Analysis of Receiver Adaptation for Layered Video Transmission over Heterogeneous Networks: A Microscopic Perspective

Po-Chin Hu†, Zhi-Li Zhang‡, Mostafa Kaveh†

† Dept. of Electrical & Computer Engineering
University of Minnesota
Minneapolis, MN55455
{phu,kaveh}@ece.umn.edu

‡ Dept. of Computer Science & Engineering
University of Minnesota
Minneapolis, MN55455
zhzhang@cs.umn.edu

Abstract— **Receiver adaptation layered multicast video has been proposed to address the issue of network bandwidth heterogeneity in multi-party video communication over networks. An important issue in this approach is how to conduct layer adaptation to achieve the "optimal level" of subscription to maximize the perceived video quality at receivers while utilize bandwidth efficiently. In the current work, optimal layers are typically determined by congestion signals fed back from networks. This approach implies that instantaneous congestion is a key parameter; however, in random traffic, such an implication somewhat under-estimates other parameters.**

In order to investigate the performance of "layer adaptation" in random traffic for improving the network-based multi-party video communication system, in this paper we present an analytical studying of the issue of receiver layer adaptation for layered video transmission. In this work, we use the notion of congestion probability as a partial and indirect measure of the perceived layered video quality at a receiver. The system's efficiency under a given receiver layer adaptation scheme is also defined and used as a metric to study the performance of a receiver layer adaptation scheme. Through the analysis, we determine the optimal receiver layer adaptation scheme, which maximizes the system efficiency while providing best sustainable video quality to a receiver. We show that the greedy receiver layer adaptation scheme is indeed optimal if the traffic generated by various layers of a layered video as well as the cross traffic are constant. However, in a more realistic setting where layered video traffic or the cross traffic varies over the time, the greedy receiver layer adaptation scheme is in general not optimal. To verify our theory, we conduct simulations using both stationary and non-stationary traffic. Our simulation results demonstrate the superior performance of the optimal receiver layer adaptation scheme.

I. Introduction

Many emerging networked multimedia applications involve multi-party video transmission. A challenging problem in one-to-many or many-to-many video communication over a shared network such as the Internet arises from its heterogeneous underlying networking environments and users' diverse requirement of video quality. For example, a video seminar broadcasted over an internetwork may be shared by users who are connected either via a high-speed network using ISDN video-phones or via a low-speed modem using PSTN video-phones.

In order to address the problem posed by network bandwidth heterogeneity, recently a receiver-driven bandwidth adaptation approach based layered video multicast has been proposed (see, e.g.,[12], [3], [9], [4], [8], [11], [13], [15]). Under this approach, a video signal is encoded into a number of layers that can be incrementally combined to provide progressive refinement [14]. Each layer of layered video is transmitted to receivers using a separate IP multicast group [5]. *Receiver-driven bandwidth adaptation* is accomplished by allowing individual receivers to join or leave one or multiple layers, based on their perceived network condition and/or the desired video quality.

In the design of receiver-driven layered video multicast schemes, two important aspects must be carefully considered: (1) *receiver layer adaptation* and (2) *receiver layer join/leave coordination*. Receiver layer adaptation refers to mechanisms and policies used by receivers to conduct bandwidth inference and to determine the number of layers to subscribe at a given time. Receiver join/leave coordination refers to mechanisms and policies used in a layered video multicast scheme to coordinate receivers' join and leave operations to minimize mutual interference and achieve the system's stability in receiver adaptation as well as to possibly obtain fair sharing of network bandwidth. In some sense, receiver layer adaptation is concerned with the microscopic behavior of individual receivers' interaction with the network in a layered video multicast scheme, whereas receiver layer join/leave coordination is concerned with the macroscopic behavior of a layered video multicast scheme in attaining the system's stability. Clearly, receiver layer join/leave coordination is critical in facilitating the proper functioning of receiver layer adaptation, thereby ensuring a stable overall system performance. On the other hand, the efficiency of the overall system performance in *steady state* will hinge largely on the receiver layer adaptation mechanisms used by individual receivers.

Most existing studies in receiver adaptation layered multicast have focused primarily on the aspect of receiver layer join/leave coordination, where innovative mechanisms such as shared-learning join experiment with exponential back-off [9] and synchronization points and deaf periods [15] are proposed to coordinate receiver join/leave operations. In contrast, receiver layer adaptation has not garnered sufficient attention. The common approach used in receiver layer adaptation (see e.g., [9], [15]) is based on variations of a simple scheme, which we refer to as the *greedy receiver layer adaptation* scheme. Under this scheme, packet loss is used as an indication of network congestion. Whenever packet loss is experienced during "normal" transmission of layered video, a layer is subsequently dropped. Whereas whenever a receiver discovers there is spare capacity (e.g., through the join experiment [9] or server-initiated probes at synchronization points [15]), an additional layer is joined. With the help of receiver layer join/leave coordination, it is *expected* that the "optimal operating point" (i.e., the optimal level of layer subscription) can be achieved in steady state and the perceived video quality at receivers can therefore be maximized. This expectation seems reasonable if the traffic generated by each layer of a layered video is relative constant, and there is no or little cross traffic sharing the bottleneck link except for the traffic generated by the various layers of the layered video (as is the case in the simulation study conducted in [9], [15]). Under this scenario, the "optimal operating point" of the bottleneck link is also well-defined, namely, the maximum number of layers the bottleneck link can sustain. However, in a more realistic setting, the traffic generated by a layered video is likely to be variable (despite that it may be generated by a constant-bit-rate codec). Furthermore, besides the layered video traffic, the network links are also shared by other traffic, especially those generated by non-adaptive continuous media applications. Hence the traffic of the links fluctuates over the time, and *random* packet loss due to "transient congestion" is likely to occur. Under this setting, several questions arise:

- What should be the appropriate definition of the "optimal operating point" for receiver layer adaptation?
- Given an appropriate definition of "optimality", is the greedy receiver layer adaptation scheme optimal?
- If not, what is the optimal receiver layer adaptation scheme?

This paper is devoted to the investigation of the above questions, in particular, the assessment of the greedy adaptation. We present an analytical model, where we limit our analysis to a single network bottleneck link shared by a receiver and some cross traffic. Both layered video traffic and cross traffic are modeled by some given but arbitrary random processes. Under the assumption that the traffic is stationary, we introduce the notion of *congestion probability* of the bottleneck link. This notion provides a partial and indirect measure of the perceived layered video quality at the receiver. Given a receiver layer adaptation scheme, we define the *system efficiency* under the given scheme to be the ratio of the amount of the total "uncongested" traffic transmitted across the bottleneck link to the total amount of traffic that can be sustained by the bottleneck link. This metric is used to study the performance of a receiver layer adaptation scheme. Hence the "optimal operating point" of receiver layer adaptation is the *maximum* system efficiency that can be attained. An *optimal* receiver layer adaptation scheme is therefore a scheme that achieves this maximum system efficiency.

