Available online at www.sciencedirect.com
ScienceDirect
Computers and Electrical Engineering 32 (2006) 394–410
Computers and Electrical Engineering
www.elsevier.com/locate/compeleceng

Optimizing the performance of tree-based reliable multicast ☆

Zuo Wen Wan [a], Michel Kadoch [b,*], Ahmed Elhakeem [c]

[a] *Avici Systems Inc, MA, USA*
[b] *Department of Electrical Engineering at Ecole de Technologie Superieure, University of Quebec, Montreal, Canada*
[c] *Department of Electrical and Computer Engineering at Concordia University, Montreal, Canada*

Received 9 June 2004; accepted 30 August 2005
Available online 5 April 2006

Abstract

The expected number of transmissions is a very important parameter to evaluate the multicast performance. It is often used to estimate the bandwidth consumption of tree-based reliable multicast where some special receivers or routers are assigned for retransmission handling of each domain. In this work, we will analyze and simulate the tree-based multicast transmissions to obtain the number of transmissions and actual bandwidth consumption. This paper will also investigate the problem of finding the optimal partitioning of multicast trees based on the bandwidth consumption.

Keywords: Reliable multicast; Topology; Optimization; Partition; Bandwidth

1. Introduction

A number of reliable multicast protocols have been proposed in recent years [1–4]. The performance analysis of basic parameters such as delay and bandwidth consumption has been compared for different kinds of protocols in some references [4–9]. For reliable multicast, loss recovery is crucial. Each lost packet must be retransmitted to the receivers that did not get the packet in the previous transmission and retransmissions. In the analysis of bandwidth and delay, one often uses an important parameter, i.e. the expected number of transmissions for one packet [2,4–7,10]. These basic performance parameters of reliable multicast depend on many factors such as the loss recovery, the loss characteristics, the network topology and the partitioning of networks. Evaluations of these parameters are very important criteria for reliable multicast.

The number of transmissions analyzes the bandwidth performance well. The topological effects are found to be significant [4]. However, it is intensive to analyze the topological factor, so intermediate links are often considered to be almost loss-free, i.e. the loss-free intermediate models [5–8,11,12]. The bandwidth is wasted

☆ This work was supported by a CRD grant from NSERC and Bell Canada. The completion of this research was made possible, thanks to the support of Bell Canada through its Bell University Research Program.
* Corresponding author.
E-mail address: kadoch@ele.etsmtl.ca (M. Kadoch).

0045-7906/$ - see front matter © 2006 Elsevier Ltd. All rights reserved.
doi:10.1016/j.compeleceng.2005.08.002

mainly by receiver links whose topology is star. Impacts of topology are neglected in this case. In general, intermediate routers may lose multicast packets, and losses may still exist in intermediate links. For example, due to the buffer overflow of intermediate routers, packets may be lost. The losses from intermediate links will result in retransmissions to many nodes and will consume more bandwidth. Impact of these losses from intermediate links on the multicast network is significant. In the old model of loss-free intermediate links, the effects of topology were not considered. If intermediate links are considered to be lost [4,6,10,13], one must consider the influence of topology. Even if they have the same number of links, networks of different topologies may have different number of retransmissions, and different bandwidth consumption. In this work, we will analyze and simulate multicast transmissions, consider loss from intermediate links, and focus on the effects of topology on multicast performance.

2. Performance evaluation

The expected number of transmissions $E[M]$ that a packet should be multicast by the source using selective reject ARQ until all group members correctly receive it can be recursively calculated [4,5]. In the analytical study of $E[M]$, one often uses the CDF (cumulative distribution function) of the total number of transmissions to consider loss from intermediate links and to analyze the effects of topology on multicast performance [4]. Let $M(k)$ be the total number of transmissions of a packet until it is received by all receivers under node k. Then CDF for node k is $F_k(m) = P[M(k) \leqslant m]$, i.e. $F_k(m) =$ Prob.[all nodes from k and below get the packet at most in m trials]. We denote $F_r(m)$, $F_k(m)$ and $F_s(m)$ to be CDF of the total number of transmission for leaf receivers (r), intermediate nodes (k) and the sender (s), respectively. Then one can obtain the following equations for $E[M]$ [4]:

$$E[M(s)] = \sum_{m=0}^{\infty} mP_s(M = m) = \sum_{m=o}^{\infty}(1 - F_s(m)), \tag{1}$$

$$F_s(m) = \prod_{c\in \text{child}(s)} F_c(m), \tag{2}$$

$$F_n(m) = \sum_{u=0}^{m-1} \binom{m}{u} p_n^u (1 - p_n)^{m-u} \prod_{c\in \text{child}(n)} F_c(m - u), \tag{3}$$

$$F_r(m) = 1 - p_r^m, \tag{4}$$

where p_k or p_r is the packet loss probability of the link leading to node k or leaf receiver r.

Needless to say, one should use recursion starting from the bottom of the nodes, and should use the numerical evaluation of Eqs. (1)–(4) to calculate $E[M]$. However, due to the consideration of topological effects and loss from intermediate links [4], the computation of $E[M]$ may be very intensive for a general topology [5]. Even if the reduction technique is used, the computation of $E[M]$ can be exponential with the number of nodes N [3]. In the worst case, it may take hours to calculate $E[M]$. Based on intensive computations of $E[M]$, it is difficult to calculate other parameters, such as bandwidth and delay, for a big multicast network, especially for real time dynamic networks. The efficient estimation of $E[M]$ is necessary. In this paper we follow a completely analytical approach for the evaluation of $E[M]$ without using recursion. We consider the loss from intermediate links and further evaluate the effects of topology on multicast by introducing a new topology parameter. This significantly simplifies the evaluation of $E[M]$, which is very useful for dynamic networks. In order to focus primarily on the effects of topology on multicast, we use homogeneous packet loss probability in the following, i.e. all links have the same packet loss probability p. Our approach can be extended to discuss networks with heterogeneous loss probability and to further analyze the effect of loss characteristics on multicast. Packet loss probability per link can be obtained by the following equation:

$$p = \#\text{lost packets}/\#\text{total transmitted packets}. \tag{5}$$

It includes link error and packet loss due to buffer overflow.

2.1. Probability of the number of transmissions

For a tree-based reliable multicast, packet loss of one intermediate node is involved in the retransmissions of many nodes. Once one intermediate node loses a packet, all nodes under the node will lose it. For example, if node k_1 loses a packet in Fig. 1, n_{k_1} nodes will not receive the packet where n_{k_1} is the number of links under the node k_1. The purpose of the next retransmission is to have the requested packet received at the n_{k_1} nodes. The requested packet must pass through the path from the sender to node k_1 to recover the loss. The number of links involved in retransmissions depends not only on the nodes that did not receive the packet, but also on the nodes along the path from the sender to node k_1. We denote by D_{k_1} the set of these links involved in retransmissions if loss occurs over link k_1, as illustrated in Fig. 1. D_{k_1}, including d_{k_1} and n_{k_1} links, reflects the effect of dependent loss over link k on multicast. Therefore, we only concentrate on whether these links within D_{k_1} correctly receive the packet in the next retransmission. If loss out of D_{k_1} recurs in the next retransmission, we need to repeat the above process until all nodes receive the packet correctly. We do not concentrate on whether the nodes that have received the packet correctly do so again on further retransmissions. Therefore, each multicast retransmission is aimed at recovering loss from the previous transmission or retransmissions.

