A RTT-based Partitioning Algorithm for a Multi-rate Reliable Multicast Protocol

Moufida Maimour, Congduc Pham

HAL Id: hal-00096954
https://hal.archives-ouvertes.fr/hal-00096954

Submitted on 20 Sep 2006

A RTT-based Partitioning Algorithm for a Multi-Rate Reliable Multicast Protocol

Moufida Maimour and Cong-Duc Pham

RESO-INRIA, LIP-ENS
ENS, 46 alle d'Italie 69364 Lyon Cedex 07 - France
email:{mmaimour,cpham}@ens-lyon.fr

Abstract. Various Internet applications involve multiple parties and usually adopt a one-to-many communication paradigm (multicast). The presence of multiple receivers in a multicast session rises the problem of *inter-receiver fairness*. Transmitting with a rate which matches the slowest receiver will limit the throughput of other receivers and thus their *satisfaction*. A multi-rate mechanism where the receivers are distributed into subgroups with similar capacities, can improve the inter-receiver fairness for multicast sessions. In this paper, we deal with the problem of receivers partitioning and propose a simple algorithm based on the receivers RTT variations where an explicit estimation of the receivers capacities is avoided. Our partitioning algorithm, although simple, performs an on-the-fly partitioning depending on the receivers' feedback. We show that our partitioning algorithm approximates and in many cases, achieves the optimal solution with a minimum computation effort.

1 Introduction

Various Internet applications involve multiple parties and usually adopt a one-to-many communication paradigm (multicast). The presence of multiple receivers in a multicast session rises the problem of *inter-receiver fairness*. Transmitting with a rate which matches the slowest receiver will limit the throughput of other receivers and thus their *satisfaction*. A multi-rate mechanism can improve the inter-receiver fairness for multicast sessions. The main advantage of a multi-rate scheme is that receivers with different needs can be served at a rate closer to their needs rather than having to match the speed of the slowest receiver. In a multi-rate session, the multicast source can transmit at different rates to different receivers either through a hierarchical scheme (layering) [1,6,7] or a replicated scheme (destination set grouping, DSG) [5]. In layered multicast, each receiver controls the rate at which it receives the data, usually by using multiple multicast groups. The receivers join and leave groups depending on their path congestion state so the amount of data being received is always appropriate. Layering schemes provide more economical bandwidth usage than DSG schemes, however layering is more complicated and requires efficient hierarchical encoding/decoding algorithms and synchronization among different layers.

In both layered and replicated schemes, an explicit or an implicit partitioning of the receivers among subgroups is required. This partitioning is performed so that the receivers satisfaction is maximized. Receivers satisfaction can be quantified using a

utility function as the ones proposed in [3,8]. Determining the optimal grouping for a multicast session is a difficult problem. In [5], the authors propose a number of grouping heuristics that are guidelines for a multicast source to make its splitting decisions. The authors in [8] consider the optimal partitioning of receivers into groups for multi-rate schemes. They formulated, for a general class of utility function, the partitioning problem as an optimization problem to maximize the sum of receiver utilities. They present a dynamic programming algorithm to solve the partitioning problem, and proved that the solution it finds is optimal. Both of the previous partitioning algorithms require the knowledge of the isolated[1] rates of the different receivers which are not easily obtained in the current Internet. Jiang et al. proposed in [4] a special (two-group) DSG protocol to be deployed in the Internet. They proposed a mechanism based on the experienced losses by a receiver to estimate its isolated rate. This mechanism can be used in loss tolerant applications; however the aim of our work is to be able to perform a partitioning in the context of fully reliable multicast applications where we privilege a conservative approach where losses are to be avoided as much as possible.

In this paper we propose a new partitioning algorithm based on the receivers RTT variations instead of their isolated rates. In this context we propose an other formulation of the utility function using the RTT variations instead of the isolated rates. Our partitioning algorithm, although simple, performs an on-the-fly partitioning depending on the receivers' feedback. Our algorithm approximates and in many cases, achieves the optimal solution without using complex computations. For instance, in [2], a computation is performed on every candidate solution before choosing the one that maximizes the receivers satisfaction. The dynamic programming algorithm proposed in [8] requires less computation effort but still be complex. The remainder of this paper is organized as follows. Our new formulation for the utility function is proposed in section 2. The description of our partitioning algorithm with a study of some of its properties are provided in section 3. Some simulation results are presented in section 4 and section 5 concludes.