Based on the analytical model we introduce, we first analyze the greedy receiver layer adaptation scheme and derive an expression for its congestion probability. We show that the greedy receiver layer adaptation scheme is indeed optimal under the constant traffic setting. However, under the variable traffic setting it is in general not optimal. We also investigate the problem of optimal receiver adaptation. Under the assumption of stationary traffic, we find that the optimal level of layer subscription does not depend on random packet loss or transient congestion encountered, but rather on the stationary statistics of the traffic. Based on this observation, we identify the optimal statistical receiver adaptation scheme which maximizes the expected system efficiency. To verify our theory, we conduct simulations using both synthesized traffic and traces from real video and network traffic.

The remainder of this paper is organized as follows. In In Section II, we present the analytical model. In Section III, we analyze the greedy receiver adaptation scheme. In Section IV we present the optimal receiver adaptation scheme. In Section V, we present the simulation results. The paper is concluded in Section VI.

II. The Analytical Model

In this section, we present an analytical model and introduce two performance metrics, *congestion probability* and *efficiency*, for the study of receiver layer adaptation for layered video transmission.

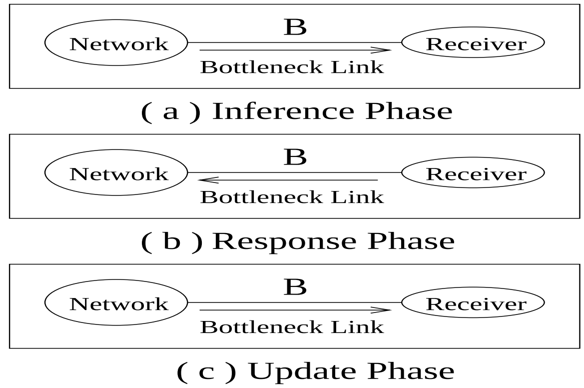


Fig. 1. Overview of the problem setting

For simplicity of analysis, we confine our attention to a single *bottleneck* link with a total bandwidth of B (see Figure 1(a)) and assume that some types of receiver layer join/leave coordination and other control mechanisms are employed that minimize or eliminate such effects as receiver join interference and leave delay to ensure the eventual stability of the system. Therefore, we will consider the *steady-state* behavior of an individual receiver who shares this bottleneck link with other cross traffic, and focus on the criteria used in a receiver layer adaptation scheme in deciding when to join or leave a layer.

In our model, we assume that both the traffic generated by a layered video and the cross traffic are modeled by some given but arbitrary (discrete-time) random processes. The random process describing the cross traffic is denoted by $\{\mathbf{B_c}(n), \in N\}$[1], where the random variable denotes the amount of the cross traffic at time n is $\mathbf{B_c}(n)$. When we do not wish to emphasize the time, we will drop the time index n. We assume that the layered video is encoded into L layers, where layer 1 is the base layer, and for $l = 2, \ldots, L$, layer l is an enhancement on top of layers $1, \ldots, l-1$. For convenience of notation, we also include a "dummy" layer, layer 0, which contains no video information. As in [9], [2], to simplify analysis we assume that the proportion of the traffic generated by the various layers are fixed. Namely, at any time n, the amount of the traffic generated by layer i is $\lambda_i B_v(n)$, where $\lambda_0 = 0$ and $\lambda_i > 0$, $i = 1, \ldots, L$ and $\sum_{i=0}^{L} \lambda_i = 1$. For $i = 0, 1, \ldots, L$, $\Lambda_i = \sum_{j=0}^{i} \lambda_j$ is the cumulative fraction of the combined video traffic of layers $0, 1, \ldots, i$. Receivers adapt the rate of layered video transmission they receive to the available capacity at the bottleneck link by dynamically adding or dropping layers. Note that this setting does not imply that the cross traffic B_c is "irresponsive" although in the analysis we consider cross traffic as a priority traffic.

Figure 1 illustrates the three phases in receiver layer adaptation for layered video transmission. Figure 1(a) shows the first phase, the *inference phase* where a subscribed number of layers of the layered video are multicasted to the receiver across the bottleneck link which is shared by other cross traffic. The

[1] In this paper we adopt the convention that the bold-faced letter denotes either a random variable or a random process, whereas a corresponding regular-faced letter denotes an instant of the random variable or a realization (i.e., sample path) of the random process.

receiver detects whether the bottleneck link is congested or not and infers the number of layers the bottleneck link can sustain at the moment. In Figure 1(b), the receiver is in the *response phase*, requesting to either join an additional number of layers or leaving a certain number of existing layers. In the last phase, the *update phase*, the newly subscribed number of layers of the layered video are transmitted to the receiver. The delay between the time the receiver's join or leave request and the time the newly subscribed layers are transmitted across the bottleneck link to the receiver is referred to as *receiver layer adaptation delay*. The receiver layer adaptation delay at time n is denoted by n_d.[2]

For a receiver layer adaptation scheme, let $\Lambda(n)$ denote the sum of the fractions of the layered video traffic subscribed at time n according to the receiver layer adaptation scheme. In other words, $\Lambda(n) = \sum_{j=0}^{i} \lambda_j$ if the first i layers are subscribed at time n according to the receiver layer adaptation. Given the adaptation delay n_d, then at time $n+n_d$, the total amount traffic $B_T(n+n_d)$ transmitted across the bottleneck link is such that $B_T(n+n_d) = \Lambda(n)B_v(n+n_d) + B_c(n+n_d)$. Hence at time $n+n_d$ whether the new subscription will cause congestion at the bottleneck is determined by whether $B_T(n+n_d) \le B$. This yields the following notion of congestion probability.

The congestion probability $P_{congestion}$ of the bottleneck link at time n under a receiver layer adaptation scheme is defined as

$$\begin{aligned} P_{congestion} &= P\{\mathbf{B_T}(n) > B\} \qquad (1) \\ &= P\{\mathbf{\Lambda}(n-n_d)\mathbf{B_v}(n) + \mathbf{B_c}(n) > B\} \end{aligned}$$

where $\mathbf{B_T}(n)$, $\mathbf{\Lambda}(n)$, $\mathbf{B_v}(n)$ and $\mathbf{B_c}(n)$ denote the corresponding random variables of $B_T(n)$, $\Lambda(n)$, $B_v(n)$, and $B_c(n)$, respectively.

