# Optimal Bandwidth Assignment for Multiple Description Coding in Media Streaming

Pengye Xia*, Xing Jin† and S.-H. Gary Chan†
* Department of Physics
† Department of Computer Science and Engineering
The Hong Kong University of Science and Technology
Clear Water Bay, Kowloon, Hong Kong
Email: ph_xpx@stu.ust.hk, {csvenus, gchan}@cse.ust.hk

***Abstract*— Multiple description coding (MDC) has been widely used in media streaming to address the bandwidth heterogeneity issue. That is, the video source encodes data into multiple descriptions. At the receiver end, the streaming quality is proportional to the number of descriptions received. In this paper, we investigate how to assign bandwidth for descriptions in order to make full use of available network bandwidth. We formulate the problem as an optimization problem. We then decompose the problem and address it by an iterative algorithm. We evaluate the algorithm through simulations. Our results show that arbitrary description rates may severely degrade system performance and an optimal solution can make efficient use of network bandwidth.**

## I. Introduction

With the popularity of broadband Internet access, there has been an increasing interest in media streaming services. Websites like YouTube and MSN Video have offered numerous video clips for on-demand watch [1]. Online live streaming services through peer-to-peer (P2P) technology have also been widely deployed (e.g., PPLive, GridMedia, Sopcast and TVants) [2], [3].

In media streaming, the Internet's intrinsic heterogeneity continues a challenging problem. End users may have different edge bandwidth for data receiving or forwarding, especially in large-scale streaming with hundreds of thousands of users. One traditional solution is to offer multi-rate video at the source side and to allow users to receive video data at different rates according to their respective bandwidth [4]. MDC is one example of multi-rate video coding method [5]. In MDC, data are encoded into several descriptions, which are independent of each other. When all the descriptions are received, the original data can be reconstructed without distortion. If only a subset of the descriptions are received, the quality of the reconstruction degrades gracefully. Therefore, in MDC, an end user can choose to receive the maximum number of descriptions under its edge bandwidth constraint.

Many MDC-based streaming schemes focus on the delivery of descriptions. For example, in COSMOS, the video source determines the coding rates of descriptions, and end users selectively subscribe descriptions according to their edge bandwidth [6]. MDC is often jointly used with multiple path delivery [7]–[9]. For example, in overlay delivery, the server encodes its media content into $M$ descriptions using MDC (where $M$ is a tunable parameter), and transmits the descriptions along $M$ different overlay trees. Usually, a host has different descendants in different trees, and the descendant of a host in one tree is not the host's descendant in other trees. Therefore, packet loss at a host or failure of the host only causes the loss of a single description (out of $M$ descriptions) at each of its descendants.

However, we note that little work has been done on the optimal bandwidth assignment or coding rates of descriptions. Description coding rates have straightforward impact to the delivery performance. If a description has a high coding rate, some network paths may not have enough bandwidth to support its delivery. The loss rate of the description will be high. On the other hand, if descriptions have low coding rates, the number of descriptions and accordingly the coding cost will be high. Therefore, we need to properly set description number and description coding rates to achieve the best system performance.

In this paper, we propose an adaptive approach to adjust description coding rates according to the user bandwidth distribution. Our target is to provide the best streaming quality under certain network bandwidth constraint. We formulate the problem and address it by an adaptive solution. We also evaluate our algorithm through simulations. Our results show that arbitrary description rates may severely degrade system performance and an optimal solution can make significant improvement on the use of network bandwidth.

The rest of the paper is organized as follows. In Section II, we describe our problem formulation. In Section III, we study how to address the problem. In Section IV, we present illustrative simulation results. We finally conclude in Section V.

## II. Problem Formulation

We consider a single video source in the network, which provides video content to end users. The source encodes streaming data into multiple descriptions. The number of descriptions and the coding rate of each description are computed by the source according to the network condition. In our formulation, we set some constraint on the user receiving rate but not user sending rate. Hence, an end user may fetch data

978-1-4244-2309-5/09/$25.00 ©2009 IEEE

from the source or from other users. Our model is applicable to both client/server networks and P2P networks.