After understanding the process of loss recovery, we can evaluate the multicast performance. Packet losses may take place at different nodes in each transmission or retransmission. The loss recovery of different nodes needs different retransmission times; therefore, the probability density of the number of transmissions M depends on which node loses the packet.

Here, we first clarify two kinds of losses. If the loss of one link is not under the subtree of another link loss, the two losses do not affect each other; for example, the losses of node k_1 (A) and k_2 (B) are independent. However, if one link is under the subtree of another link loss, the losses are dependent; for example, the loss of node B clearly results in the loss of node C. In the following, loss refers to an independent loss that does not affect each other, and dependent loss is considered by factor D_k.

2.1.1. No loss occurs

The probability that no loss occurs in the first transmission is

$$\text{Prob.(no loss)} = (1-p)^N. \tag{6}$$

Readers are urged to read Table 1 for multicast notations.

2.1.2. Only one loss occurs

Assume that only one loss occurs in the first transmission; e.g. node k_1 loses the packet. The probability that only node k_1 loses data is

$$\text{Prob.(only node } k_1 \text{ loses)} = p(1-p)^{N-n_{k_1}-1}, \tag{7}$$

where n_{k_1} is the number of links under node k_1.

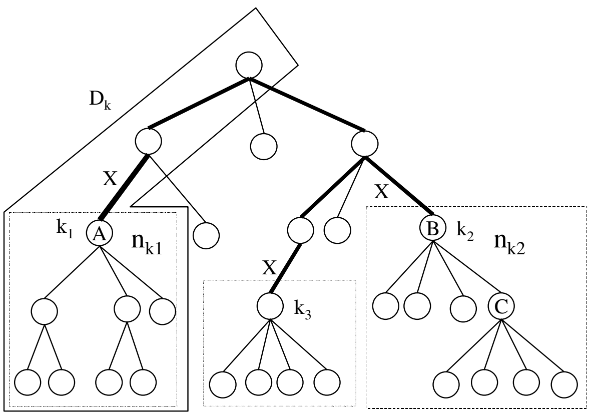


Fig. 1. One example of a general tree.

Table 1
Notations for a reliable multicast

N is the total number of links
p is the packet loss probability of each link
M, m is the number of transmissions (including retransmissions) for one packet
d_k is the number of links from the sender to node k
n_k is the total number of links under node k
$d_{k_1k_2\cdots k_i}$ is the total number of links from the sender to node $k_1,k_2,\ldots,k_i$, $i=1,2,\ldots$
D_{k_1} is the set of links involved in retransmissions if loss occurs over link k_1
$D_{k_1k_2\cdots k_i}$ $(i=1,2,\ldots,j)$ is the set of links involved in retransmissions for independent loss over link $k_1,k_2,\ldots,k_i$
$P(M=m)$ is the probability of the mth transmission for one packet
$P(M=m\mid k_1k_2\cdots)$ is the conditional probability of m retransmissions after node k_1, node $k_2,\ldots$, lose the packet in a previous transmission

2.1.3. j independent losses occur

The probability that node $k_1,k_2,\ldots,k_j$ lose the packet in the first transmission is

$$\text{Prob.(node } k_1,k_2,\ldots,k_j \text{ lose)} = \begin{cases} p^j(1-p)^{N-\sum_{s=1}^{j} n_{k_s}-j}, & k_2 \notin D_{k_1},\ldots,k_j \notin D_{k_1k_2\cdots k_{j-1}} \\ 0, & \text{otherwise} \end{cases}, \tag{8}$$

where n_{k_s} is the number of links under node k_s $(s=1,2,\ldots,N)$. $d_{k_1k_2\cdots k_i}$ is the total number of links from the sender to node $k_1,k_2,\ldots,k_i$. For example, $d_{k_1k_2}=4$, $d_{k_2k_3}=4$, and $d_{k_1k_2k_3}=6$ (in bold) in Fig. 1. $D_{k_1k_2\cdots k_i}$ $(i=1,2,\ldots,j)$ is the set of links involved in retransmissions for independent loss $k_1,k_2,\ldots,k_i$, i.e. including $d_{k_1k_2\cdots k_i}$ and $n_{k_s}(s=1,2,\ldots,i)$ links.

After we know the probability which nodes lose the packet in one transmission, we can recursively calculate the probability density of the number of transmissions M. The probability of only one transmission is obvious:

$$P(M=1)=(1-p)^N. \tag{9}$$

If some nodes lose the packet in the first transmission and the sender can recover these losses in the first retransmission, one can obtain the probability of two transmissions:

$$P(M=2)=\sum_{k_1=1}^{N} p(1-p)^{N-n_{k_1}-1}P(M=1\mid k_1)+\frac{1}{2!}\sum_{k_1=1}^{N}\sum_{k_2\notin D_{k_1}} p^2(1-p)^{N-n_{k_1}-n_{k_2}-2}P(M=1\mid k_1k_2)+\cdots, \tag{10}$$

where $P(M=1\mid k_1k_2\cdots k_s)$ is the conditional probability that one retransmission is needed to make the multicast successful if nodes $k_1,k_2,\ldots,k_s$ lose the packet in the first transmission.

Similarly, we can recursively obtain the probability of m transmissions:

$$\begin{aligned} P(M=m)=&\sum_{k_1=1}^{N} p(1-p)^{N-n_{k_1}-1}P(M=m-1\mid k_1)+\cdots \\ &+\frac{1}{j!}\sum_{k_1=1}^{N}\sum_{k_2\notin D_{k_1}}\cdots\sum_{k_j\notin D_{k_1\cdots k_{j-1}}} p^j(1-p)^{N-\sum_{s=1}^{j} n_{k_s}-j}P(M=m-1\mid k_1k_2\cdots k_j)+\cdots, \end{aligned} \tag{11}$$

where $P(M=m-1\mid k_1k_2\cdots k_j)$ is the conditional probability of $m-1$ retransmissions after node k_1, node $k_2,\ldots,k_j$ lose the packet in previous transmission; for example, $P(M=m-1\mid k_1)$ is the conditional probability of $m-1$ retransmissions if node k_1 loses the packet.