2 Utility Function

In a multicast session, the satisfaction of a receiver R_i can be quantified using a utility function (or a receiver fairness measure) that maps the reception rate of the receiver to a fairness value normalized to the range $[0.0, 1.0]$. Authors in [3,4] proposed $U_i(r) = \min(r_i, r) / \max(r_i, r)$ as a fairness function, where r_i and r are respectively the isolated rate and the R_i receiver's reception rate. This utility function has a value of 1 if the reception rate equals the receiver isolated rate, non-decreasing in the range $[0, r_i]$ and non-increasing in the range $[r_i, \infty]$. This utility function assumes that the isolated rates of the different receivers are known. We propose an other expression of the utility function using the RTT variation experienced by the receiver as a response to a given transmission rate. The RTT variation is a measure which can be easily obtained in the Internet using a ping-like mechanism.

---

[1] The isolated rate [3] is the rate that a receiver would obtain if unconstrained by the other receivers in the group, assuming max-min link sharing.
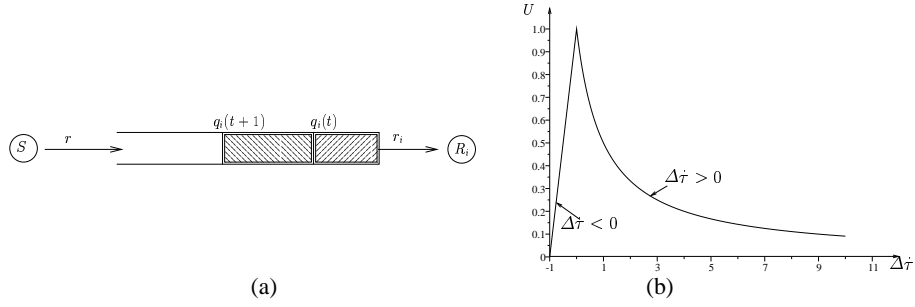
Fig. 1. (a) Queue length variation. (b) The inter-receiver fairness as a function of the relative RTT variation.

For the purpose of defining the utility function, we consider the case of one receiver with r_i as the isolated rate. Let r be the transmission rate of the sender in bits/s. When the transmission rate exceeds the bottleneck rate, a queue of packets will build up within the connection. We suppose that the receiver sends periodically (every T seconds) a probing packet toward the source in order to estimate its RTT to the source. Let $q_i(t)$ be this queue size in packets at time t and $\Delta q_i = q_i(t+1) - q_i(t)$ be the positive or negative variation in the queue length during a given time period T upon the reception of the subsequent probing packet (figure 1a). Noting by S the packet size, the queue builds up when $r > r_i$ during T with $\Delta q_i = (r - r_i)\,T/S$. The RTT variation experienced by the connection during the last T period, $\Delta \tau_i = \Delta q_i\,S/r_i = (r - r_i)\,T/r_i$ giving $\Delta \dot{\tau}_i = (r - r_i)/r_i$ where $\Delta \dot{\tau}_i = \Delta \tau_i/T$ can be seen as the *relative* RTT variation experienced by receiver R_i. We can write $r/r_i = 1 + \Delta \dot{\tau}_i$ and the individual receiver utility can be expressed as a function of the relative RTT variation as follows (figure 1b):

$$U_i(r) = \begin{cases} \frac{1}{1+\Delta\dot{\tau}_i} & \text{if } \Delta\dot{\tau}_i \geq 0 \\ 1 + \Delta\dot{\tau}_i & \text{if } \Delta\dot{\tau}_i \in [-1,0[\end{cases} \tag{1}$$

We note that the utility function is not defined[2] for $\Delta\dot{\tau}_i < -1$ which corresponds to a negative reception rate ($r < 0$). The receiver that experiences a positive RTT variation could experience losses since its reception rate is greater than its isolated rate. In the case of a negative RTT variation, the receiver will be unsatisfied since it has more bandwidth resources. In a similar way to [3], we define the utility function of a single-rate multicast session as the weighted sum of the individual utility values of receivers in the session: $U(r) = \sum_{i=1}^{n} \alpha_i U_i(r)$ subject to $\sum \alpha_i = 1$ and $\alpha_i \in [0,1], i = 1, \ldots, n$ where n is the number of the receivers in the multicast session. A multi-rate multicast session consists of one or more subgroups split from an original multicast session. The session utility function in this case is defined as the summation of the utility values obtained by all multicast subgroups, using the single-rate utility measure in each subgroup. More specifically, if a multicast session of receivers $\{R_1, R_2, \ldots, R_N\}$ is split into K subgroups $\{G_1, G_2, \ldots, G_K\}$ with different transmission rates $g_1, g_2, \ldots, g_K$,