To provide a somewhat more comprehensive measure of the overall system performance, we introduce the notion of *efficiency* of the whole system as a metric to study the performance of receiver layer adaptation schemes. This metric is formally defined using sample paths of the involved traffic processes. For a given receiver layer adaptation scheme, let $B_T(n) = B_c(n)+\Lambda(n-n_d)B_v(n)$ denote the total amount of traffic transmitted across the bottleneck link at time n. Then the (asymptotic) efficiency of the whole system (here the bottleneck link) under the given receiver layer adaptation scheme, in short, the *system efficiency*, is defined as

$$\mathit{Eff} = \lim_{N\to\infty} \frac{\sum_{n=1}^{N} B_T(n)\mathbf{1}\{B_T(n) \le B\}}{B \cdot N} \qquad (2)$$

where $\mathbf{1}\{B_T(n) \le B\}$ is an indicator function, i.e., $\mathbf{1}\{B_T(n) \le B\} = 1$ if $B_T(n) \le B$, and 0 otherwise.

In the above definition we have assumed that any video information transmitted at the time of congestion (i.e., $B_T(n) > B$) becomes "noisy" and useless, and thus not counted as part of the "goodput". This is a reasonable assumption (albeit somewhat conservative) if "uniform dropping" [2] is employed in network routers. In this case, any packet has the potential to be dropped at the time of congestion. Intuitively, *Eff* is the ratio of the amount of the "uncongested" video traffic transmitted using a given receiver layer adaptation scheme to the total amount of traffic that can be sustained by the bottleneck link.

Assuming that the involved traffic processes are stationary and *ergodic* [10], the *system efficiency* defined in (2) provides an approximate measure of the expected "goodput" of the bottleneck link under a given receiver layer adaptation scheme. Clearly $0 \le \mathit{Eff} \le 1$. The following inequality[3] relates the system efficiency and the congestion probability of the bottleneck link

$$\mathit{Eff} \le 1 - P_{congestion}. \qquad (3)$$

From this relation, we see that an aggressive receiver layer adaptation scheme is unlikely to have a high efficiency if it results in a high congestion probability. In this paper, we say that a receiver layer adaptation scheme is *optimal* if it achieves the maximum system efficiency.

The system efficiency is defined from the perspective of the whole system performance. We can also introduce a notion of efficiency from the perspective of a receiver. This notion of efficiency, referred to as the *receiver efficiency* and denoted by $\mathit{Eff}_{receiver}$, can be defined as follows. For a given receiver adaptation scheme,

$$\mathit{Eff}_{receiver} = \lim_{N\to\infty} \frac{\sum_{n=1}^{N} \Lambda(n-n_d)B_v(n)\mathbf{1}\{B_T(n) \le B\}}{B \cdot N}. \qquad (4)$$

Intuitively, $\mathit{Eff}_{receiver}$ measures the "goodput" of the "uncongested" video traffic received by the receiver. We see that the receiver efficiency is always upper bounded by the system efficiency (i.e., $\mathit{Eff}_{receiver} \le \mathit{Eff}$). In the rest of the paper, we use the system efficiency *Eff* as the primary metric (together with the congestion probability $P_{congestion}$) to study the performance of receiver layer adaptation schemes. For ease of reference, we have listed all the notation used in this paper in Table I.

III. Analysis of the Greedy Receiver Layer Adaptation Scheme

In this section, following the analytical model and notation introduced in Section II, we model the steady-state behavior of the greedy receiver layer adaptation and compute its (steady-state) congestion probability. We assume that the traffic processes involved ($\mathbf{B_v}$ and $\mathbf{B_c}$) are *stationary* and *ergodic*. Hence we can analyze a time domain realization (i.e., sample path) of the traffic processes for the ensemble average.

Under the greedy receiver layer adaptation scheme, a receiver infers the spare capacity of the bottleneck link at time n and attempts to subscribe to the maximum number of layers that can

[2] Depending on the context, n_d denotes either the receiver layer adaptation delay for a receiver layer subscription request sent at time n, or the receiver layer adaptation delay for a receiver subscription request sent at a time earlier than n and the updated subscription taking effect at time n.

[3] This inequality can be proved as follows. Since $B_T(n)\mathbf{1}\{B_T(n) \le B\} \le B\mathbf{1}\{B_T(n) \le B\}$, from (2), we have $\mathit{Eff} \le \lim_{N\to\infty} \frac{\sum_{n=1}^{N} \mathbf{1}\{B_T(n)\le B\}}{N}$. Under the assumption of stationarity and ergodicity, we have $\mathit{Eff} \le P\{\mathbf{B_T} \le B\} = 1 - P_{congestion}$.

Notation	*Description*
B	the bandwidth of a bottleneck link
$B_c(n)$	the amount of cross traffic in the bottleneck link at time n
$B_v(n)$	the total traffic generated by a layered video at time n
$B_T(n)$	total traffic of the bottleneck link at time n
$l(n)$	the maximal number of layers sustainable
L	total number of layers of the encoded video
n_d	receiver layer adaptation delay
λ_i	the fraction of layer i traffic to the encoded video
Λ^{max}	cumulative fraction of the layers inferred by the receiver
Λ_i	cumulative fraction of first i layers, i.e. $\Lambda_i = \sum_{j=0}^{i} \lambda_j$
$\Lambda(n)$	the cumulative fraction of layers subscribed to at time n

TABLE I
LIST OF NOTATION.

be sustained without congestion by the bottleneck link (see Figure 1). The maximum number of sustainable layers at time n , denoted by $l(n)$, can be expressed as

$$l(n) = \max \left\{ l \geq 0 : (\sum_{j=0}^{l} \lambda_j) B_v(n) \leq B - B_c(n) \right\}. \quad (5)$$

Therefore, under the greedy receiver layer adaptation scheme the fraction of the layered video traffic to be subscribed at time n, is

$$\Lambda^{max} = \sum_{i=0}^{l(n)} \lambda_i = \max \left\{ \Lambda_i : \Lambda_i B_v(n) \leq B - B_c(n) \right\}. \quad (6)$$

Taking the receiver layer adaptation delay n_d into account, the total amount of traffic transmitted across the bottleneck link at time n, after the receiver subscription request made at time $n - n_d$ has taken effect, is

$$\begin{aligned} B_T(n) &= (\sum_{i=0}^{l(n-n_d)} \lambda_i) B_v(n) + B_c(n) \\ &= \Lambda^{max}(n - n_d) B_v(n) + B_c(n). \end{aligned} \quad (7)$$

Hence the congestion probability of the bottleneck link at time n under the greedy receiver layer adaptation scheme is given by

$$P_{congestion} = P\{\mathbf{\Lambda^{max}}(n - n_d)\mathbf{B_v}(n) + \mathbf{B_c}(n) > B\}. \quad (8)$$

Because of the dependency among $\mathbf{\Lambda^{max}}(n - n_d)$, $\mathbf{B_v}(n)$, and $\mathbf{B_c}(n)$, evaluation of (8) is in general not an easy task. To simplify the task, we assume that $\mathbf{B_v}(n)$ and $\mathbf{B_c}(n)$ are independent. In addition, n_d is sufficiently large so that the correlation between $\mathbf{\Lambda^{max}}(n - n_d)$, $\mathbf{B_v}(n)$ $\mathbf{B_c}(n)$) can be ignored.[4] Under these independence assumptions, we can re-write (8) as a sum of conditional probability density functions. In the following derivation we will drop the time indices n and $n - n_d$, thanks to the assumption of stationarity and independence.

