

SCALABLE MULTIPLE DESCRIPTION CODING FOR VIDEO DISTRIBUTION IN P2P NETWORKS

J. R. Taal, J. A. Pouwelse, R. L. Lagendijk

Information & Communication Theory Group

Faculty of Electrical Engineering, Mathematics and Computer Science,

Delft University of Technology,

The Netherlands

{J.R.Taal,J.A.Pouwelse,R.L.Lagendijk}@ewi.tudelft.nl

ABSTRACT

In this paper we present and evaluate a method for video distribution through a peer-to-peer (P2P) network using Multiple Description Coding in a multicast tree. Our method solves the problem of serving many clients with heterogeneous bandwidths. Every node forwards descriptions to other nodes. Nodes that have sufficient bandwidth are then able to receive more descriptions, and thereby to increase the quality. We use a combination of a layered video codec and erasure coding to generate multiple descriptions. Given different client-bandwidth distribution functions, we optimize a fairness criterion such as the (average) quality for all clients. Since we optimize the average quality for all clients as a group, the performance for individual clients may be very different. We discuss the reduction in quality for clients that have bandwidths other than average as opposed to a single description stream. Finally, we discuss the efficiency, in terms of bit rate and redundancy, for all clients as a group, and individually.

1. INTRODUCTION

Video streaming over the Internet is becoming increasingly popular with news sites and broadcasting organizations offering complete TV-programs on their websites. However, the visual quality of these streams is often not very good, especially the frame rate is not comparable with real TV. For efficient streaming to many clients, multicast is a promising solution to handle the workload.

There are two approaches to multicasting [1], namely multicasting at the IP level and at the application level. The first approach uses Internet routers which send out a single IP packet to multiple receivers, the second uses end-nodes to forward IP packets. IP-level multicasting involves modifying the Internet infrastructure. Routers have to maintain state and best-effort delivery is no longer sufficient. Application-level multicasting does not suffer from these drawbacks, but the performance cannot be as good as IP level multicast in terms of both bandwidth and

latency. We investigate application-level multicasting to a large number of broadband Internet users.

In a simple multicast system, a single compressed video stream is used, with a fixed data rate [2], [3], [4]. For most broadband Internet users, the height of their monthly subscription fee determines their maximum Internet upload/download bandwidth. We define the *rate matching* problem as the inability of a video distribution system of matching the video data rate with the receivers' current Internet bandwidth. The ideal system continuously adapts the video stream data rate to exactly match the maximum bandwidth of each receiver.

In this paper we focus on the rate matching problem. The novel method presented here uses an estimation for the bandwidth of the nodes to calculate the optimal allocation of the rates for all layers. We call this method "Multiple Rate Support for heterogeneous client bandwidths in a P2P context". We use Multiple Description Coding to generate a number of descriptions. Clients are then able to select the number of descriptions they want to receive according to their own bandwidth.

In Section 2, we formulate the problem and present our solution. The models we used and several optimization criteria are discussed in Section 3. In Section 4 we will investigate the behavior of this system under different assumptions regarding the client bandwidths. We end this paper with a Discussion (Section 5) and some concluding remarks.

2. MULTIPLE DESCRIPTION CODING FOR APPLICATION LEVEL MULTICAST

A well-known solution to the rate matching problem is to offer layered video[5]. Clients with higher bandwidth request more layers and therefore achieve a higher image quality than the clients with less bandwidth. This, however, poses another problem for the multicast tree: clients that have received only a few layers, are not able to forward higher layers. This means that any client that is not an end-

node¹ or leaf of the multicast tree, should in fact receive all layers. Consequently, losing a lower layer (more important) will have serious implications for the quality of all clients that depend on this node.

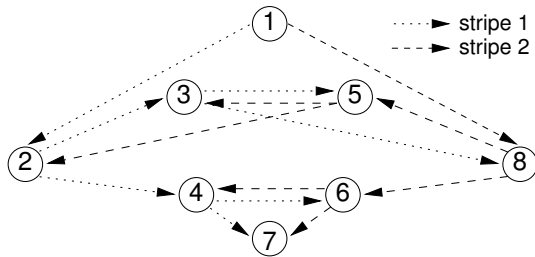


Fig. 1. Example of p2p layout with two multicast overlays.