---

[2] The interested reader can refer to [10] for more details.

then $U(g_1, g_2, \ldots, g_K) = \sum_{j=1}^{K} \sum_{i=1}^{n_j} \alpha_{i,j} U_{i,j}(g_j)$, subject to $\sum_{i,j} \alpha_{i,j} = 1$ and $\alpha_{i,j} \in [0,1]$. We have $\sum_j n_j = N$ where n_j is the number of the receivers in subgroup G_j. $U_{i,j}(g_j)$ and $\alpha_{i,j}$ are respectively the utility function and the weight associated to the ith receiver of the jth subgroup. Since we are concerned by a fully reliable multicast, the transmission rate g_j for the receivers in subgroup G_j has to match the minimum rate of the subgroup isolated rates, i.e. $\forall j$, $g_j = \min_{i \in G_j} r_i$

3 Partitioning Algorithm

Given a set of receivers $\{R_1, R_2, \ldots, R_N\}$ with isolated rates $r_1 < r_2 < \ldots < r_N$, the problem consists in splitting this set of receivers into K subgroups (K is less than a maximum number G) to make a partition $P = \{P_0, P_1, \ldots, P_{K-1}\}$ of the receivers so that the overall session utility is maximized. We aim to determine the optimal solution or at least an approximated one (without prior knowledge of the number of subgroups) such that the global receivers utility is greater than a given threshold. Our partitioning algorithm is based on the RTT variation experienced by every receiver and is executed on-the-fly while the source is increasing its rate. The source begins sending packets at a specific minimum rate and increases its rate as long as it does not receive a feedback indicating a relative RTT variation greater than a given threshold $\dot{\epsilon}$ (algorithm 1). The aim of algorithm 1 is to partition the receivers among subgroups of "similar" capacities. Initially, P_0 is the set of all the receivers. The source periodically checks if any receiver from P_0 experienced a $\Delta \dot{\tau}$ greater than a given threshold b; if so P_0 is split up into two subgroups $P_1 = \{R_j \in P_0, \Delta \dot{\tau}_j > a\}$ and $P_0 = P_0 - P_1$. Only feedbacks from P_0 are considered by the source for the purpose of next splittings. The source will continue increasing its rate and would if necessary, split P_0 again into two groups P_0 and P_2. The splitting process will continue until G subgroups have already been built or the P_0 is no more "split-able" (with only one element or members of P_0 have similar capacities).

3.1 The Convergence Criteria

One of the most important convergence criterion of the algorithm is the fact that P_0 is no more split-able. Here we demonstrate that the sufficient condition for P_0 to be no longer split-able is (as stated in algorithm 1):

$$\forall j < N, R_j \in P_0, \qquad \frac{1 + \Delta \dot{\tau}_{j+1}}{1 + \Delta \dot{\tau}_j} \geq \rho . \tag{2}$$

For this purpose, we first consider two consecutive[3] receivers R_j and R_{j+1} with isolated rates $r_j < r_{j+1}$. Let γ be the multiplicative factor by which the source rate could be multiplied at every period. Suppose that at a given time the two receivers belong to the same group, this means that none of their RTT variations is greater than the b threshold: $\Delta \dot{\tau}_j = (r - r_j)/r_j < b$ and $\Delta \dot{\tau}_{j+1} = (r - r_{j+1})/r_{j+1} < b$. At the

---

[3] we consider two receivers R_i and R_j as *consecutive* (in this order) if their corresponding isolated rates r_i and r_j satisfies $r_i < r_j$ and $\forall k, r_k < r_i < r_j$ or $r_i < r_j < r_k$.