2.2. Analytical approximation of E[M]

Each multicast subgroup meets the condition of $Np<1$. If $Np>1$, the sender needs to multicast the packet so many times that the multicast protocol cannot work properly. Then, we need to design loss recovery to reduce the number of retransmissions. The local recovery and partitioning of a group can be efficiently used to limit the scope of loss recovery so that $Np < 1$ for each subgroup. Thus, we consider only the case of $Np < 1$ in this work.

For a multicast subgroup of $Np < 1$, we may expand the above probability density function of M according to the order of loss probability p. In the following, we will consider the function until the 3rd order approximation of p or Np. From (9), one may have

$$P(M = 1) = (1 - p)^N = 1 - Np + \binom{N}{2}p^2 - \binom{N}{3}p^3 + \cdots \tag{12}$$

In order to calculate $P(M = 2)$, one needs to retransmit only once after the losses take place. For example, if only node k_1 loses the packet in the first transmission, the probability that the first retransmission will be successful is $P(M = 1 \mid k_1) = (1 - p)^{d_{k_1}+n_{k_1}}$, where d_{k_1} is the number of links from the sender to node k_1 and n_{k_1} is the number of links under node k_1. Similarly, one can obtain

$$P(M = 1|k_1k_2\cdots k_j) = (1 - p)^{d_{k_1k_2\cdots k_j}+\sum_{s=1}^{j} n_{k_s}}, \quad k_2 \notin D_{k_1}, \ldots, k_j \notin D_{k_1k_2\cdots k_{j-1}}, \tag{13}$$

where $d_{k_1k_2\cdots k_j}$ is the number of links from the sender to node $k_1, k_2, \ldots, k_j$. Substituting (13) in (10), one has

$$P(M = 2) = \sum_{k_1=1}^{N} p(1 - p)^{N+d_k-1} + \frac{1}{2!}\sum_{k_1=1}^{N}\sum_{k_2\notin D_{k_1}} p^2(1 - p)^{N+d_{k_1k_2}-2} + \frac{1}{3!}\sum_{k_1=1}^{N}\sum_{k_2\notin D_{k_1}}\sum_{k_3\notin D_{k_1k_2}} p^3(1 - p)^{N+d_{k_1k_2k_3}-3} + \cdots \tag{14}$$

After expanding the equation above, one may obtain the following approximation:

$$P(M = 2) = Np - \left[\binom{N}{2} + \sum_{k_1=1}^{N}(n_{k_1} + d_{k_1})\right]p^2 + a_{23}p^3 + \cdots, \tag{15}$$

where

$$a_{23} = \frac{1}{3!}\sum_{k_1=1}^{N}\sum_{k_2\notin D_{k_1}}\sum_{k_3\notin D_{k_1k_2}} 1 + \sum_{k_1=1}^{N}\binom{N + d_k - 1}{2} - \frac{1}{2!}\sum_{k_1=1}^{N}\sum_{k_2\notin D_{k_1}}(N + d_{k_1k_2} - 2). \tag{16}$$

Similarly, one may use (15) to evaluate $P(M = 2|k_1k_2\cdots k_j)$. If one only considers the 3rd order approximation, one achieves the following result:

$$P(M = 3) = \sum_{k_1=1}^{N}(d_{k_1} + n_{k_1})p^2 + a_{33}p^3 + \cdots, \tag{17}$$

where

$$a_{33} = -\sum_{k_1=1}^{N}\left[\left(N - \frac{3}{2} + \frac{d_{k_1} - n_{k_1}}{2}\right)(d_{k_1} + n_{k_1}) + \sum_{k_2\in D_{k_1}}(d'_{k_2} + n'_{k_2})\right] + \frac{1}{2!}\sum_{k_1=1}^{N}\sum_{k_2\notin D_{k_1}}(d_{k_1k_2} + n_{k_1} + n_{k_2}). \tag{18}$$

Using (17), one can obtain the 3rd approximation of $P(M = 4)$:

$$P(M = 4) = a_{43}p^3 + \cdots, \tag{19}$$

where

$$a_{43} = \sum_{k_1=1}^{N}\sum_{k_2\in D_{k_1}}(d'_{k_2} + n'_{k_2}). \tag{20}$$

So far, we have derived until the 3rd order approximation for the probability density of the number of transmissions and retransmissions. One may find the expected number of transmissions easily using (12), (15), (17), and (19):

$$E[M] = \sum_{m=1}^{\infty} mP(M = m) \approx 1 + Np + \left[xN - \binom{N}{2}\right]p^2 + \left[\binom{N}{3} + y\right]p^3, \tag{21}$$

where

$$x = \frac{1}{N}\sum_{k=1}^{N}(n_k + d_k), \tag{22}$$

$$y = -2\binom{N}{3} + 2a_{23} + 3a_{33} + 4a_{43}. \tag{23}$$

After some calculations, y may be simplified to the following equation [10]:

$$y = \frac{5}{6}\sum_{k=1}^{N}\left[(d_k)^2 - (n_k)^2\right] + \frac{xN}{6}. \tag{24}$$

Assume that $d_k + n_k \sim x$. We can estimate the y value as

$$y \sim \frac{5x}{6}\sum_{k=1}^{N}(d_k - n_k) + \frac{xN}{6} = xN. \tag{25}$$

We may ignore the effect of y on $E[M]$ due to $yp^3 \sim xNp^3 \ll xNp^2$ for the small p value and to $y \ll \binom{N}{3}$ for the large N value.

2.3. Discussions

From (21), $E[M]$ depends on the number of links N, the packet loss probability p, and the topology parameter. The estimation of $E[M]$ is simple and it just takes seconds to evaluate $E[M]$, which is very useful for parameter evaluation of real time dynamic networks. The number of transmissions $E[M]$ increases with p and N. p reflects on the packet loss probability over a link. It depends on link error and packet loss due to buffer overflow. p does not change as $E[M]$. Different topologies have different values of x. x is defined as the topology parameter which reflects the effects of topology on reliable multicast due to the consideration of intermediate link loss.

From [10], we have $\sum_{k=1}^{N} n_k = \sum_{k=1}^{N}(d_k - 1)$; therefore, x is written as follows:

$$x = \frac{1}{N}\sum_{k=1}^{N}(n_k + d_k) = \frac{2}{N}\sum_{k=1}^{N} n_k + 1 = \frac{2}{N}\sum_{k=1}^{N} d_k - 1. \tag{26}$$

If each node knows the number of links away from the sender, or the number of links downstream, we can use x in (26) to analyze the change of distribution trees, and to further reflect the topological effects on multicast performance. The x value reaches the biggest value for the linear topology and the smallest value for the star topology if they have the same number of links. For a linear topology, $n_k + d_k = N$, so it is easy to get an x and a y value from (26) and (24); $x = N$ and $y = N^2$. For a star topology, $d_k = 1$ and $n_k = 0$, so $x = 1$ and $y = N$ from (26) and (24). This approximation (21) for the star and linear topologies coincides perfectly with the exact solution until the 3rd order result.