Splitstream is an application-level multicast system that uses principles which can also be applied to solving the rate matching problem [6]. Splitstream splits the video stream into two pieces, called stripes. The most simple case is creating a separate stripe for the even frames and another stripe for the odd frames. Splitstream uses such a system to construct a multicast tree, as is shown in Figure 1. The source node (1) produces two stripes. The multicast tree for stripe 1 is split at node 2 to nodes 3 and 4 and then to nodes 5, 6, 7 and 8. The multicast tree for stripe 2 splits at node 8 to nodes 5 and 6 and then to nodes 2, 3, 4 and 7.

In this paper we propose to employ Multiple Description Coding[7], [8] to alleviate the problem of layer dependencies in the multicast tree. Each client receives a number of descriptions according to its bandwidth. The crucial difference with layered coding is that each description can be independently decoded. The higher the bandwidth, and consequently the more descriptions can be received, the higher the quality. Since there is no dependency between different descriptions, non-leaf clients can still exchange descriptions, without actually having *all* descriptions. This setup resolves the problems of rate matching, node failure, and packet loss. The solution, however, poses an additional burden on the multicast tree formation. Instead of having only one multicast tree, we now need different multicast trees for all descriptions. Since the number of descriptions will be limited, we do not see this as an unsurmountable hurdle.

The p2p solution with different multicast trees for each description, also introduces another kind of error resilience. When nodes fail, or data is lost, there is a limited effect on the quality. First, because other descriptions can still be used to reconstruct a reasonable image. Secondly, because a failing node will not generate a cascade effect, since other nodes have redundant information available.

The MDC encoding of each frame consists of layered coding into M layers, followed by Multiple Description

packetization and erasure coding of the data. The lower layers will be better protected by the erasure coding than the (less important) higher layers. This scheme of adding erasure codes to generate descriptions is similar to the work of Padmanabhan et al. [9] and Mohr et al.[10]. In our system, MD packets are formed by applying the erasure code to each layer separately, and then splitting up the resulting data into M descriptions. The more descriptions that are received, the more layers can be successfully decoded. This method has several advantages. First, we have control over the rate of each layer, and therefore over the amount of redundancy. Second, this method is easily extended to a larger number of descriptions. Most methods specify how to generate two description but it is not trivial how to generate more than two descriptions.

Suppose the video is encoded in M descriptions. By offering a large number of descriptions, each client has freedom to choose which peers to receive descriptions from, but also have freedom in the number of descriptions it wants to receive. Each client chooses the number m out of M descriptions it wants to receive, according to its bandwidth. We assume that we know the distribution function of the bandwidth of our clients, for example a distribution as shown in Figure 3(a), where clients with high bandwidths are present but not common. Analyzing such a distribution, we can predict that most of the clients we want to serve have sufficient bandwidth to receive m_0 out of M description (each of size R_D). This means that when designing our MDC coder, we may assume that receiving less than m_0 out of M does not happen or is uncommon. This also means that although we add redundancy by MDC, we do not have to spend much redundancy for the cases that less than m_0 descriptions are received.

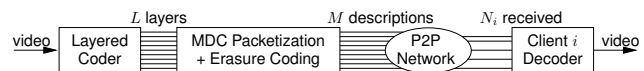


Fig. 2. Block diagram for Multiple Description coding in a p2p context

3. MULTIPLE RATE SUPPORT FOR HETEROGENEOUS CLIENT BANDWIDTHS

A. Multiple Description and Packetization scheme

Our Multiple Description Coder consist of a layered coder and a packetizer, as shown in Figure 2. The layered coder generates a base layer and $(M - 1)$ enhancement layers, such that for decoding layer L_l , first all layers L_1, \dots, L_{l-1} should be decoded successfully. We generate multiple descriptions by combining erasure codes of each layer into packets. By using erasure codes such as Reed Solomon codes, we can ensure that when a client receives only a subset of the descriptions, it can still successfully decode a number of layers. As shown in Figure 3(c), each description contains information of all layers using

¹We refer to clients and nodes interchangeably.