[4]In the simulation study of Section V, we will investigate the impact of n_d on the congestion probability and system efficiency of a receiver layer adaptation scheme.

Let $f_{\mathbf{B_T}}(x|\mathbf{\Lambda^{max}} = \Lambda)$ denote the conditional probability density function (pdf) of the total traffic $\mathbf{B_T}$ conditioning on $\mathbf{\Lambda^{max}} = \Lambda$. Then

$$\begin{aligned} P_{congestion} &= \sum_{i=0}^{L} P\{\mathbf{\Lambda^{max}}\mathbf{B_v} + \mathbf{B_c} \geq B | \mathbf{\Lambda^{max}} = \Lambda_i\} \times \\ &\qquad P\{\mathbf{\Lambda^{max}} = \Lambda_i\} \\ &= \sum_{i=0}^{L} [\int_B^\infty f_{\mathbf{B_T}}(x|\mathbf{\Lambda^{max}} = \Lambda_i)dx] P\{\mathbf{\Lambda^{max}} = \Lambda_i\} \end{aligned} \quad (9)$$

where under the greedy receiver layer adaptation scheme

$$P\{\mathbf{\Lambda^{max}} = \Lambda_i\} = P\{\Lambda_i \mathbf{B_v} \leq B - \mathbf{B_c} \leq \Lambda_{i+1}\mathbf{B_v}\}, \quad (10)$$

Since the pdf of the sum of two independent random variables is the convolution of their pdf's [10], from (7) we have

$$f_{\mathbf{B_T}}(x|\mathbf{\Lambda^{max}} = \Lambda_i) = \frac{1}{\Lambda_i} f_{\mathbf{B_v}}(\frac{x}{\Lambda_i}) * f_{\mathbf{B_c}}(x) \quad (11)$$

where $f_{\mathbf{B_v}}(x)$ and $f_{\mathbf{B_c}}(x)$ are the pdf's of $\mathbf{B_v}$ and $\mathbf{B_c}$, and $*$ denotes the convolution operator.

We now derive an expression for $P\{\mathbf{\Lambda^{max}} = \Lambda_i\}$ in terms of the pdf's of the random variables involved. For convenience, we introduce a new random variable $\mathbf{Z} = B - \mathbf{B_c}$. Let $f_{\mathbf{Z}}$ denote the pdf of this new random variable. From (10) and by conditioning on $\mathbf{B_v}$, we have

$$\begin{aligned} P\{\mathbf{\Lambda^{max}} = \Lambda_i\} &= \int_0^\infty (\int_{\Lambda_i y}^{\Lambda_{i+1} y} f_{\mathbf{Z}}(z)dz) f_{\mathbf{B_v}}(y)dy \\ &= \int_0^\infty \int_{\Lambda_i y}^{\Lambda_{i+1} y} f_{\mathbf{Z}}(z) f_{\mathbf{B_v}}(y)dzdy \end{aligned} \quad (12)$$

Substituting (11) (12) into (9), we have

$$\begin{aligned} P_{congestion} = \sum_{i=0}^{L} \left(\int_B^\infty \frac{1}{\Lambda_i} f_{\mathbf{B_v}}(\frac{x}{\Lambda_i}) * f_{\mathbf{B_c}}(x)dx \right) \times \quad (13) \\ \int_0^\infty \int_{\Lambda_i y}^{\Lambda_{i+1} y} f_{\mathbf{Z}}(z) f_{\mathbf{B_v}}(y)dzdy. \end{aligned}$$

In (13) we have expressed the congestion probability of the greedy receiver layer adaptation scheme as a function of the pdf's of the layered video traffic and the cross traffic. Given that the pdf's $f_{\mathbf{B_v}}(x)$ and $f_{\mathbf{B_c}}(x)$ are known, we can use (13) to compute the (steady-state) congestion probability.

IV. THE OPTIMAL RECEIVER LAYER ADAPTATION SCHEME

In this section we determine the optimal receiver layer adaptation scheme through analysis. In particular, we show that the greedy receiver layer adaptation scheme is not optimal in general. However, with constant layered video traffic and constant cross traffic, the greedy receiver layer adaptation scheme coincides with the optimal receiver layer adaptation scheme.

As before we assume that $\mathbf{B_v}$ and $\mathbf{B_c}$ are stationary and ergodic. For $i = 0, 1, \ldots, L$, define

$$M_i = E[(\Lambda_i \mathbf{B_v} + \mathbf{B_c})\mathbf{1}_{\{\Lambda_i \mathbf{B_v} + \mathbf{B_c} \leq B\}}] \quad (14)$$

where E denotes the expectation and $\mathbf{1}_{\mathcal{E}}$ is an indicator function: $\mathbf{1}_{\mathcal{E}} = 1$ if and only if the event $\mathcal{E}$ holds.

Intuitively, M_i is the average amount of "uncongested" traffic transmitted across the bottleneck *given that i layers of the layered video are subscribed*. By "uncongested" traffic we mean the total traffic transmitted across the bottleneck link that does not cause congestion at the time of transmission, i.e., $B_T \leq B$. We will refer to the vector $\mathbf{M} = [M_0, M_1, ..., M_L]^T$ as the (uncongested) *mean traffic vector*. Here the superscript T denotes the transpose of a vector or matrix.

Assuming that $\mathbf{B_v}$ and $\mathbf{B_c}$ are independent, we can express M_i in terms of the pdf's of $\mathbf{B_v}$ and $\mathbf{B_c}$ as follows:

$$M_i = \int_0^B \frac{x}{\Lambda_i} f_{\mathbf{B_v}}(\frac{x}{\Lambda_i}) * f_{\mathbf{B_c}}(x) dx \tag{15}$$

where $*$ denotes the convolution operator.

For $i = 0, 1, \ldots, L$, define $S_i = P\{\mathbf{\Lambda^{max}} = \Lambda_i\}$, where $P\{\mathbf{\Lambda^{max}} = \Lambda_i\}$ is given by (10). S_i denotes the probability that at given time the receiver infers that it can subscribe to i layers of the video traffic without causing congestion at the bottleneck link. We refer to the vector $\mathbf{S} = [S_0, S_1, ..., S_L]^T$ as the *layer inference vector*. Clearly $\sum_{i=0}^{L} S_i = 1$.