2.4. Hierarchical reliable multicast

For tree-based multicast protocols, receivers are organized hierarchically in a tree. Some special receivers or routers, for example, designated receivers (DR) in reliable multicast transport protocol (RMTP [1]), manage a group of receivers or a domain. In this paper, we will use repair routers (RR) to represent these special receivers or routers, i.e. RRs have repair functions and handle retransmissions. In fact, the multicast group is partitioned into different subgroups according to the repair routers [1,3,10]. RRs are responsible for these subgroups. If some receivers lose the packet, they request retransmissions from their RRs. If an RR does not receive a data packet, it will ask for retransmission from the sender. Placing r RRs that can send repairs to requesting nodes, one has $r+1$ subgroups, e.g., in Fig. 2 we have 3 subgroups after putting 2 RRs in this multicast group, and in Fig. 3 we have 2 subgroups for 1 RR. We denote by S_i the subgroup covered by RR_i. S_0 is the subgroup not covered by any RR, i.e. the sender is responsible for the retransmissions of S_0. It is obvious that multicast performance depends on RR locations because RRs are responsible for the retransmissions of these subgroups. Based on actual bandwidth consumption of different RR locations, we try to find the optimal RR locations.

We define C as the total bandwidth consumed by one source multicast packet over all links and all transmissions, including retransmissions. The bandwidth consumption of the whole multicast group is the summation of the bandwidth consumed by all subgroups. Therefore, for r RRs, one may find the total bandwidth consumption $E[C]$:

$$E[C] = \sum_{i=0}^{r} E[C_i], \tag{27}$$

where C_i is the total bandwidth consumed by one source multicast packet over all links in subgroup S_i whose retransmissions are handled by one RR. If B is defined as the bandwidth consumed by a multicast packet per link, averaged over all links in a multicast group, one may obtain the following equation:

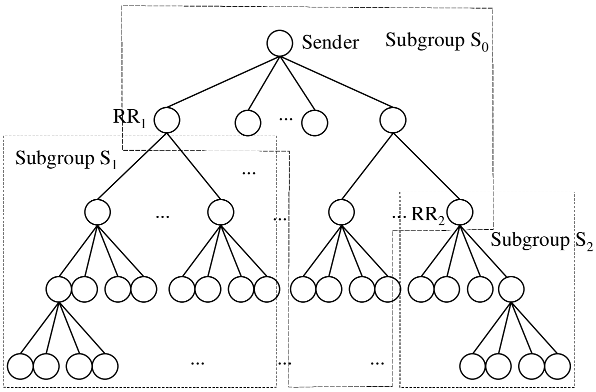


Fig. 2. One example of homogeneous topology structure – k-ary tree. w is the number of branches for each intermediate node.

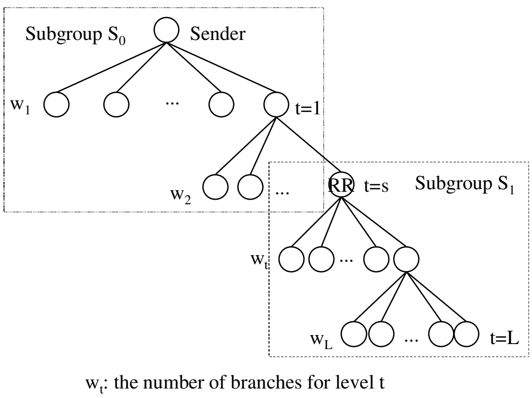


Fig. 3. One example of heterogeneous topology for different subgroups. Note $w_1 = w_2 = \cdots = w_L$ is homogeneous topology.

$$E[B] = \frac{E[C]}{N}, \tag{28}$$

where N is the number of links. Average bandwidth $E[C_i]$ of subgroup S_i can be obtained by counting all links affected by multicast transmissions in simulation. It may also be estimated analytically:

$$E[C_i] \approx E[M_i]N_i, \tag{29}$$

where $E[C_i]$ and $E[M_i]$ are the expected bandwidth consumption and the expected number of transmissions for subgroup S_i respectively, and N_i is the number of links for subgroup S_i. Therefore, substituting (29) and (27) into (28), one may find the expected bandwidth consumption $E[B]$:

$$E[B] = \frac{1}{N}\sum_{i=0}^{r} N_i f_i(N_i), \tag{30}$$

where $f_i(N_i) = E[M_i]$, $i = 0, 1, \ldots, r$. One may use Lagrange multipliers to obtain optimal condition, i.e. we need to minimize the function subject to the following constraint:

$$\sum_{i=0}^{r} N_i - N = 0. \tag{31}$$

Introducing Lagrange multiplier λ, we form a new objective function $g(N_0, N_1, \ldots, N_r)$ which is to be minimized:

$$g(N_0, N_1, \ldots, N_r) = \frac{1}{N}\sum_{i=0}^{r} N_i f_i(N_i) + \lambda\left(\sum N_i - N\right). \tag{32}$$

We find its partial derivatives with respect to $N_0, N_1, \ldots, N_r$ and set them equal to zero:

$$\frac{\partial}{\partial N_i} g(N_0, N_1, \ldots, N_r) = \frac{1}{N}\left[f_i(N_i) + N_i \frac{\partial f_i(N_i)}{\partial N_i}\right] + \lambda = 0, \quad i = 0, 1, \ldots, r. \tag{33}$$

Thus we have the following optimal condition of r RRs:

$$f_0(N_0) + N_0 \frac{\partial f_0(N_0)}{\partial N_0} = f_1(N_1) + N_1 \frac{\partial f_1(N_1)}{\partial N_1} = \cdots = f_r(N_r) + N_r \frac{\partial f_r(N_r)}{\partial N_r}. \tag{34}$$

One can clearly see from (34) that $N_0 = N_1 = N_2 = \cdots = N_r = N/(r+1)$ is a solution to the optimal problem only for the case of homogeneous topology as will follow.

2.4.1. Homogeneous topologies for different subgroups

If each subgroup has the same topology, then $E[M_0]$ and $E[M_1]$ should have the same functional dependence on the number of links. Many topologies fit into this case: in the linear topology, each subgroup always has the same topology, no matter where RRs are placed; in Fig. 2, all subgroups have a similar topology; in Fig. 3, with $w_1 = w_2 = \cdots = w_L$, the subgroups also have the same topologies when RRs are placed in the intermediate nodes. Thus, they have the same functional dependence with N, i.e. $f_0(N) = f_1(N) = f(N)$. From analytical approximation (21) in the case that y is neglected, if the x of each subgroup has the same functional dependence with N, each x has the same $f(N)$ function. From (21) and (26), the subgroups have the same function x and the same functional dependence on N, as long as each level has the same number of links for two subgroups. In the case of the same functional dependence with N, i.e. $f_0(N_0) = f_1(N_1) = \cdots = f_r(N_r) = f(N)$, or $x_0(N_0) = x_1(N_1) = \cdots = x_r(N_r)$, the optimal conditions of RRs are obvious from (34):

$$N_0 = N_1 = N_2 = \cdots = N_r. \tag{35}$$

Therefore, RR should be optimally placed to partition the group into such subgroups having the same number of links. This only applies to homogeneous topology structures.