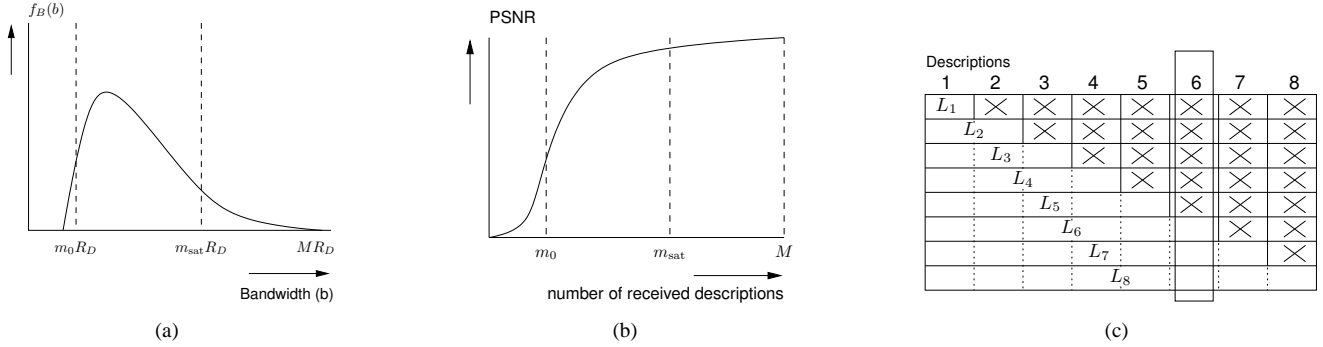


Fig. 3. (a) Example of a PDF of the bandwidth of all clients. (b) Model for quality vs. number of received descriptions. (c) Example of the packetization scheme for $M = 8$. Each column represents one description. Horizontally, each layer L_l consist of l data blocks which are protected by an erasure code with $M - l$ erasure code blocks, represented by (X)

erasure codes. For instance, Layer 1 is just copied in all descriptions. An intermediate layer L_l is first split up in l equal data sized blocks and then the remaining $M - l$ erasure code blocks with the same size are generated. Any number l of these blocks is then sufficient to regenerate the original l data blocks of this layer. The last layer is split up such that all descriptions are needed to reassemble that layer. The total number of layers M is equal to the number of descriptions. The rates for each layer can vary and therefore the amount of data in each row.

This scheme is very flexible since all integral combinations of the number of data blocks and the number of erasure code blocks are available. Since the rate of each layer is not fixed, we still have control over the redundancy as added by the erasure codes. For instance, by using a small base layer and large enhancement layers, the total redundancy will be low.

A layer L_l is encoded at rate R_l . The total rate of each description R_D is then

$$R_D = \sum_{l=1}^L \frac{R_l}{l}$$

Let $D(R)$ be the rate distortion function of our video coder. Similar to layered coding, Multiple description coding comes with a cost, due to reduced efficiency and redundancy in the descriptions. Furthermore, due to the fact that clients only receive an integral number of descriptions they often do not fully use their bandwidth. For now, we assume that the layering has no effect on the coding efficiency.

Each client with bandwidth b receives $m = \lfloor \frac{b}{R_D} \rfloor$ descriptions. After decoding these m descriptions, this client then has distortion

$$D \left(\sum_{l=0}^{\lfloor \frac{b}{R_D} \rfloor} R_l \right), \text{ with } R_0 = 0.$$

B. Fairness criteria

In this paper we try to optimize the performance of the system for all clients. Since many individual clients are contributing to this system, and all have to deal with some redundancy or reduced efficiency, we want to optimize a measure of fairness. This fairness criterion should weigh the cost and the gain for each client. Since any choice for such a fairness criterion is rather arbitrary, we will evaluate several criteria.

Let $f_B(b)$ be a probability distribution of the bandwidth b^2 of a population of clients. Figure 3(a) shows an example of such a distribution. Given such $f_B(b)$, we now want to optimize the performance given by this fairness criterion. The parameters to be optimized are the number of descriptions M and the rates R_1, R_2, \dots, R_M .

$$\max_{M, R_1, \dots, R_M} \text{crit}(f_B(b), D(R)) \quad (1)$$

given $D(R)$ and $f_B(b)$

The first fairness criterion we introduce is the (inverse of) average distortion over all clients:

$$\text{crit}_1 : d_{\text{av}}^{-1} = \left(\int_0^B f_B(b) D \left(\sum_{l=1}^{\lfloor \frac{b}{R_D} \rfloor} R_l \right) db \right)^{-1}. \quad (2)$$