Suppose the source encodes streaming data into $m$ descriptions. Let a $m$-dimensional vector $\vec{d}$ denote the coding rates of descriptions. Suppose there are $n$ users in the system. Each user $i$ has a maximum receiving rate $c_i$. Given $\vec{d}$, a user selects a combination of descriptions to maximize the summation of description rates under the constraint $c_i$ (i.e., the summation of description rates should not exceed $c_i$). We use a $n \times m$ matrix $K$ to denote the selection of descriptions by users. Each row of $K$ indicates the selection of descriptions for a certain user. An element of $K$ is either 0 or 1. $K_{ij}$ with value 1 means the user $i$ will select to receive description $j$.

Given the above definitions, we compute the residual bandwidth at user $i$ as

$$v_i = c_i - \sum_{j=1}^{m} K_{ij} d_j.$$

Clearly, $v_i$ cannot be less than 0. Our target is to minimize the overall residual bandwidth over all end users. Formally, we try to solve

$$U^{\star} = \min_{\alpha} U_{\alpha}(\vec{d}),$$

where

$$U_{\alpha}(\vec{d}) = \sum_{i=1}^{n} w_i f(v_i),$$

subject to

$$\sum_{j=1}^{m} d_j = \alpha.$$

Here $w_i$ is a weighting factor which indicates the importance of user $i$, and $f$ is a convex function. In the formulation, we consider that the summation of all description rates equals to a constant $\alpha$. Function $U_{\alpha}$ is the summation function of residual bandwidth over all users, with a given $\alpha$. $U^{\star}$ is the minimum among all the $U_{\alpha}$ values.

In the above formulation, $m, n, c, w$ and $f$ are given. We aim to find the values of $K$ and $d$ to approach $U^{\star}$. In order to achieve a solvable version of the formulation, it is helpful to make some assumption on $f$. In reality, the residual bandwidth at a user is often much smaller than its bandwidth upper bound. The convex function $f$ can hence be approximated as a quadratic function, which can be obtained from the Taylor's expansion. For example, $e^x - 1 = x + x^2 + O(x^3)$. Therefore, we require $f_i$ to satisfy:

$$f_i(x) = (x + b_i)^2 + constant.$$

Here the coefficient before $x$ has been combined into the weighting factor $w_i$. The constant clearly does not affect the solution. We hence ignore it in the following.

## III. Problem Solution

The above problem can be decomposed as follows. Firstly, we set a constant streaming rate $\alpha$. We address the optimization problem $U_{\alpha}$ to identify vector $\vec{d}$. Secondly, with a series of $\alpha$ values, we select the minimum $U_{\alpha}$ value, which gives the result of $U^{\star}$. Therefore, a key issue is how to address the optimization problem $U_{\alpha}$ for a constant $\alpha$. That is:
Minimize

$$U_{\alpha}(\vec{d}) = \sum_{i=1}^{n} w_i f_i[(c_i - \sum_{j=1}^{m} K_{ij} d_j)],$$

where

$$f_i(x) = (x + b_i)^2,$$

subject to

$$\sum_{j=1}^{m} d_j = \alpha.$$

To address the problem, we divide it into two layers. We use an algorithm to iteratively address the two layers of problems. The output for one layer will be the input for another layer in the next round of calculation.

The sub layer problem is how to subscribe descriptions by a user to minimize its residual bandwidth. We assume the subscription of one user does not affect the others, and hence all the $n$ users' subscriptions are independent of each other. The solution to this sub layer problem will give the matrix $K$ and an $n$-dimensional vector $\vec{v}$ which indicates the residual bandwidth at users.

The master layer problem is how to update the description rates based on $K$. Denote the change of $\vec{d}$ as $\vec{x}$. Since the summation of all the elements of $\vec{d}$ is a constant, the summation of all the elements of $\vec{x}$ must be 0. Therefore, $\vec{x}$ has only $(m-1)$ free elements. Then $\vec{x}$ should satisfy the following conditions:

$$\begin{cases} K\vec{x} = \vec{v}, \\ \sum_{i=1}^{m} x_i = 0. \end{cases}$$

As $\vec{x}$ has only $(m-1)$ free dimensions, the above condition can be rewritten as:

$$A\vec{t} = \vec{v},$$

where

$$A_{ij} = K_{ij} - K_{im}, \forall i \in [1, n], j \in [1, m-1].$$

Here $\vec{t}$ is a $(m-1)$ dimensional vector which contains the first $(m-1)$ elements in $\vec{x}$ and $A$ is an $n \times (m-1)$ matrix.