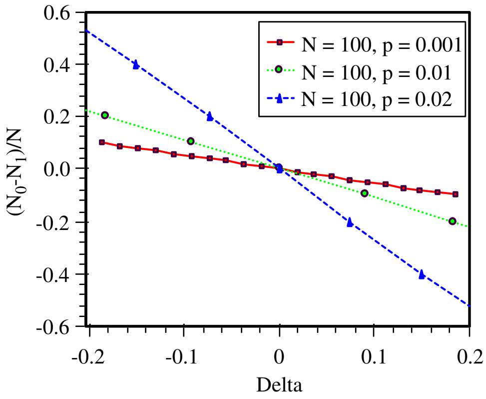


Fig. 4. The effects of topology on subgroup size.

2.4.2. Heterogeneous topologies for different subgroups

If each subgroup has a different topology, then $E[M_0]$ and $E[M_1]$ will have a different functional dependence on the number of links; for example, in Fig. 3, if $w_1 = w \gg 1$, $w_2 = \cdots = w_L = 1$, one subgroup resembles a linear topology and another resembles a star topology when one RR is placed in tth level. The function $f_i(N_i)$ of each subgroup can be obtained from (21) i.e.

$$f_i(N_i) = x_i N_i p^2 + g(N_i), \tag{36}$$

where

$$g(N_i) = 1 + N_i p - \binom{N_i}{2} p^2 + \binom{N_i}{3} p^3. \tag{37}$$

Substituting (36) into (34), one gets the optimal condition for 2 subgroups:

$$\frac{\partial[N_1 g(N_1)]}{\partial N_1} - \frac{\partial[N_0 g(N_0)]}{\partial N_0} = \Delta, \tag{38}$$

where

$$\Delta = \left[\frac{\partial(N_0^2 x_0)}{\partial N_0} - \frac{\partial(N_1^2 x_1)}{\partial N_1}\right] p^2. \tag{39}$$

Δ reflects the effect of topology on the optimal partitioning of multicast trees, shown in Fig. 4. The subgroup with the larger topology parameter should have a smaller size. In the case of a larger value of Np (e.g. $Np = 2$, $p = 0.02$), the topology effect is significant and the subgroup size is different. The more different topology, the more different size.

Thus, $N_0 = N_1$ will not be the optimal condition for this case. In order to meet the optimal condition (34), the linear-like subgroup should have a small size while the star-like subgroup should have a larger size.

3. Simulations

For a single-source multicast tree, a sender is located at the top of the tree. The sender multicasts the repair packet until each node correctly receives it at least one time. Different packets may have different transmission times due to loss. M is a random sample of the number of transmissions. Furthermore, links actually affected by each transmission are different. The bandwidth consumption depends on the locations where loss takes place. In order to obtain the total bandwidth consumption C, we need to summarize all links actually affected by all multicast transmissions. C is a random sample of actual bandwidth consumption. Thus the sample mean $E[C]$ ($E[M]$) of bandwidth (the number of transmissions) can be averaged over many packets. Fig. 5 shows our

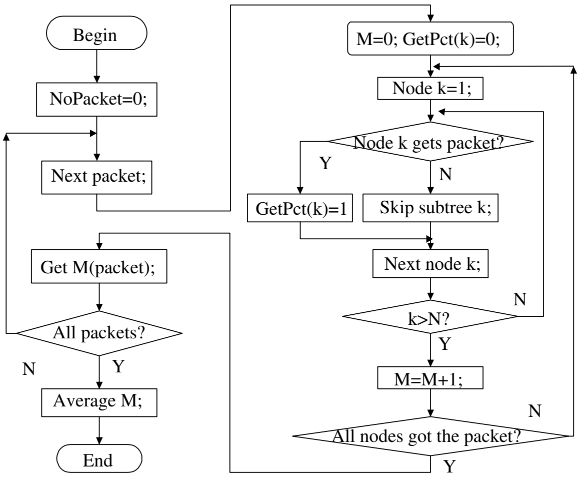


Fig. 5. The simulation flowchart for one subgroup.

simulation flowchart to obtain $E[M](E[C])$. Control traffic condition, such as NAK packets is assumed ideal with neglected bandwidth consumption. Data packets are sent from the sender.

In Fig. 5, we give the simulation flowchart where we average many packets to obtain the average number of transmissions. Let M_i denote the number of transmissions for one specific multicast packet i. Averaged over θ packets (e.g. $\theta = 10{,}000$),

$$E[M] = \sum_{i=1}^{\theta} M_i/\theta. \tag{40}$$

The sample variance σ^2 of a random sample of size θ may be written as follows:

$$\sigma^2 = \frac{\theta \sum_{i=1}^{\theta} M_i^2 - \left(\sum_{i=1}^{\theta} M_i\right)^2}{\theta(\theta - 1)}. \tag{41}$$

After evaluating the mean μ (an estimate of $E[M]$) and standard deviation of a random sample, we can calculate the $(1 - \alpha)\%$ confidence interval for μ in the following:

$$E[M] - z_{\alpha/2}\frac{\sigma}{\sqrt{\theta}} < \mu < E[M] + z_{\alpha/2}\frac{\sigma}{\sqrt{\theta}}, \tag{42}$$

where $z_{\alpha/2}$ is the z-value with $\theta - 1$ degrees of freedom, leaving an area of $\alpha/2$ to the right.

4. Results and discussions

4.1. The number of transmissions

We simulate multicast transmissions and obtain the average number of transmissions $E[M]$ over many packets. In these figures, we average the number of transmissions over 10,000 packets. This was found to be enough to provide us with a confidence interval of 95% about the true mean of each result in all the figures (Figs. 6–8, 10 and 11) and a max error of 3% from true mean. We also compare simulation results with approximate and exact results for some topologies.

Figs. 6–8 give the comparisons of several topologies among simulation results, the exact (i.e. the earlier recursion) solution (1)–(4) and approximate solutions (21) for $E[M]$. For small $E[M]$ value, they match perfectly.

For an example of a general topology shown in Fig. 9, we also give comparisons between analysis and simulations for homogeneous and heterogeneous loss probability. Fig. 10 compares the analysis and simulation

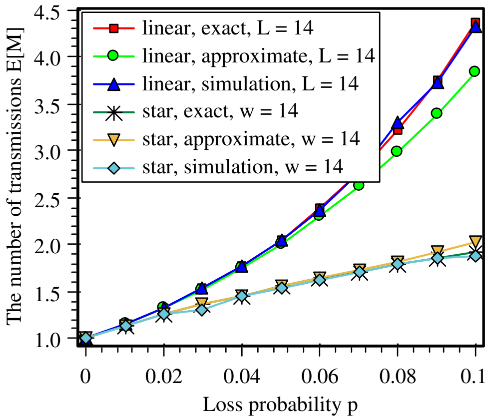


Fig. 6. Comparisons of $E[M]$ corresponding to analysis and simulation for the linear topology and the star topology. They have the same number of links.