When we try to optimize the perceptual video quality for all clients, it may be worthwhile looking at the average PSNR over all clients³:

$$\text{crit}_2 : \text{PSNR}_{\text{av}} = \int_0^B f_B(b) \text{PSNR} \left(\sum_{l=1}^{\lfloor \frac{b}{R_D} \rfloor} R_l \right) db. \quad (3)$$

²rate R an bandwidth b are interchangeable and both are measured in [bps]

³We take the average over all individual PSNRs instead of the PSNR using the average distortion

As we will see in the next section, these two criteria do not take into account inefficiency or redundancy or any fairness. For example all clients with bandwidth between mR_D and $(m + 1)R_D$ receive the same information and have the same quality, but do not achieve the same efficiency, because for clients with bandwidth $b > mR_D$, $(b - mR_D)$ is left unused.

We define an efficiency metric eff for a client with bandwidth b , which incorporates the unused bandwidth. We divide the quality for each client with quality that the client would have had when it received a single description video at the full available bandwidth. This means that clients that are able to use all their available bandwidth will have a higher efficiency than clients with the same quality but with unused bandwidth.

$$eff(R_0, \dots, R_L, b) = \frac{PSNR\left(\sum_{i=1}^{\lfloor \frac{b}{R_D} \rfloor} R_i\right)}{PSNR(b)}$$

$$crit_3 : \int_0^B f_B(b) eff(R_0, \dots, R_L, b) db \quad (4)$$

C. Evaluation

Optimizing the three aforementioned criteria, is in fact a rate allocation problem. We try to find the rates for each layer that while taking into account the packetization scheme, yields the lowest distortion or highest PSNR, depending on the criterion. We know that this kind of optimization can become quite hard, especially when the R-D function depends on the rate of lower layers. For now we assume that the R-D function does not depend on the layering, and there is no negative effect on the quality by generating more than one layer. The R-D function we use is an average R-D function for a short sequence at SDTV-resolution using the Dirac[11] video encoder.

In the next section we evaluate our method with several client bandwidth distributions. We optimize (1) with each criterion separately and compare the average distortion, average PSNR, the average efficiency and the average redundancy in each case.

The redundancy is defined as the difference between the total bandwidth a client uses to receive as much descriptions as possible and the rate that would be necessary to achieve the same quality with a single description coder. This redundancy contains both the overhead used by the erasure codes and the rates of the unused layers: $\rho = mR_D - \sum_{i=0}^m R_i$. The relative redundancy of a client with bandwidth b and number of received descriptions m is defined by:

$$\rho(b)\% = \frac{mR_D - \sum_{i=0}^m r_i}{b} 100(\%).$$

The average redundancy is then:

$$\rho_{av}(\%) = \int_0^B f_B(b) \frac{mR_D - \sum_{i=0}^m R_i}{b} db 100(\%).$$

4. RESULTS

Optimizing (1) consists of optimizing the number of descriptions M and the rates of the layers R_i . For simplicity we first fix M to 8 descriptions. Later we will discuss optimization of M .

Figure 4(a) shows the results for a Gaussian client distribution with the average-distortion criterion (d_{av}). The vertical lines indicate the partitioning of the client-bandwidth distribution into selections of clients that receive the same number of descriptions. Clients with bandwidth $b > MR_D$, all receive M descriptions. Figure 4(b) shows for the same case, the achieved quality for clients with bandwidth b . As we can see, there are only four quality levels (besides the 13.5dB layer, which is fact the zero-rate quality), this is because layers one to four have gotten zero rate assigned and only layers five to eight have rate $R_i > 0$.

This is a good example of the flexible design of the packetization structure; only certain ratios of data – erasure code are needed for this optimal allocation. This means that although we have $M = 8$ descriptions, we only need to encode 4 layers. We observed that this criterion puts in much effort to decrease the distortion for clients with low bandwidth while clients with more bandwidth have a marginally higher quality. The reason that this criterion treats all clients with bandwidth $b > 2000$ the same is twofold. First, at this point ($b = 2000$ kbps) the R-D function saturates, and increasing the PSNR for clients with high bandwidth costs more than for clients with less bandwidth. Secondly, since an increase of the quality for clients with higher bandwidths also increases the R_D , this negatively affects the quality of all other clients.