Algorithm 1 Partitioning Algorithm. Input: P_0, a set of receivers. Output: a partition $\{P_0, PK-1, \ldots, P_2, P_1\}$ where $K \leq G$

Require: $N > 1$ and $a < b < \dot{\epsilon}$
  $\rho \leftarrow \frac{a+1}{b+1}$
  the source rate $r \leftarrow r_{min}$
  $P_0 \leftarrow \{R_j, j = 1, \ldots, N\}$, the set of all the receivers
  $i \leftarrow 1$
  Periodically,
  if $\exists j, R_j \in P_0$ such that $\Delta\dot{\tau}_j > b$ **then**
    $P_i \leftarrow \{R_j \in P_0, \Delta\dot{\tau}_j > a\}$
    $P_0 \leftarrow P_0 - P_i$
    $i \leftarrow i + 1$
    if $\exists j, R_j \in P_0$ such that $\Delta\dot{\tau}_j > \dot{\epsilon}$ **then**
      decrease r
    else
      increase r
    end if
  end if
  until $i = G$ or $|P_0| = 1$ or $\forall j < N, R_j \in P_0, \frac{1+\Delta\dot{\tau}_{j+1}}{1+\Delta\dot{\tau}_j} \geq \rho$

---

next period the source rate becomes γr rather than r. Receivers R_j and R_{j+1} will no longer continue to be in the same subgroup if their new relative RTT variations $\dot{\tau}_j{}^+$ and $\dot{\tau}_{j+1}^+$, satisfy: $\Delta\dot{\tau}_j{}^+ = (r - r_j)/r_j > b$ and $\Delta\dot{\tau}_{j+1}^+ = (r - r_{j+1})/r_{j+1} < a$. Hence $(b+1)r_j < \gamma r < (a+1)r_{j+1}$ and $r_j/r_{j+1} < (a+1)/(b+1) = \rho$ which is the necessary condition so that the two receivers do not belong to the same group according to our partitioning algorithm. Consequently, the sufficient condition for the two receivers to belong to the same group is that the ratio of their isolated rates is greater or equals ρ, i.e. $r_j/r_{j+1} \geq \rho$. Since the isolated rates are not available, using the relative RTT variation instead, the sufficient condition becomes:

$$\forall j \in 1..N - 1, \frac{r_j}{r_{j+1}} = \frac{1 + \Delta\dot{\tau}_{j+1}}{1 + \Delta\dot{\tau}_j} \geq \rho . \tag{3}$$

Algorithm 1 converges if G groups are already constructed or when the remaining receivers (i.e. those in P_0) have so similar capacities such that a split is no more possible. This depends on the a and b parameters as stated by condition (2), which is a generalization of (3) for every two consecutive receivers in P_0 (assuming of course that $|P_0| > 1$).

3.2 Lower Bound Guarantee on the Utility Function

We present here an interesting feature of our partitioning algorithm which consists in assuring a lower bound on the resulting utility value depending on the algorithm parameters. We consider one subgroup $G_j, j = 1, \ldots, K$ containing the receivers $R_{j1}, R_{j2}, \ldots R_{jn_j}$ with isolated rates $r_{j1} < r_{j2} < \ldots < r_{jn_j}$. The reception rate

of the G_j's receivers is $g_j = \min_{i \in G_j} r_i = r_{j1}$. Without loss of generality, we put $\alpha_{i,j} = 1/N, \forall i, j$, where N is the number of receivers in the whole multi-rate multicast session. The overall session utility satisfies[4]:

$$U(g_1, g_2, \ldots, g_K) \geq \frac{1}{N} \sum_{j=1}^{K} \frac{1 - \rho^{n_j}}{1 - \rho} \ . \tag{4}$$

This last relation (4) is very interesting since it provides a lower bound on the session utility function independently of the isolated rates distribution. It is quite obvious that if we want to have a higher utility value, we have to choose a higher ρ. It is worth noting that the number of subgroups is proportional to this parameter.

3.3 Numerical Results

For numerical results about the lower bound guarantees on the utility function, we consider the particular case when the number of receivers in each subgroup is the same i.e. $\forall j, n_j = N/K$. In this case, inequation (4) becomes:

$$U(g_1, g_2, \ldots, g_K) \geq \frac{K}{N} \frac{1 - \rho^{N/K}}{1 - \rho} = U_{min} \ . \tag{5}$$