Note that from the definitions of M_i and S_i in (15) and (10), the (uncongested) mean traffic vector $\mathbf{M}$ and the layer inference vector $\mathbf{S}$ depend only on the statistics of the total layered video traffic and the cross traffic, *not* on any specific receiver layer adaptation scheme.

Now consider an arbitrary receiver layer adaptation scheme. Let $\mathbf{\Lambda}$ be a random variable denoting the number of layers subscribed under the said receiver layer adaptation scheme at any given time. For $i = 0, 1, \ldots, L$, let $D_i = P\{\mathbf{\Lambda} = \Lambda_i\}$. The vector $\mathbf{D} = [D_0, D_1, \ldots, D_L]$ represents the probability distribution function of $\mathbf{\Lambda}$ for the said receiver layer adaptation scheme. We express $\mathbf{D}$ in terms of $\mathbf{S}$ by introducing the following notation.

For $0 \leq i, j \leq L$, define $T_{ji} = P\{\mathbf{\Lambda} = \Lambda_j | \mathbf{\Lambda^{max}} = \Lambda_i\}$. Intuitively, T_{ji} represents the probability that the said receiver layer adaptation scheme decides to subscribe j layers when it infers that i layers can be sustained by the bottleneck link without congestion. We refer to the $(L+1) \times (L+1)$ matrix $\mathbf{T} = [T_{ji}]$ as the *transition matrix*. It is easy to see that $\mathbf{D} = \mathbf{TS}$. Note that for the greedy receiver layer adaptation scheme, $\mathbf{T}^{greedy}$ is the identity matrix, i.e., $T_{ji}^{greedy} = 1$ if and only if $j = i$. Hence $\mathbf{D}^{greedy} = \mathbf{T}^{greedy}\mathbf{S} = \mathbf{S}$.

We now proceed to derive a succinct matrix expression for *Eff*, which is defined in (2) on a sample path basis in Section II. Given the ergodicity of the traffic processes involved, we can re-write (2) as follows

$$\mathit{Eff} = \frac{\int_0^B x f_{\mathbf{B_T}}(x) dx}{B}. \tag{16}$$

Similarly to the derivation in (9), by conditioning on $\mathbf{\Lambda}$ and assuming that $\mathbf{\Lambda}$, $\mathbf{B_v}$ and $\mathbf{B_c}$ are independent, we have

$$\int_0^B x f_{\mathbf{B_T}}(x) dx = \sum_{i=0}^{L} \left(\int_0^B \frac{x}{\Lambda_i} f_{\mathbf{B_v}}(\frac{x}{\Lambda_i}) * f_{\mathbf{B_c}}(x) dx \right) \times P\{\mathbf{\Lambda} = \Lambda_i\} \tag{17}$$

where for $i = 0$, $\frac{x}{\Lambda_0} f_{\mathbf{B_v}}(\frac{x}{\Lambda_0})$ is assumed to be an impulse (Dirac delta) function.

Substituting (17) into (16) and using (15), we now can express the system efficiency under a given receiver layer adaptation scheme in a succinct matrix form as follows:

$$\mathit{Eff} = \frac{\mathbf{M}^T\mathbf{D}}{B} = \frac{\mathbf{M}^T(\mathbf{TS})}{B}. \tag{18}$$

Let k be such that $M_k = \max_{0 \leq i \leq L} M_i$. Define $\mathbf{D}^* = [D_0^*, D_1^*, \ldots, D_L^*]$ where $D_i^* = 1$ if $i = k$, and $D_i^* = 0$ if $i \neq k$. We claim that $\mathbf{D}^*$ has the maximum system efficiency. This is because for any arbitrary $\mathbf{D} = [D_0, D_1, \ldots, D_L]$ such that $D_i \geq 0$, $0 \leq i \leq L$, and $\sum_{i=0}^{L} D_i = 1$,

$$\sum_{i=0}^{L} M_i D_i \leq \sum_{i=0}^{L} M_k D_i = M_k = \sum_{i=0}^{L} M_i D_i^*.$$

Hence under the optimal receiver layer adaptation, the probability distribution vector for $\mathbf{\Lambda}^{optimal}$ is $\mathbf{D}^*$. Furthermore, the optimal receiver layer adaptation has the following transition matrix $\mathbf{T}^{optimal}$: for $0 \leq j, i \leq L$,

$$T_{ji}^{optimal} = \begin{cases} 1, \ if \ j = k \\ 0, \ if \ j \neq k. \end{cases} \tag{19}$$

This implies that the *optimal* level of subscription is the one which yields the *largest* mean of "uncongested" traffic at the bottleneck. Hence in steady state, the optimal receiver layer adaptation scheme is for a receiver to always join this optimal number of layers no matter what the number of layers inferred is at the current moment. In other words, the optimal receiver layer adaptation is determined by the (stationary) statistics of both the layered video traffic and the cross traffic, not by the *instantaneous* state of the network. This observation sheds light on how the optimal receiver layer adaptation scheme should be designed in practice.

Lastly, we note that since the optimal receiver layer adaptation scheme has a transition matrix $\mathbf{T}^{optimal}$ which has all 1's in row k, and 0 in the other rows, whereas the greedy receiver layer adaptation scheme has a transition matrix $\mathbf{T}^{optimal}$ which is an identity matrix. Hence, the greedy receiver layer adaptation scheme is in general *not* optimal. It is optimal if and only if $\mathbf{T}^{greedy}\mathbf{S} = \mathbf{T}^{optimal}\mathbf{S}$. This can only happen if $S_k = 1$ and $S_i = 0$ for $i \neq k$. We show that it is exactly the case when both the layered video traffic and the cross traffic are constant. In this case, $k = \max\{i : \Lambda_i B_v + B_c \leq B\}$ and $M_k = \Lambda_k B_v + B_c \geq M_i$ for $0 \leq i \leq L$. Hence $S_k = 1$ and $S_i = 0$ for $i \neq k$. As a result, $\mathbf{T}^{greedy}\mathbf{S} = \mathbf{T}^{optimal}\mathbf{S}$ and $\mathit{Eff}^{greedy} = \mathit{Eff}^{optimal}$.

V. Simulation Results

In this section we conduct simulations to verify the theoretical analysis we have developed in the previous sections. The simulations are performed using MATLAB, which provides a matrix-based simulation toolkit. We choose this numerical simulator because the focus of this paper is on the analytical behavior of receiver layer adaptation schemes, not on the protocol implementation or control mechanisms associated with these schemes.

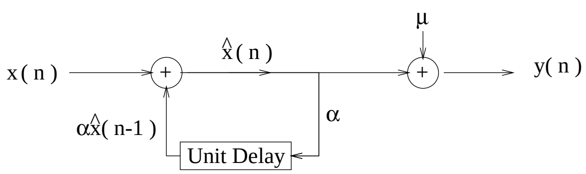


Fig. 2. Traffic generating scheme: $x(n)$ is the zero-mean white Gaussian samples, α controls the randomness (burstiness) of traffic, and μ determines the mean.