This consideration led us to look at the average PSNR, which might perceptually be favorable over average distortion. Figures 4(c) show the result for the $PSNR_{av}$ criterion. This optimization criterion gave a surprising result, since it only assigned rate to the last layer. Furthermore, this $PSNR_{av}$ criterion still does not take into account the fact that clients do not use all their bandwidth effectively. In Figure 4(d) the results for the efficiency criterion eff_{av} are shown. Our eff_{av} obtains a similar average $PSNR$, but still uses only one layer. We conclude that this method does not work well for this kind of distribution. Mainly because a very wide range of bandwidths has to be supported.

Table I summarizes the average distortion, PSNR, efficiency and redundancy for all three criteria. It also shows that the eff_{av} also obtains to lowest redundancy of the three cases. This is not surprising if we recall (4), where we find $PSNR(b)$ in the denominator.

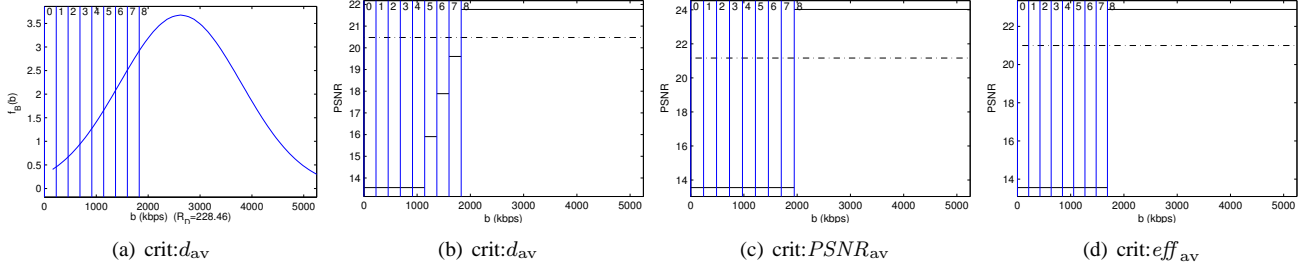


Fig. 4. Optimal rate allocation. (a,b,c) partitioning of the bandwidth distribution for different criteria. (d) the quality for clients depending on their bandwidth using the d_{av} criterion.

Client distribution	Optimization criterion	Metric			
		d_{av}	$PSNR_{av}$	eff_{av}	ρ_{av} (%)
gauss	d_{av}	(*)0.26	20.5	0.46	26.4
	$PSNR_{av}$	0.34	(*)21.2	0.45	21.7
	eff_{av}	0.30	21.0	(*)0.53	(*)16.0
exp	d_{av}	(*)0.24	21.6	0.32	37.1
	$PSNR_{av}$	0.33	(*)22.9	0.25	32.7
	eff_{av}	0.32	22.8	(*)0.36	(*)27.7

Optimal values are annotated with (*).

TABLE I

AVERAGE DISTORTION, PSNR, EFFICIENCY AND REDUNDANCY.

The distribution shown in Figure 5(a) gives other results. It shows an exponential client distribution with the optimal partitioning using the d_{av} criterion. The corresponding quality curve is shown in Figure 5(b). In this case there are six, more or less constantly increasing PSNR levels. The $PSNR_{av}$ criterion (5(c)) gives a somewhat higher average PSNR, but only yields two different quality levels. With the eff_{av} criterion (5(d)), there are three quality levels, where only the bulk of clients with bandwidth $b > 3000$ gets a significant high quality. In Figure 5(e) we show the efficiency for each client. In this case, only the bulk has a high efficiency, but clients with a low bandwidth gain relatively little from joining this P2P video distribution. In Figure 5(f), we show the redundancy for each client (not normalized with the client distribution). Especially clients with $b \approx 2500$ are paying a lot in terms of redundancy.

Distribution	Criterion	M	d_{av}	$PSNR_{av}$	eff_{av}	ρ_{av}
exp	d_{av}	4	0.24	21.6	0.40	25.8
		8	0.24	21.6	0.30	37.4
		12	0.24	21.6	0.34	38.6
		16	0.24	21.6	0.06	48.8

TABLE II

THE AVERAGE DISTORTION, PSNR, EFFICIENCY AND REDUNDANCY FOR VARYING M .