The above linear system will be overdetermined if the number of users is larger than the number of descriptions (i.e., $n > m$), which is often the case in reality. For an overdetermined linear system, there is either no solution or one unique solution. It's impossible to get a solution if all the equations are independent or some of them are contradictive. An approximate solution can be obtained by computing the least square solution for the system, i.e,. minimizing $\|A\vec{t} - \vec{v}\|$.

Note that

$$\begin{aligned}\|A\vec{t}-\vec{v}\|^2 &= \sum_{i=1}^{n}([A\vec{t}]_i-\vec{v}_i)^2\\ &= (At-v)^T(At-v).\end{aligned}$$

If we require

$$\begin{aligned}\frac{d}{dt}(At-v)^T(At-v) &= 2A^TAt-2A^Tv\\ &= 0,\end{aligned}$$

we have

$$A^TAt=A^Tv,$$

or, equivalently,

$$t=(A^TA)^{-1}A^Tv.$$

Therefore, $t$ can be derived as above. We now prove that an approximate least square solution minimizes $U_\alpha$ for a given matrix $K$. We first consider a simple case, and then generalize the result. The proof is as follows.

Consider the case $w_i=1, b_i=0, \forall 1\le i\le n$. Then

$$U_\alpha(\vec{d})=\sum_{i=1}^{n}[c_i-\sum_{j=1}^{m}K_{ij}(d_j+x_j)]^2.$$

Since

$$v_i=c_i-\sum_{j=1}^{m}K_{ij}d_j,$$

and

$$\sum_{j=1}^{m}x_j=0,$$

we have

$$\begin{aligned}U_\alpha(\vec{d}) &= \sum_{i=1}^{n}(v_i-\sum_{j=1}^{m}K_{ij}x_j)^2\\ &= \sum_{i=1}^{n}(v_i-\sum_{j=1}^{m-1}K_{ij}x_j+K_{im}\sum_{j=1}^{m-1}x_j)^2\\ &= \sum_{i=1}^{n}(v_i-\sum_{j=1}^{m-1}A_{ij}t_j)^2\\ &= \sum_{i=1}^{n}([A\vec{t}]_i-\vec{v}_i)^2\\ &= \|A\vec{t}-\vec{v}\|^2.\end{aligned}$$

We see that the least square solution indeed minimizes the $U_\alpha$ function. Now we extend the proof to the case with weighting factors and general quadratic functions.

$$\begin{aligned}U_\alpha(\vec{d}) &= \sum_{i=1}^{n}(\sqrt{w_i}v_i+\sqrt{w_i}b_i-\sqrt{w_i}\sum_{j=1}^{m-1}A_{ij}t_j)^2\\ &= \sum_{i=1}^{n}([I_wAt]_i-[I_w(v+b)]_i)^2\\ &= \|I_wAt-I_w(v+b)\|^2,\end{aligned}$$

where $\forall x,y\in[1,n]$,

$$I_w(x,y)=\begin{cases}\sqrt{w_x}, & \text{if } x=y,\\ 0 & \text{if } x\neq y.\end{cases}$$

Here $I_w$ is an $n\times n$ diagonal matrix representing the weighting factors. Using the similar method as above, we have

$$\|I_wAt-I_w(v+b)\|^2=(I_wAt-I_w(v+b))^T(I_wAt-I_w(v+b)).$$

If we require

$$\begin{aligned}&\frac{d}{dt}(I_wAt-I_w(v+b))^T(I_wAt-I_w(v+b))\\ =\ & 2A^TI_w^TI_wAt-2A^TI_w^TI_w(v+b)\\ =\ & 0,\end{aligned}$$

we have

$$t=(A^TI_w^TI_wA)^{-1}A^TI_w^TI_w(v+b).$$

This is the solution for the master layer problem with given $K$ and weighting factors. Variable $t$ indicates the amount of change that needs to be applied to the description rate vector. Given $t$, an updated $\vec{d}$ can be calculated, which is then applied to the sub layer problem to calculate $K$ for the next round.

## IV. Illustrative Numerical Results

We have done simulations to evaluate our algorithm. Suppose the maximum receiving rate at users, $c$, are uniformly distributed within $[0,1]$. We consider that the source has enough bandwidth to support all users. We set the the same convex function $f$ for all users and set $b_i=0$.