In order to verify the correctness of our analysis as well as to test its robustness we carry out simulations using both synthesized traffic and traces from real video and network traffic. The synthesized layered video traffic and cross traffic are generated using Gaussian random signals. The Gaussian random signals are generated using a first-order *Auto Regressive* (AR) model as shown in Figure 2. Under this model, the simulated traffic $y(n)$ is generated using zero-mean white Gaussian samples $x(n)$ which pass through a feedback loop with a parameter α where $\alpha^2 < 1$, and then are added witha mean μ. The mean $E[Y(n)]$ of the generated traffic is μ, and the variance $Var[Y(n)]$ is $Var[X(n)]/(1-\alpha^2)$. The mean and variance are controlled by the parameters μ and α. In our simulations we have chosen μ and α such that the probability of generating negative signals is extremely small, and in simulated traffic they actually never appear.

We consider two scenarios. In the first scenario, we use stationary synthesized layered video traffic and cross traffic. The simulation results serve as validation tests for our theoretical analysis and the optimal receiver layer adaptation we identified through analysis. In the second scneario, we use traces from real video and network traffic to further validate our analysis. The simulation results obtained in these scenarios demonstrate both the robustness of our analysis as well as the superior performance of the optimal statistical receiver layer adaptation scheme.

A. Scenario I: Stationary Traffic

Table II lists the traffic and system parameters used in our simulations. In the discrete-time simulations we perform, the time unit is assumed to be $1/30$ second, which corresponds to one frame unit of time, given a video of 30 frames/sec frame rate. The receiver layer adaptation delay $n_d = 30$ is thus about 1 sec, which we deem it to be roughly realistic. At the end of this subsection we will investigate the impact of n_d on the performance of the greedy receiver layer adaptation scheme, and show that the performance data does justify our belief.

Given the traffic statistics, the layer inference matrix **S** and (uncongested) mean traffic matrix **M** calculated using (12) and (15) are shown below:

$$\mathbf{S} = [0.0186\ 0.2427\ 0.4546\ 0.2190\ 0.0520\ 0.0130]^T, \quad (20)$$

$$\mathbf{M} = [749820\ 830670\ 676130\ 268420\ 62073\ 12343]^T. \quad (21)$$

From (21) we see that M_1 is the largest among all M_i, $i = 0, 1..., 5$. Hence the optimal level of subscription is 1. Therefore the optimal receiver layer adaptation scheme is always to join layer 1.

mean of layered video	500000
mean of cross-traffic	750000
α of layered video	0.6
α of cross-traffic	0.7
variance of layered video	7.6562x10^9
variance of cross-traffic	4.902x10^9
λ_i	0.2
L	5
B	1000000
n_d delay time	30
Simulation time	10000

TABLE II

TRAFFIC AND SYSTEM PARAMETERS USED IN SCENARIO I

$$T_0 = \begin{pmatrix} 1&1&1&1&1&1\\ 0&0&0&0&0&0\\ 0&0&0&0&0&0\\ 0&0&0&0&0&0\\ 0&0&0&0&0&0\\ 0&0&0&0&0&0 \end{pmatrix} \quad T_1 = \begin{pmatrix} 0&0&0&0&0&0\\ 1&1&1&1&1&1\\ 0&0&0&0&0&0\\ 0&0&0&0&0&0\\ 0&0&0&0&0&0\\ 0&0&0&0&0&0 \end{pmatrix} \quad T_2 = \begin{pmatrix} 0&0&0&0&0&0\\ 0&0&0&0&0&0\\ 1&1&1&1&1&1\\ 0&0&0&0&0&0\\ 0&0&0&0&0&0\\ 0&0&0&0&0&0 \end{pmatrix}$$

$$T_{\text{greedy}} = \begin{pmatrix} 1&0&0&0&0&0\\ 0&1&0&0&0&0\\ 0&0&1&0&0&0\\ 0&0&0&1&0&0\\ 0&0&0&0&1&0\\ 0&0&0&0&0&1 \end{pmatrix}$$

$$T_3 = \begin{pmatrix} 0&0&0&0&0&0\\ 0&0&0&0&0&0\\ 0&0&0&0&0&0\\ 1&1&1&1&1&1\\ 0&0&0&0&0&0\\ 0&0&0&0&0&0 \end{pmatrix} \quad T_4 = \begin{pmatrix} 0&0&0&0&0&0\\ 0&0&0&0&0&0\\ 0&0&0&0&0&0\\ 0&0&0&0&0&0\\ 1&1&1&1&1&1\\ 0&0&0&0&0&0 \end{pmatrix} \quad T_5 = \begin{pmatrix} 0&0&0&0&0&0\\ 0&0&0&0&0&0\\ 0&0&0&0&0&0\\ 0&0&0&0&0&0\\ 0&0&0&0&0&0\\ 1&1&1&1&1&1 \end{pmatrix}$$

Fig. 3. The simulated schemes, where $\mathbf{T}_1$ is the optimal scheme, and $\mathbf{T}_{greedy}$ is the greedy scheme.

In addition to comparing the performance of the optimal receiver layer adaptation and the greedy receiver layer adaptation, we also consider several other receiver layer adaptation schemes. This would help illustrate that the optimal receiver layer adaptation indeed has the *best* performance among all these schemes, not just better than the greedy receiver layer adaptation scheme. All the schemes considered in our simulations are shown in their transition matrix representation in Figure 3. $\mathbf{T}_0$ is the trivial scheme under which the receiver receives no layers of the video traffic. $\mathbf{T}_1$ is the optimal scheme. For $i = 2, \ldots, 5$, $\mathbf{T}_i$ is the scheme which always joins i layers. In particular, $\mathbf{T}_5$ is the most aggressive scheme, which receives all the five layers of the video traffic. $\mathbf{T}_{greedy}$ is the greedy receiver layer adaptation scheme.

Simulation results for these schemes are summarized in Table III. In the table, "Eff Theory" denotes the theoretical system efficiency computed using (18). "Eff Simulation" denotes the system efficiency obtained from the simulations (using the sample-path system efficiency formula (2)). "Ptheory" and "Psimulation" denote respectively the congestion probability computed from theory and observed from the simulations.