So far we have only looked at optimizing the rates R_1, \dots, R_M . We now look at finding an optimal M . Although we presented the optimization of M and rates R_l separately, in general these parameters cannot be optimized

independently.

We used the “exponential” distribution and varied the M . The results of the rate allocation and optimization of the d_{av} criterion are shown in Table II. It shows that increasing the number of layers, not necessarily decreases the average distortion. On the other hand, the number of quality levels does increase with the number of descriptions. Furthermore the d_{av} criterion does not take into account redundancy (and does not try to minimize that), therefore the efficiency drops with increasing M . We predict that other criteria, definitely will benefit from an increasing M .

Some remarks regarding the optimization of M : First, the number of descriptions will probably be bounded by the capabilities of the underlying P2P network. Secondly, we have neglected the inefficiency by layered coding. It is not likely that a high number of layers is feasible without a great loss of coding efficiency.

5. DISCUSSION & CONCLUSION

We investigated the use of Multiple Description Coding in the case where multiple clients with heterogeneous bandwidths receive a number of descriptions according to their bandwidths. We are dealing with many clients that are connected through a peer-to-peer network. Although this is a typical case where layered coding can offer the wanted scalability, we assumed that the P2P network has no way of guaranteeing the delivery of lower layers. In that case Multiple Description Coding may be the only way to offer scalable video to these clients. Since all clients pay in terms of bandwidth and redundancy, our method uses a fairness criterion to ensure that the clients are fairly treated. We have shown different bandwidth distributions, and the resulting qualities for all clients. One of the results is that in some cases, many layers will have a rate of zero, which in practice means that we have less layers/quality levels than descriptions. This also means that the number of quality levels is not as high as we expected, which was one of the goals of this method. On the other hand, if the results of the optimization yield a low number of quality

