEE International Conference on Communications, Conference Record, 1998, pp. 623 -629.

[9] Ross K. W., "Multiservice Loss Models for Broadband Telecommunication Networks", Springer-Verlag, London, 1995.

[10] Shacham N., Yokota H., Admission Control Algorithms for Multicast Sessions with Multiple Streams IEEE Journal on Selected Areas in Communications 15(3), 1997, pp. 557-566.

[11] Stasiak M. and Zwierzykowski P., Analytical Model of ATM Node with Multicast Switching, Mediterranean Electrotechnical Conference, 1998, pp. 683 -687.

[12] Yang Y., On Blocking Probability of Multicast Networks, IEEE Transactions on Communications 46 (7), pp. 957 -968, July 1998.