From the table, it is no surprise that the congestion probability for $\mathbf{T}_0$ is 0. As additional layers are subscribed, both the theoretical/simulated congestion probabilities for $\mathbf{T}_i$ increase from around $0.0187/0.0149$ for $\mathbf{T}_1$ to $0.9870/0.9871$ for $\mathbf{T}_5$. Again it is no surprise that $\mathbf{T}_5$ has the worst congestion probability both in theory and in simulation. In comparison, $\mathbf{T}_{greedy}$ has a moderate congestion probability (around 0.3416/0.3421) both in theory and simulation. Discounting the trivial scheme $\mathbf{T}_0$, the optimal receiver adaptation scheme has the lowest congestion probability. Similarly, the optimal receiver layer adaptation scheme also results in the best system efficiency both in theory and simulation. Whereas, the greedy receiver layer adapta-

Eff Theory	Eff Simulation	Ptheory	Psimulation	Adapt. Scheme
0.7498	0.7517	0	0	$\mathbf{T}_0$
0.8307	0.8321	0.0187	0.0149	$\mathbf{T}_1$
0.6761	0.6825	0.2613	0.2542	$\mathbf{T}_2$
0.2684	0.2599	0.7159	0.7254	$\mathbf{T}_3$
0.0621	0.0617	0.9349	0.9367	$\mathbf{T}_4$
0.0124	0.0125	0.9870	0.9871	$\mathbf{T}_5$
0.5848	0.5854	0.3416	0.3421	$\mathbf{T}_{greedy}$

TABLE III

SIMULATION RESULTS FOR SCENARIO I.

tion scheme only has a system efficiency slightly below 0.6 both in theory and in simulation. Note in particular that because no video layers are subscribed, $\mathbf{T}_0$ has a system efficiency around 0.75. This relatively high system efficiency is obtained at the expense of the receiver, since it receives no video information at all. Last but not the least, we point out that the theoretical results we obtained through analysis are very close to the simulation results. This empirically verifies the correctness of our analysis.

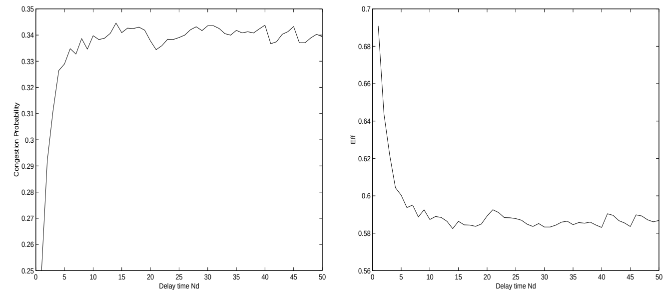
In our analysis we have assumed that the receiver layer adaptation delay n_d is relatively large such that the independence among the involved processes can be assumed. We now investigate the impact of n_d on the greedy receiver layer adaptation scheme[5]. In Figure 4, the (simulated) congestion probability and system efficiency for the greedy receiver adaption scheme are plotted as a function of n_d. From the plots, we see that the greedy receiver layer adaptation scheme has fairly good performance when n_d is relatively small. When n_d becomes large, its performance degrades quickly. In particular, when n_d is larger than 5, its performance seems to stabilize and fluctuate within a small range. This shows that after $n_d = 5$, the correlation among the traffic can be ignored. Note that $n_d = 5$ corresponds roughly to a receiver layer adaptation delay of $5/30$ sec, or approximately 167 msec. This is comparable to expected delay in receiver layer adaptation which includes such factors as delay in request for joining/leaving a multicast group, delay in grafting/pruning the new receiver to/from an existing multicast tree, and delay in finally forwarding video data to the new receiver. Lastly, as an aside, we comment that with $n_d = 1$, the system efficiency (≈ 0.6909) and the congestion probability (≈ 0.2499) of the greedy receiver layer adaptation scheme are still worse than those ($Eff^{optimal} = 0.8307$, and $P^{optimal}_{congestion} = 0.01870$) of the optimal receiver layer adaptation scheme.

B. Scenario II: Real Traffic Traces

In the previous scenario, simulations are conducted using synthesized traffic to verify the correctness and robustness of our analysis. In this section we carry out simulations using traces from real video and network traffic to further demonstrate the validity of our analysis.

The setup of the simulations is similar to that of Scenario I, except that the bandwidth of the bottleneck link is now 400 Kb/sec. Since true layered video traces are not available, we

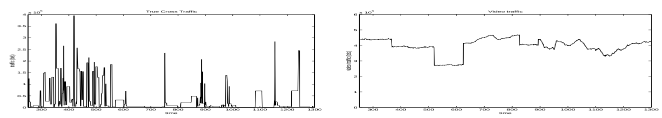
[5]Since the receiver layer adaptation schemes $\mathbf{T}_i$, $i = 0, 1, \ldots, 5$, always join i layers (in steady state). Hence n_d has minimal or no impact on their performance in steady state.



(a) Congestion Probability *vs.* n_d

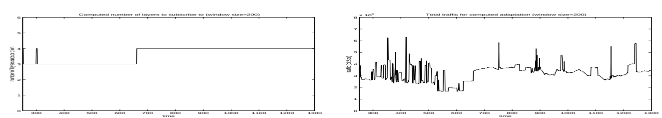
(b) *Eff vs.* n_d

Fig. 4. Impact of n_d on the performance of the greedy scheme.



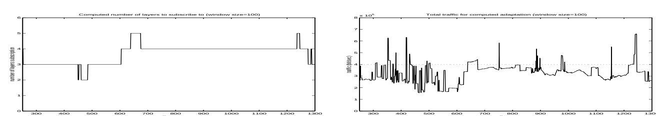
(a) cross traffic trace

(b) video trace



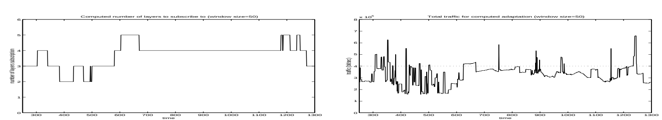
(c) subscribed layers: optimal statistical scheme, $W = 200$

(d) total traffic: optimal statistical scheme, $W = 200$



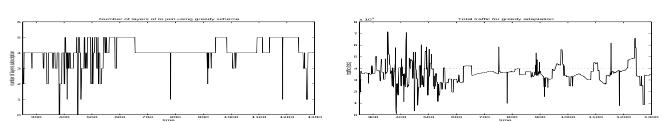
(e) subscribed layers: optimal statistical scheme, $W = 100$

(f) total traffic: optimal statistical scheme, $W = 100$



(g) subscribed layers: optimal statistical scheme, $W = 50$

(h) total traffic: optimal statistical scheme, $W = 50$



(i) subscribed layers: greedy scheme

(j) total traffic: greedy scheme

Fig. 5. Scenario II: the number of layers subscribed over the time under various schemes and their corresponding total traffic B_T.