Figures 2a,b,c plot the minimum utility value (U_{min}) as a function of the number of the built subgroups for 12 and 48 receivers and for different values of the ρ parameter ($0.8, 0.85, 0.9$, and 0.95). Figure 2a shows for $\rho = 0.9$, a minimum utility of 0.8 with just 2 subgroups which corresponds to 80% of the maximum utility achieved with a partition of one receiver per subgroup. For 48 receivers with $\rho = 0.9$, only 4 subgroups are sufficient to achieve a minimum utility of 80%. Figure 2c shows the minimum utility as a function of the number of subgroups for different numbers of the receivers ($\rho = 0.9$). We can see that for 6 receivers the minimum utility is at least 80% which increases with the number of subgroups. For 48 receivers, we need approximately 3 and 10 subgroups to assure a minimum utility of 50% and 80% respectively. Increasing the number of receivers increases the required number of subgroups to achieve a minimum receivers satisfaction. An other observation is that increasing the number of subgroups improves the utility value. Figures 2d,e plot the utility gain as a function of the number of subgroups for 12 and 48 receivers respectively. We can see that for 12 receivers and $\rho = 0.8$, with two subgroups we already have a gain of 160% and with 4 subgroups, the gain is about 210%. We can note that independently of ρ, the gain does not increase significantly when increasing the number of subgroups. For instance (figure 2e), with 5 subgroups we can achieve a gain of 2 for $\rho = 0.8$. If we double the number of subgroups from 5 to 10, the gain is only improved by 0.1.

4 Simulation Results

In order to get an insight into the proposed partitioning algorithm, simulations are performed and for a given number of receivers, the isolated rates are randomly generated

---

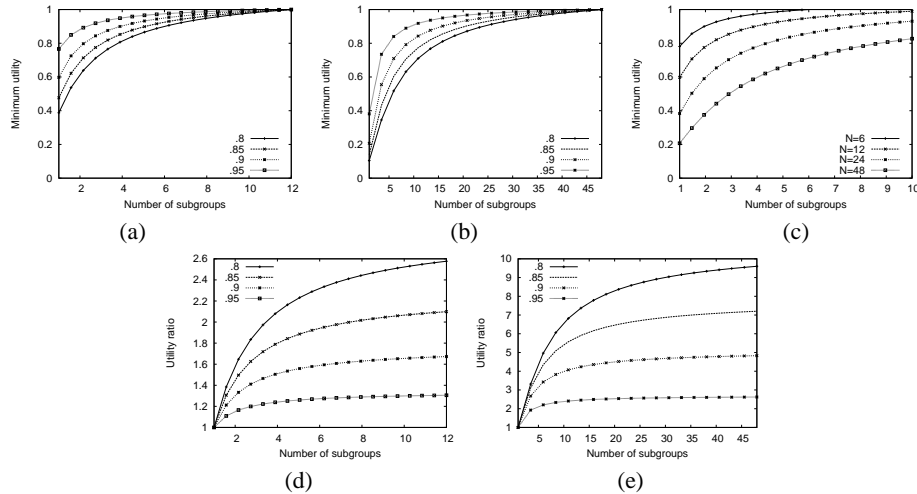[4] For the demonstration, refer to [10]

Fig. 2. Lower bound on the session utility (a) 12 receivers, (b) 48 receivers, (c) $\rho = 0.9$, varying the number of the receivers. Utility gain (d) 12 receivers, (e) 48 receivers

following a uniform distribution with different parameters. Our partitioning algorithm is then applied on the resulting set of generated rates and this process is repeated for different values of ρ. A large number of simulations are performed for every set of parameters in order to get stable results. For every simulation, the number of subgroups built is recorded with the corresponding utility value.

4.1 Comparing with the Optimal Solution

In this section, we are comparing the performances of our partitioning algorithm to those of an algorithm that produces the optimal solution. The optimal algorithm consists in computing for each possible partition the corresponding utility value. Then the minimum utility value is saved in order to be compared with the utility achieved by our partitioning algorithm. A large number (about 1500) of simulations have been performed on two sets of rates. The first set consist in rates generated following a uniform distribution between 5 and 10. The second set presents a higher degree of heterogeneity with a uniform distribution between 5 and 55. We have limited ourselves to a partition of only two subgroups.