In some cases, the solution series jump around the local minimum. To avoid infinite loop, the maximum number of rounds is set to 20, which is large enough for the algorithm to converge. The final solution is selected with the smallest $U_\alpha$ if the series do not converge to a certain point after 20 rounds. In the simulations, the system consists of 100 users with uniformly distributed bandwidth. The number of descriptions is variated from 4 to 9. The streaming rate $\alpha$ is set to 1.0.

Figure 1 shows $U_\alpha$ versus the number of descriptions $m$. $U_\alpha$ decreases as the number of descriptions increases. Clearly, the more descriptions, the smaller granularity for one description. Using more descriptions can hence better handle network heterogeneity.

From the figure, $U_\alpha$ is almost 0 when the stream is divided into 7 descriptions. Further increment in the description number does not significantly reduce the $U_\alpha$ value. On the other hand, the number of descriptions should be kept small in order to reduce the coding cost. Therefore, in this case, 7 is a good choice for the number of descriptions.

We also tune the streaming rate $\alpha$ to evaluate the algorithm performance. We set all the weighting factors to the same value, and set the number of descriptions to 7. Figure 2 shows $U_\alpha$ versus streaming rate. From the figure, $U_\alpha$ decreases as the streaming rate increases from 0.2 to 1.0, and increases as the streaming rate increases from 1.0 to 2.0. $U_\alpha$ hits the minimum value when the streaming rate reaches 1.0. Certainly,

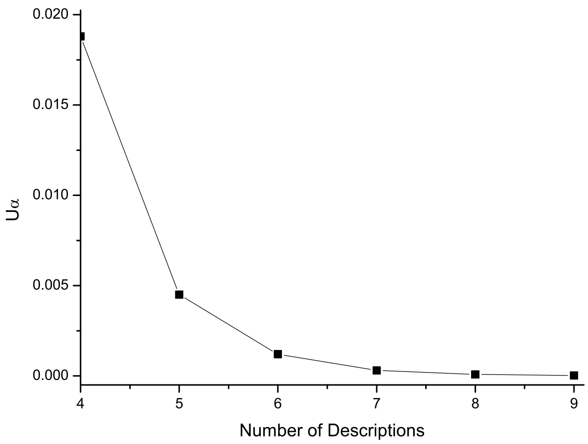



Fig. 1. $U_\alpha$ versus number of descriptions $m$.

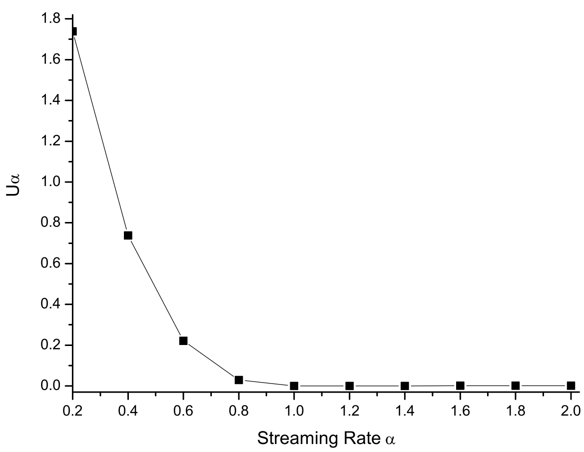


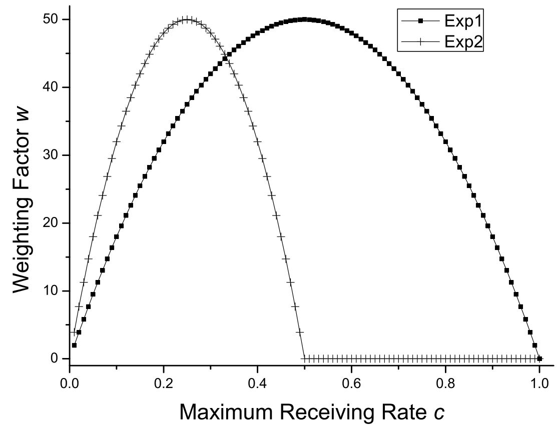

Fig. 2. $U_\alpha$ versus streaming rate $\alpha$.

the optimal streaming rate may vary away from 1.0 if the bandwidth or weighting factor distribution changes.