Simulated $Eff_{200}^{optimal}$	0.7263	Congestion Prob.	0.0657
Simulated $Eff_{100}^{optimal}$	0.7035	Congestion Prob.	0.0933
Simulated $Eff_{50}^{optimal}$	0.6942	Congestion Prob.	0.1105
Simulated Eff^{greedy}	0.6605	Congestion Prob.	0.1924

TABLE IV
SIMULATION RESULTS FOR SCENARIO II

generate a layered video trace from an JPEG video trace. The JPEG video trace we use is a 1300 frame sequence from the movie *Sleepless in Seattle* [6], which is shown in Figure 5 (b). The frame rate is 30 frames per second, hence a frame interval is 1/30 sec. We divide each frame of the JPEG sequence evenly into five segments to generate a video trace with five layers. Although the generated layered video trace may not represent a *true* layered video trace, we believe that it still reflects the multiple time-scale variability exhibited by the original video trace. The cross traffic comes from a portion of a WAN traffic trace from a real network[6]. Since the original format of the WAN traffic trace is in terms of packet arrival times, we have to transform it into a cross traffic trace that suits our simulation setting. This is done by measuring the average rate of the WAN traffic trace during the interval and sampling traffic at every frame interval (i.e., 1/30 sec). Since the bandwidth of the bottleneck link is set to 400 Kb/s, traffic samples which are larger than 400 Kb/sec are truncated. Figure 5 (a) shows the cross traffic trace thus generated. It is clear that the cross traffic is quite bursty.

Since both video and cross-traffic traces are non-stationary, we use *slide window* with size W=200,100 and 50 to estimate the statistics of the traces to conduct the statistical adaptation simulations. The results are shown in Figure 5. Figures 5 (c), (e), (g) and (i) (in the left column) show the number of layered subscribed under the three versions of the optimal statistical receiver layer adaptation scheme and the greedy receiver layer adaptation. The corresponding total traffic (i.e., the aggregation of the cross traffic and the subscribed layered video traffic) transmitted across the bottleneck link under the various schemes are shown in Figures 5 (d), (f), (h) and (j). From the figures we can see that the optimal statistical receiver layer adaptation scheme is less sensitive to the burstiness of the cross traffic than the greedy receiver adaptation scheme. In particular, a larger sliding window leads to less oscillation in layer join/leave activities.

The congestion probability and system efficiency of all the schemes are shown in Table IV. From the table we see that the three versions of the optimal statistical receiver layer adaptation scheme all outperform the greedy receiver layer adaptation scheme both in terms of the system efficiency and the congestion probability. These simulation results based on the real traffic traces again demonstrate empirically the robustness of our analysis and the superiority of the optimal statistical receiver layer adaptation scheme.

[6]The cross traffic trace is downloaded from http://ita.ee.lbl.gov/html/contrib/BC.html.

VI. CONCLUSIONS

In this paper we have presented an analytical model to investigate the issue of optimal receiver layer adaptation for layered video multicast transmission. Based on this model, we introduced the notion of congestion probability as an indirect measure of the perceived layered video quality at a receiver. The notion of system efficiency is introduced and used as a metric to study the performance of a receiver layer adaptation scheme. Through analysis, we have determined the optimal receiver layer adaptation scheme which maximizes the system efficiency while providing best sustainable video quality to a receiver. We have showed that the greedy receiver layer adaptation scheme is indeed optimal if the traffic generated by various layers of a layered video as well as the cross traffic are constant. However, in a more realistic setting where layered video traffic or the cross traffic varies over the time, the greedy receiver layer adaptation scheme is in general not optimal. To verify our theory, we conducted simulations using both synthesized traffic and traces from real video and network traffic. Our simulation results demonstrate that the optimal statistical receiver adaptation scheme yields optimal results closely.

REFERENCES

[1] C. Albuquerque, B. Vickers, and T. Suda, "An End-to-End Source-Adaptive Multi-Layered Multicast (SAMM) Algorithm," *ICS Technical Report 98-31*, University of California, Irvine, September 1998.

[2] S. Bajaj, L. Breslau, and S. Shenker, "Uniform versus Priority Dropping for Layered Video," *Proceedings of ACM SIGCOMM'98*, Vancouver BC., September 1998.

[3] T. Brown, S. Sazzad, C. Schroeder, P. Cantrell, and J. Gibson, "Packet Video for Heterogeneous Networks using CU-SeeMe," *Proceedings of IEEE ICIP'96*, Lausanne, Switzerland, September 96.

[4] N. Chaddha, and A. Gupta, "A Framework for Live Multicast of Video Streams over the Internet," *Proceedings of IEEE ICIP'96*, Laussane, Switzerland, September 1996.

[5] S. E. Deering, "Multicast Routing in a Datagram Internetwork," Ph.D. Thesis, Stanford University, December 1991.

[6] Wu-chi Feng, "Video-on-Demand Services: Efficient Transportation and Decompression of Variable Bit Rate Video", Ph.D. Thesis, Univ. of Michigan, April 1996.

[7] P.-C. Hu, Z.-L. Zhang and M. Kaveh, "A Framework for Proxy-Based Receiver Adaptation for Layered Video Transmission in Multicast Networks," *Proceedings of IEEE ICIP'99*, Kobe, Japan, October 1999.

[8] X. Li, S. Paul, P. Pancha, and M. Ammar, "Layered Video Multicast with Retransmission (LVRM): Evaluation of Error Recovery Schemes," *Proceedings of IEEE NOSSDAV*, St. Louis, Missouri, May 1997.

[9] S. McCanne, V. Jacobson, and M. Vetterli, " Receiver-driven Layered Multicast," *Proceedings of ACM SIGCOMM'96*, pp. 117-130, Stanford, CA, August 1996.

[10] A. Papoulis, "Probability, Random Variables, and Stochastic Processes, 3rd edition," McGraw-Hill, Singapore, 1991.

[11] R. Ramanujan, A. Ahamad, and K. Thurber, "Traffic Control Mechanism to Support Video Multicast over IP Networks," *Proceedings of IEEE International Conference on Multimedia Computing and Systems '97*, pp. 85-94, Ottawa, Ont., June 1997.

[12] N. Shacham, "Multipoint Communication by Hierarchically Encoded Data," *Proceedings IEEE INFOCOM'92*

[13] T. Streams, L. Wu, R. Sharma and B. Smith, "An Architecture for Multicasting Layered Video," *Proceedings of NOSSDAV'97*, St. Louis, Missouri, May 1997.

[14] D. Taubman and A. Zakhor, "Multirate 3-D Subband Coding of Video," *IEEE Transactions on Image Processing*, pp. 572-588, vol. 3, no.5, Sep. 1994.

[15] L. Vicisano, J. Crowcroft, and L. Rizzo, "TCP-like Congestion Control for Layered Multicast Data Transfer," *Proceedings of IEEE INFOCOM'99*, pp. 996-1003, vol. 3, New York, NY, 1999.