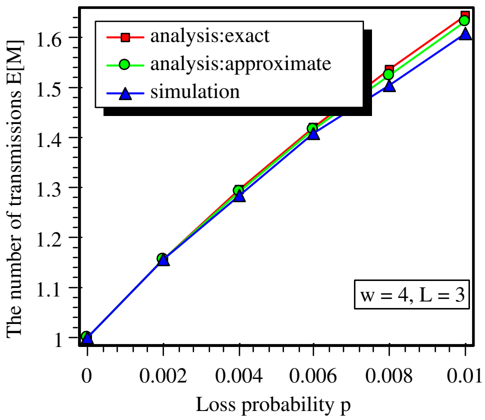


Fig. 7. Comparisons of $E[M]$ corresponding to analysis and simulation for 4-ary tree in Fig. 2.

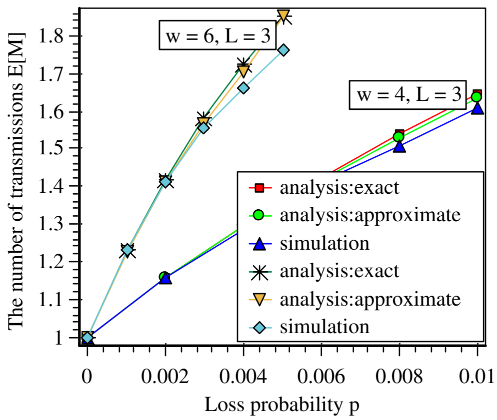


Fig. 8. Comparisons of $E[M]$ corresponding to analysis and simulation for the topology in Fig. 3. Here $w_1 = w_2 = \cdots = w_L = w$.

results where all links have the same loss probability, and Fig. 11 compares a few values where different links have different loss probability. In Fig. 11, the loss probability of each link is the summation of the value of the

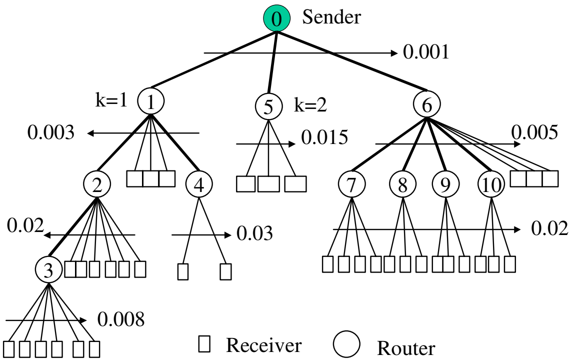


Fig. 9. One example of a general multicast network.

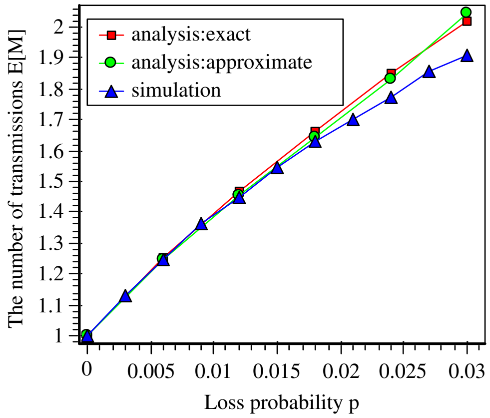


Fig. 10. Comparisons of $E[M]$ corresponding to analysis and simulation for the topology in Fig. 9 assuming that all links have the same loss probability p.

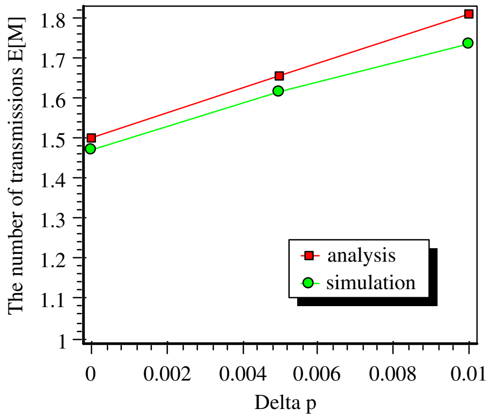


Fig. 11. A comparison of $E[M]$ corresponding to analysis and simulation for a heterogeneous network in Fig. 9 where the loss probability of each link is summation of value of the link in Fig. 9 and delta p.

link in Fig. 9 and delta p. The first set of data (i.e. delta p is 0) is for the network in Fig. 9 while the second and third set of data in Fig. 11 (i.e. delta p are 0.005 and 0.01, respectively) are for the increase value of delta p for loss probability of each link. The analysis and simulation results in Fig. 11 match closely.

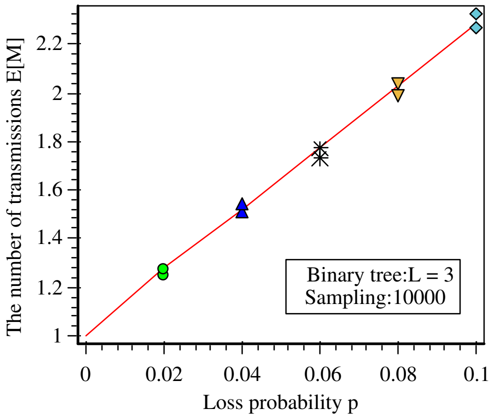


Fig. 12. 95% confidence interval. L is depth.

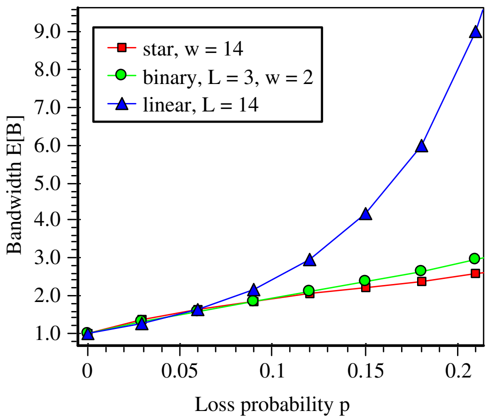


Fig. 13. The bandwidth consumption of different topologies versus the packet loss probability.

In Fig. 12, we use one example of binary trees to give the confidence interval. It is obvious that the more the packets are averaged, the more accurate the simulation $E[M]$ value will be.