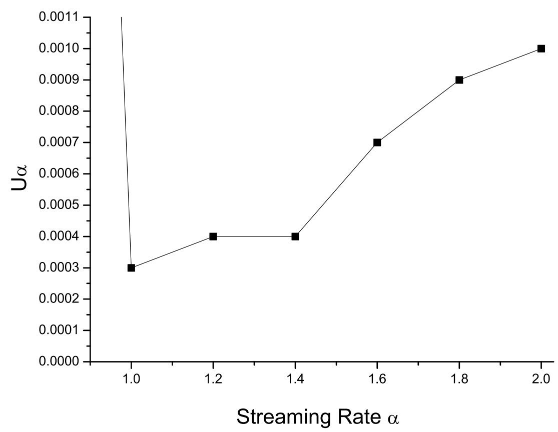
The trend of the curve can be explained as follows. With a very low streaming rate, it is hard to satisfy the requirement of users with high bandwidth capacity. The overall residual bandwidth is hence high. If the streaming rate is too high, the resultant description rate is also high and the users with low bandwidth capacity cannot be fulfilled. Therefore, it is better to use a proper streaming rate in between.

The corresponding optimal partition for each of the streaming rate $\alpha$ is listed in Table I. Then for these 100 users with uniformly distributed bandwidth and the same weight, the streaming rate should be set to 1.0 to achieve the minimum $U^{\star}$.

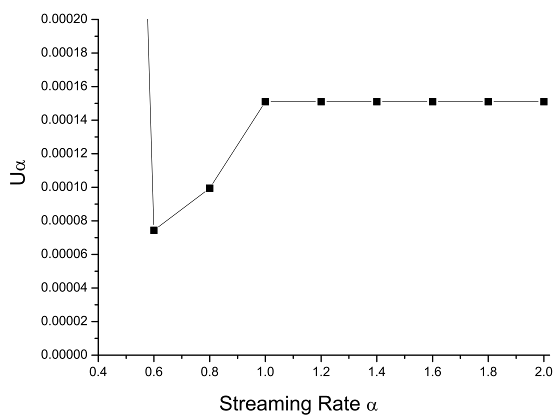
We also investigate the impact of weighting factor. We set the number of descriptions to 7. We increase the weighting factor for one user and keep the weighting factors for all other




(a)




(b)




(c)

Fig. 3. (a) Settings of weighting factors in two experiments. (b) Results of Exp1.(c) Results of Exp2.

TABLE I

OPTIMAL PARTITION FOR CORRESPONDING STREAMING RATE.

| α | optimal partition | | | | | | | $U_\alpha$ |
|---|---|---|---|---|---|---|---|---|
| 0.2 | 0.0031 | 0.0078 | 0.0122 | 0.0176 | 0.0276 | 0.0549 | 0.0768 | 1.7388 |
| 0.4 | 0.0095 | 0.0116 | 0.0187 | 0.0334 | 0.0600 | 0.1087 | 0.1581 | 0.7381 |
| 0.6 | 0.0071 | 0.0144 | 0.0275 | 0.0544 | 0.0931 | 0.1702 | 0.2333 | 0.2215 |
| 0.8 | 0.0084 | 0.0172 | 0.0368 | 0.0688 | 0.1278 | 0.2288 | 0.3121 | 0.0288 |
| 1.0 | 0.0136 | 0.0249 | 0.0503 | 0.0838 | 0.1560 | 0.2845 | 0.3869 | 0.0003 |
| 1.2 | 0.0169 | 0.0350 | 0.0628 | 0.1247 | 0.2109 | 0.2804 | 0.4692 | 0.0004 |
| 1.4 | 0.0179 | 0.0365 | 0.0592 | 0.1097 | 0.2236 | 0.4085 | 0.5446 | 0.0006 |
| 1.6 | 0.0181 | 0.0362 | 0.0652 | 0.1228 | 0.2085 | 0.4504 | 0.6988 | 0.0008 |
| 1.8 | 0.0206 | 0.0375 | 0.0665 | 0.1102 | 0.2408 | 0.4813 | 0.8431 | 0.0008 |
| 2.0 | 0.0183 | 0.0353 | 0.0653 | 0.1253 | 0.2519 | 0.5018 | 1.0020 | 0.0010 |

TABLE II

IMPACT OF WEIGHTING FACTOR.

| Weighting factor | Residual Bandwidth for the User |
|---|---|
| 1 | 0.0028 |
| 100 | 0.0001 |
| 10000 | 0.0000 |

users as 1. Table II shows the variation in the user's residual bandwidth. As the weighting factor of a user increases, the residual bandwidth at the user decreases. This is because the algorithm preferentially uses the bandwidth of important users.