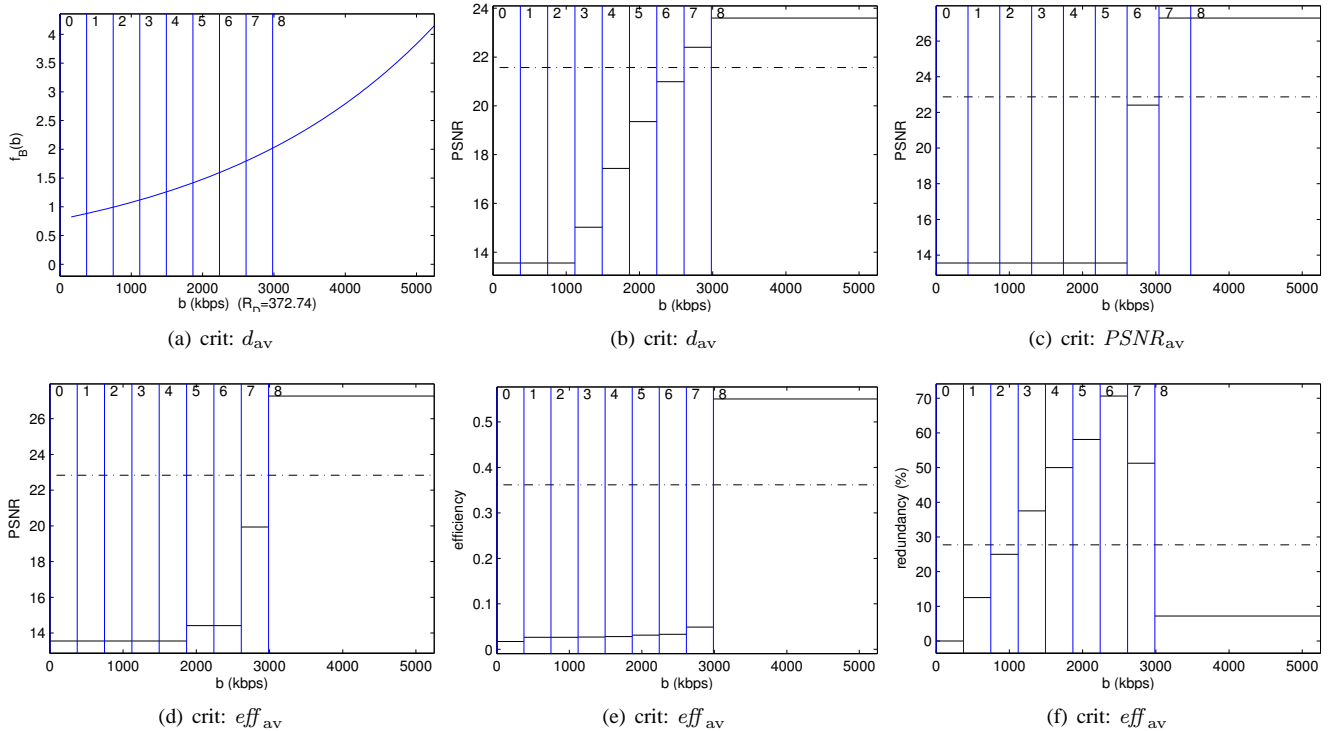


Fig. 5. Optimal rate allocation for the “exponential” distribution, using the eff_{av} criterion (a,b,d,e), d_{av} criterion (c) and $PSNR_{av}$ criterion (f). (a) shows $f_B(b)$ with the R_D partitioning. (b,c,f) shows the PSNR for individual clients depending on their bandwidth. (d) shows the efficiency for each client. (e) shows the relative redundancy for each client

levels, apparently there is no need for more quality levels to yield a fair rate allocation.

For exponentially increasing distributions with a cutoff at a certain bandwidth, or similar distributions, the optimal allocation generates different layers (and different quality levels). By using a different fairness criterion we are able to take into account the efficiency for each individual client.

Concluding, we feel this method deserves further investigation. First, there may be other fairness criteria, that better weigh the interests of individual clients. Furthermore, to better support other distributions, we may have to look into asymmetrical MDC, where descriptions not necessarily have the same size, and descriptions may contain different amounts of data from each layer.

REFERENCES

- [1] Y.-H. Chu, S. G. Rao, and H. Zhang, “A case for end system multicast,” in *ACM SIGMETRICS 2000*, (Santa Clara, CA), pp. 1–12, ACM, June 2000.
- [2] Y. Guo, K. Suh, J. Kurose, and D. Towsley, “P2Cast: peer-to-peer patching scheme for VoD service,” in *Proc. of the 12th Int. Conf. on World Wide Web*, pp. 301–309, 2003.
- [3] T. Do, K. A. Hua, and M. Tantaoui, “P2VoD: Providing fault tolerant video-on-demand streaming in peer-to-peer environment,” in *to appear in the Proc. of the IEEE Int. Conf. on Communications (ICC 2004)*, jun 2004.
- [4] D. A. Tran, K. Hua, and T. Do, “A peer-to-peer architecture for media streaming,” *IEEE Journal on Selected Areas in Communications, Special Issue on Advances in Service Overlay Networks*, vol. 22, pp. 121–133, jan 2004.
- [5] A. Reibman, Y. Wang, X. Qiu, Z. Jiang, and K. Chawla, “Transmission of multiple description and layered video over an EGPRS wireless network,” in *Proc. IEEE Int. Conf. on Image Processing*, vol. 2, pp. 136–139, October 2000.
- [6] M. Castro, P. Druschel, A.-M. Kermarrec, A. Nandi, A. Rowstron, and A. Singh, “Splitstream: high-bandwidth multicast in cooperative environments,” in *Proc. of the 19th ACM symposium on Operating systems principles*, pp. 298–313, 2003.
- [7] V. A. Vaishampayan, “Design of multiple description scalar quantizers,” *IEEE Trans. on Information Theory*, vol. 39, pp. 821–834, May 1993.
- [8] J. Apostolopoulos, T. Wong, W. tian Tan, and S. Wee, “On multiple description streaming with content delivery networks,” in *Proc. IEEE INFOCOM*, July 2002.
- [9] V. N. Padmanabhan, H. J. Wang, and P. A. Chou, “Resilient peer-to-peer streaming,” in *Proceedings of the 11th IEEE International Conference on Network Protocols*, p. 16, IEEE Computer Society, 2003.
- [10] A. Mohr, E. Riskin, and R. Ladner, “Unequal loss protection: Graceful degradation of image quality over packet erasure channels through forward error correction,” *JSAC special issue on Error-Resilient Image and Video Transmission*, 1999.
- [11] “<http://www.bbc.co.uk/rd/projects/dirac/>.”