For each multicast transmission, not all links are affected. If an intermediate link loses the packet in one transmission, all nodes under this link do not receive the packet. Hence actual bandwidth consumption depends on the topology. In our simulation, we summarize all links actually affected by the multicast transmissions to get the bandwidth, then we average the bandwidth over 10,000 packets to obtain the average bandwidth consumption $E[C]$. Fig. 13 shows the actual bandwidth consumption of different topologies. Actual bandwidth consumption of these topologies differs slightly for small loss probability. With the increasing loss probability, the bandwidth difference of different topologies becomes larger and larger.

4.2. The optimal RR locations for homogeneous packet loss probability

From the analysis in (34), the policy for placing RRs depends on the types of topologies of different subgroups. Multicast networks with homogeneous topology structure have similar topologies after partitioning according to RR locations. Thus, optimal placements depend only on the number of links of each subgroup, i.e. each subgroup should have the same size. We use k-ary trees to illustrate this case, shown in Figs. 14 and 15, Tables 2 and 3. When RRs are placed in a network of a k-ary tree, each subgroup has a similar topology. Thus, RRs are always placed to have the same topology and the same number of links for each subgroup. We

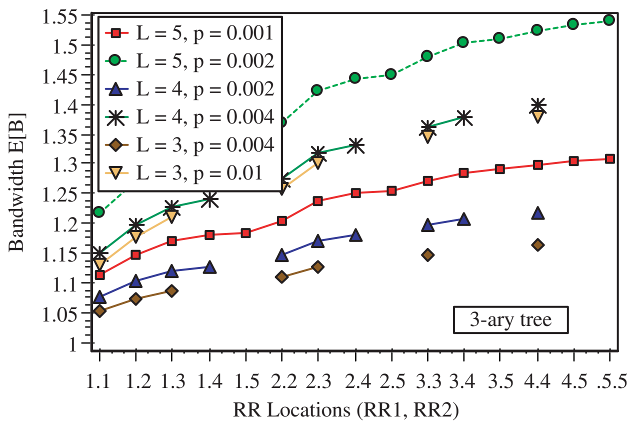


Fig. 14. Bandwidth consumption $E[B]$ of 2 RRs for 3-ary trees.

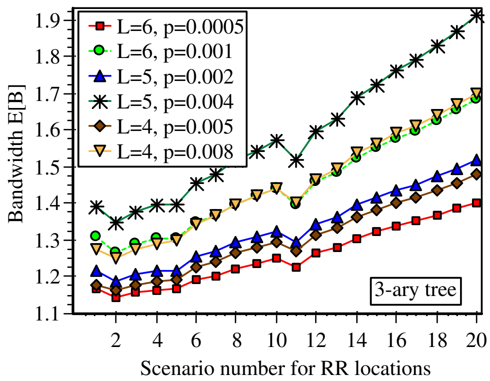


Fig. 15. Bandwidth $E[B]$ of 3 RRs for 3-ary trees. RR locations corresponding to the x axis are shown in Table 2.

Table 2
Levels of 3 RRs corresponding to each scenario number on the x axis of Fig. 15

	Scenario number of x axis					
	1	2	3	4	5	6
RRs levels	1,1,1	1,1,2	1,1,3	1,1,4	1,2,2	1,2,3
7	8	9	10	11	12	13
1,2,4	1,3,3	1,3,4	1,4,4	2,2,2	2,2,3	2,2,4
14	15	16	17	18	19	20
2,3,3	2,3,4	2,4,4	3,3,3	3,3,4	3,4,4	4,4,4

Table 3
Optimal RR levels for k-ary trees

#RRs	w			
	2	3	4	5
2 RRs	1,2	1,1	1,1	1,1
3 RRs	2,2,2	1,1,2	1,1,1	1,1,1

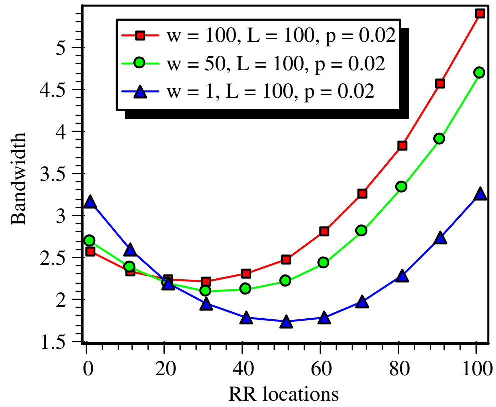


Fig. 16. Bandwidth $E[B]$ of 1 RR for heterogeneous topology in Fig. 3. $w_1 = w$, $w_2 = \cdots = w_L = 1$. Note $w_1 = 1$ is homogeneous topology.

can see these results from Fig. 15 and Table 3. For example, levels 1,1,2 are the optimal placements for 3 RRs in the case of 3-ary tree because this is the best RR combination having the same topology and the same size for 4 subgroups. Pure linear topology also has similar results. No matter how we partition the linear topology, each subgroup always has the linear topology. Thus each subgroup should have the same number of links to get the best performance of multicast. Optimal placements of RRs are trade-off between the sender and each RR. When each subgroup has the same number and the same topology, the whole group has the best performance.

When each subgroup has different topology, e.g. topology in Fig. 3, results of having the same size for each subgroup are not applicable. Fig. 16 gives the optimal placements of RRs, where $w_1 = w$, $w_2 = \cdots = w_L = 1$ and the location of RR is the level of RR. For example, optimal placement for $w = 100$ in Fig. 16 (square curve) is in level 20, that means, $N_0 = 100 + 20 = 120$ and $N_1 = 100 - 20 = 80$. We have different number of links for each subgroup to have the optimal placements of RRs. When we have the same size, $N_0 = 100$ and $N_1 = 100$ for $w = 100$ in Fig. 16 (square curve), i.e. RR is in level 1, multicast performance is not optimal. The same number of links for each subgroup does not result in the best performance of multicast for very different topologies. Subgroup S_1 is a linear topology and subgroup S_0 is like star topology. Their topologies are very different. Optimal results are that S_1 should have less number of links and S_0 should have more number of links. It may as well be that linear topology has larger number of retransmissions than star topology for the same number of links, thus linear topology should have less number of links than star topology so that the whole population reaches a trade-off between two subgroups.

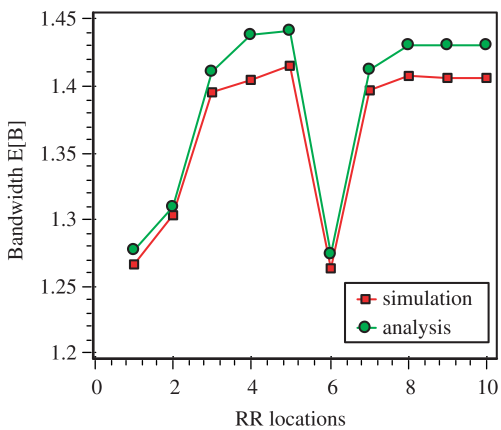


Fig. 17. Bandwidth $E[B]$ of 1 RR for the network with heterogeneous loss probability in Fig. 9.