In addition, we change the distribution of weighting factors to evaluate our algorithm. Figure 3(a) shows the settings of weighting factors in two independent experiments. Figures 3(b) and 3(c) show their resulting $U_\alpha$, respectively. The optimal rate is 1.0 in the first experiment, and is a lower value of 0.6 in the second experiment. This is because in the second experiment, we impose heavier weights on users with low bandwidth. Then with the same number of descriptions, a smaller streaming rate will lead to averagely smaller description rates, which can better fulfill the users with lower bandwidth capacity.

## V. CONCLUSION AND FUTURE WORK

In this paper, we study how to set description coding rates for MDC in video streaming. We formulate the problem as an optimization problem and propose an iterative algorithm to address it. Using our algorithm, an optimal streaming rate and a set of optimal description rates could be computed.

While the algorithm has been shown to be efficient through simulations, there are still many practical issues unaddressed. One challenge is how frequently the descriptions rates should be adjusted. If the network is highly dynamic, a highly frequent adjustment may better serve users. However, the cost for calculation would accordingly increase. We need to achieve proper tradeoff between the solution performance and the cost. Another challenge is to refine the problem formulation by considering very small descriptions. That is, some description rates from the optimal solution may be too low for practical MDC encoding. We should set a lower bound for the description coding rate, and prevent the algorithm from generating descriptions with lower rates than the bound.

At the current stage, we have not considered how users fetch descriptions. In a client/server network, users can simply fetch data from the server. But nowadays a P2P system has become a better choice for video streaming because of its high scalability. We plan to study the description delivery problem in a P2P system in the future. The major challenge is that description delivery in P2P environment is highly dynamic and flexible. The subscription of descriptions by a user depends on not only its edge bandwidth, but also peers that possess the desired descriptions (so-called overlay parents). Given different ways of identifying and connecting to overlay parents, the subscription method of descriptions could be completely different. We hence need to explore different overlay construction methods.

## REFERENCES


[1] YouTube. [Online]. Available: http://www.youtube.com/

[2] PPLive. [Online]. Available: http://www.pplive.com

[3] Y. Tang, J.-G. Luo, Q. Zhang, M. Zhang, and S.-Q. Yang, "Deploying P2P networks for large-scale live video-streaming service," *IEEE Commun. Mag.*, vol. 45, no. 6, pp. 100–106, June 2007.

[4] B. Li and J. Liu, "Multirate video multicast over the Internet: An overview," *IEEE Network*, vol. 17, no. 1, pp. 24–29, Jan. 2003.

[5] V. Goyal, "Multiple description coding: compression meets the network," *IEEE Signal Processing Magazine*, vol. 18, no. 5, pp. 74–93, Sept. 2001.

[6] M.-F. Leung and S.-H. Chan, "Broadcast-based peer-to-peer collaborative video streaming among mobiles," *IEEE Trans. Broadcasting*, vol. 53, no. 1, pp. 350–361, March 2007.

[7] V. Padmanabhan, H. Wang, P. Chou, and K. Sripanidkulchai, "Distributing streaming media content using cooperative networking," in *Proc. ACM NOSSDAV'02*, May 2002, pp. 177–186.

[8] M. Castro, P. Druschel, A.-M. Kermarrec, A. Nandi, A. Rowstron, and A. Singh, "SplitStream: High-bandwidth multicast in cooperative environments," in *Proc. ACM SOSP'03*, Oct. 2003, pp. 298–313.

[9] R. Tian, Q. Zhang, Z. Xiang, Y. Xiong, X. Li, and W. Zhu, "Robust and efficient path diversity in application-layer multicast for video streaming," *IEEE CSVT*, vol. 15, no. 8, pp. 961–972, Aug. 2005.