Figure 3a shows that for the first set of trials, 35% of the trials achieved the optimal solution (this corresponds to the box centered in 0). The three other boxes correspond to the percentage of trials for which the obtained utility is 95%, 90% and 85% of the optimal solution. An interesting remark is that all the obtained utility values are at least 85%. The results of increasing the heterogeneity (the second set of simulations) are shown in figure 3b. In this case, the partitioning algorithm gives a larger range of utility values. The reason is mainly the fact that our splitting is performed on-the-fly and setting the maximum number of subgroups to 2 limits the performance of the algorithm. Actually the algorithm could take the decision to split the original set of receivers without having

an accurate estimation of the capacity of the remaining receivers. This behavior depends on the ρ parameter which has to be well chosen to reflect the heterogeneity degree. Figure 3c is an other way to compare the two sets of simulations for which the cumulative percentage of success is plotted as a function of the decreasing degree of approximability to the optimal solution. We have at least 35% and 9% of trials with the optimal solution for the first and second sets respectively. All of the trials achieve at least 80% and 55% of the optimal utility value for the first and second set of simulations.
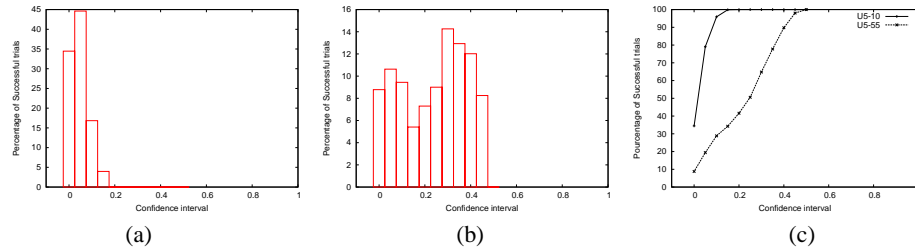

Fig. 3. Solution optimality. Uniform distribution between (a) 5 and 10, (b) 5 and 55. (c) comparison of the two sets of results.

4.2 Varying the number of subgroups

All the figures of this section show the different points that correspond to the resulting pairs of utility value (or the utility gain) and the number of subgroups. The utility gain is computed using the case where there is only one subgroup as a reference. The curves of figure 4a1 and figure 4a2 show the results of the execution of our algorithm with 12 receivers for which the isolated rates are uniformly distributed between 5 and 10. Figure 4a1 confirms the numerical results. The minimum utility according to figure 2a is 0.4 and 4a1 shows a minimum utility value of approximately 0.6. We can also see that with only 2 subgroups we can achieve a mean utility of 0.9. In figure 4a2, we can see the influence of increasing the number of subgroups. We note that with 2 subgroups, we can increase the receivers' satisfaction by 25%. Increasing the number of subgroups beyond 3 does not increase significantly the receivers' satisfaction (a maximum gain of 1.4 instead of 1.3 with 10 subgroups!). We have similar results for 48 receivers (figure 4b1-b2).

Figure 5 presents the results of having more heterogeneous receivers whose isolated rates are uniformly distributed between 5 and 55. Compared to the first set, with the same number of subgroups, we note that the utility value decreases due to the higher heterogeneity degree. However, we still achieve 0.8 with only 3 subgroups for 12 receivers and with 8 subgroups for 48 receivers. Once again, increasing the number of subgroups beyond a given threshold does not provide significant improvements. We can conclude that independently of the receivers heterogeneity degree, we do not need to increase the number of subgroups beyond a given threshold which is proportional to the receivers heterogeneity degree.
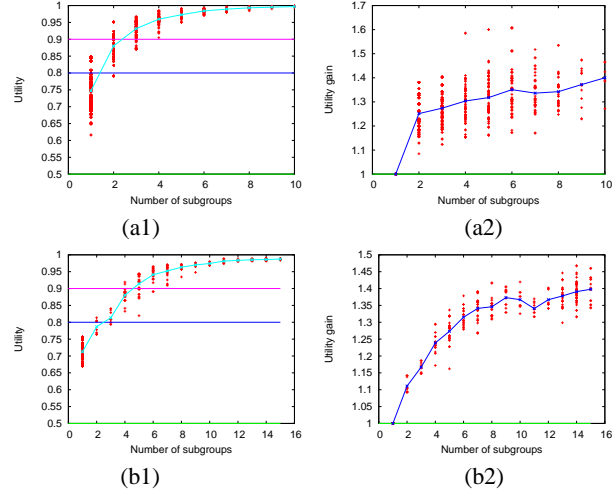
Fig. 4. Utility and utility gain for receivers' isolated rates distributed uniformly between 5 and 10. (a) 12 receivers, (b) 48 receivers.