4.3. The optimal RR locations of heterogeneous packet loss probability

Heterogeneous packet loss probability means that different links may have different packet loss probability. The approach used in this work can be extended to analyze networks with heterogeneous loss probability. Due to the limitation of pages, we mainly focus on homogeneous loss probability to emphasize the effects of topology on multicast performance. More analysis results on heterogeneous loss probability will be considered in the future work. For a multicast network with heterogeneous packet loss probability, the optimal RR locations can be found based on the actual bandwidth consumption. Fig. 17 shows the results for the topology with heterogeneous loss probability in Fig. 9. When RR is in node 1 or 6, the network has small bandwidth consumption.

5. Conclusions

The number of transmissions is a very important parameter. In this work, we have analyzed and simulated the tree-based multicast transmissions and compared simulation results for the number of transmissions with exact and approximate analysis results. We have also obtained the actual bandwidth consumption traversed by multicast transmissions. Based on the actual bandwidth consumption, the effect of topology on optimal RR placements has been investigated for QoS multicast. The multicast group with homogeneous topology should be partitioned to have the same size of each subgroup so that the whole group has the best performance. The same cannot be said for topologies where each subgroup has very different topology. In contribution we have seen that as the number of RRs increases, the optimal RR locations of *k*-ary trees go deeper down in the hierarchy.

References

[1] John C. Lin, Sanjoy Paul. RMTP: A reliable multicast transport protocol. In: Proc IEEE INFOCOM San Francisco, USA: 1996. p. 1414–24.

[2] Christos Papadopoulos, Guru Parulkar, George Varghese. An error control scheme for large-scale multicast applications. In: Proc IEEE INFOCOM San Francisco, USA: 1998. p. 1188–96.

[3] Markopoulou Athina P, Tobagi Fouad A. Hierarchical reliable multicast: performance analysis and placement of proxies. Proceedings of the second international workshop on networked group communication (NGC 2000), November. Palo Alto, USA: ACM Press; 2000. p. 27–35.

[4] Bhagwat Pravin, Mishra Partho P, Tripathi Satish K. Effect of topology on performance of reliable multicast communication. Proc IEEE INFOCOM 1994;2:602–9.

[5] Nonnenmacher Jorg, Biersack Ernst W, Towsley Don. Parity-based loss recovery for reliable multicast transmission. IEEE/ACM Trans Networking 1998;6(4):349–61.

[6] Yamamoto MiKi, Yamaguchi Makoto, Hashimoto Takashi, Ikeda Hiromasa. Performance evaluation of reliable multicast communication protocol with network support. Proc IEEE ICC 2000:1736–41.

[7] Mailhofer Christian. A bandwidth analysis of reliable multicast transport protocols. Proceedings of the second international workshop on networked group communication (NGC 2000), November. Palo Alto, USA: ACM Press; 2000. p. 15–26.

[8] Pingali Sridhar, Towsley Don, Kurose Jim. A comparison of sender-initiated and receiver-initiated reliable multicast protocols. Proceedings of the sigmetrics conference on measurement and modeling of computer systems, May. New York: ACM Press; 1994. p. 221–30.

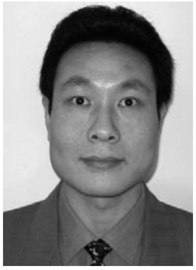
[9] Pavlin Radoslavov, Christos Papadopoulos, R. Govindan, Deborah Estrin. A comparison of application-level and router-assisted hierarchical schemes for reliable multicast. In: Proc IEEE INFOCOM, 2001. p. 229–38.

[10] Zuo Wen Wan, Michel Kadoch, Ahmed Elhakeem. Performance evaluation of tree-based reliable multicast. In: The 12th IEEE international conference on computer communications and networks (ICCCN '03). Dallas TX, USA: October 20–22, 2003. p. 67–73.

[11] Lacher Martin, Nonnenmacher Jorg, Biersack Ernst W. Performance comparison of centralized versus distributed error recovery for reliable multicast. IEEE/ACM Trans Networking 2000;8(2):224–38.

[12] Dan Rubenstein, Sneha Kasera, Don Towsley, Jim Kurose. Improving reliable multicast using active parity encoding services (APES). In: Proc IEEE INFOCOM, 1999. p. 1248–55.

[13] Kasera Sneha K, Kurose Jim, Towsley Don. Buffer requirements and replacement policies for multicast repair service. Proceedings of the second international workshop on networked group communication (NGC 2000), November. Palo Alto, USA: ACM Press; 2000. p. 5–14.



Zuo Wen Wan received the B.S. and M.S. degrees from University of Electronic Science and Technology of China, Ph.D. degree from the department of electrical and computer engineering at Concordia University, Canada. He joined in Avici Systems Inc. USA in 2005. His interest includes multicast communications, routing protocols, and telecommunication networks.

Michel Kadoch is active in research mostly in performance analysis and network management and control in wired as well as wireless networks. He is the director of the research laboratory LAGRIT at ETS. For the last 3 years, he has managed and participated actively in a research program on Multicast in high speed networks sponsored by Bell Canada and NSERC. Professor Kadoch has published many articles and is the author of a book "Protocoles et réseaux locaux" (Édition ETS). He has been involved for many years with the industry namely Teleglobe Canada, CAE, and Communication Canada. He has been a consultant with Bell South, BC Tell, and British Telecom UK, on issues related to installation and optimization of SONET in telephone and data networks.

Ahmed Elhakeem received his Ph.D. degree from the Southern Methodist University, Dallas, TX, in 1979. He spent the next two years working as a Visiting Professor in Egypt, after which he moved to Ottawa, Canada, in 1982. He assumed research and teaching positions at Carleton and Manitoba Universities, and later moved to Concordia University, Montreal, Canada, where he has been a Professor in the Electrical and Computer Engineering Department since 1983. He has published numerous papers in IEEE and in international journals in the areas of spread spectrum and networking. He is a well-known expert in these areas and serves as a consultant to various companies. His current research interests include interconnected wireless LANs, Error correction for IP Multicast, wideband networks, switching architectures, CDMA networks, multi protocols, software Radios, and reconfigurable networks. He is a co-author of the book *Fundamentals of Telecommunications Networks* (New York: Wiley, 1994). He has chaired and organized numerous technical sessions in IEEE conferences, and was the Technical Program Chairman for IEEE Montech '86 in Montreal, Canada. More recently, he was the Key Guest Editor for four issues of the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS on Code Division Multiple Access, namely CDMA I, II, III and IV, appearing in May '94, June '94, October '96, and December '96. Dr. Elhakeem was the Communications chair of IEEE Montreal and the TCCC Rep. to ICC '99. He served as an associate editor for the IEEE Communications Letters Journal (1996–1999). He is a professional engineer in Ontario and a senior member of IEEE. Dr. Elhakeem has served as a consultant to many companies, and to Communication Research Center and Industry Canada.