5 Conclusion

In this paper, we proposed a simple partitioning algorithm which does not require the knowledge of the receivers' isolated rates. Our algorithm performs an on-the-fly partitioning algorithm as soon as it receives feedback from the receivers. The knowledge of the RTT variation experienced by every receiver is required but there is no assumption on how the RTT variations are measured (therefore a simple ping method is suitable). The partitioning algorithm can be used by a multi-rate protocol (with a layered or replicated scheme) in order to adapt its number of layers or rates so that the global receivers' satisfaction is improved. One of the nice features of our algorithm is that it assured a minimum utility value depending on the value of ρ. We have shown that our algorithm converges, or at least approximates the optimal solution with a minimum computation effort. A future direction will consist in a deeper study of our algorithm in order to know how its parameters can be chosen according to the receiver's heterogeneity degree.

In this paper, we did not consider the partitioning dynamics due to space limitation. We suppose that the algorithm converges rapidly so that the initial partitioning is not disturbed by receivers changing their isolated rates. After the initial partitioning was performed, if any receiver experienced a RTT variation such that $|\Delta\dot{\tau}| > \dot{\epsilon}$ for a sufficiently long period, then a decision to move this receiver to a lower or a higher subgroup could be taken. Moreover, we have considered that the source executes the partitioning algorithm. In order to be more scalable, solutions with routers contribution seem to be very promising. We studied the possibility of executing the partitioning algorithm by the routers instead of the source. In this case if we choose a replicated scheme, for the implementation of our approach, we can distribute the data replication burden among some receivers instead of overwhelming the source [9].
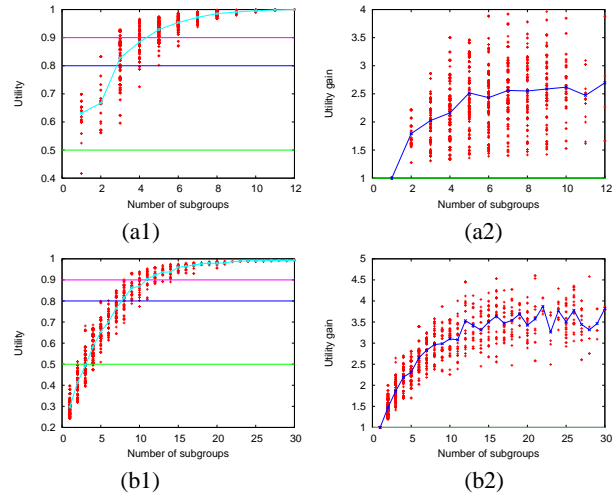
Fig. 5. Utility and utility gain for receivers' isolated rates distributed uniformly between 5 and 55. (a) 12 receivers, (b) 48 receivers.

References

1. L. Vicisano at al. Tcp-like congestion control for layered multicast data transfer. In *INFO-COM*, pages 1–8, San Fransisco, USA, 1998.
2. M. D. Amorim et al. Improving user satisfaction in adaptive multicast video. *IEEE/KICS Journal on Communications and Networks*, 4(3), Sep. 2002.
3. T. Jiang et al. Inter-receiver fairness: A novel performance measure for multicast ABR sessions. In *SIGMETRICS*, pages 202–211, June 1998.
4. T. Jiang et al. Inter-receiver fair multicast communication over the internet. In *NOSSDAV*, pages 103–114, June 1999.
5. T. Jiang et al. On the use of destination set grouping to improve inter-receiver fairness for multicast abr sessions. In *IEEE INFOCOM'00*, March 2000.
6. V. Jacobson et al. Receiver-driven layered multicast. In *ACM SIGCOMM'96, Stanford, CA*, pages 117–130, August 1996.
7. X. Li et al. Layered video multicast with retransmission (lvrm): Evaluation of hierarchical rate control. In *INFOCOM'97*, 1997.
8. Y. Yang et al. Optimal partitioning of multicast receivers. Technical Report TR-2000-10, Univ. of Texas at Austin, May 2000.
9. M. Maimour and C. Pham. Amca : an active-based multicast congestion avoidance algorithm. Technical Report RR-4689, INRIA, January 2003. Also available as a LIP/ENS Research Report under 2003-07.
10. M. Maimour and C. Pham. A rtt-based partitioning algorithm for a multi-rate reliable multicast protocol. Technical Report RR-4779, INRIA, March 2003.