

# Buffer Schemes for VBR Video Streaming over Heterogeneous Wireless Networks

Guang Ji, Ben Liang, and Aladdin Saleh

**Abstract**—With the co-existence of different wireless networks, which exhibit largely different bandwidth and coverage characteristics, much interest has been involved in integrating these networks to support smooth and efficient multimedia services. In this paper, we present an analytical framework for variable-bit-rate (VBR) video streaming in a two-tier wireless network with VBR channels. We derive the expected number of jitters and average buffering delay during video playback as measures of system performance. Our objective is to discover heterogeneous networking attributes that may influence the streaming performance, in terms of the tradeoff between jitter frequency and buffering delay. Through experimenting with a wide range of fixed, separate, and jointly optimal jitter-recovery buffering schemes, based on buffering delay, buffered data, and buffered playback duration, we quantify the benefit of incorporating user location information in streaming over heterogeneous wireless networks.

## I. INTRODUCTION

Media streaming applications have distinctive Quality of Service (QoS) requirements, such as delay sensitiveness and loss tolerance. In addition, the varying wireless environment brings in dramatic fluctuation of network bandwidth [1], which makes the streaming technology even more challenging. Meanwhile, recent trends indicate that wide-area cellular network (CELL), e.g., 3G network, Wireless Local Area Networks (WLANs), e.g., IEEE 802.11, will co-exist to offer seamless wireless multimedia services [2]. Such integration enables the users to enjoy better streaming performance while exploiting the complementary advantages of different networks.

In order to protect against the influence of the wireless link fluctuation, transmitted video packets are temporarily stored at the receiver buffer, which sustains streaming when network throughput is low. When there are not enough data in the buffer to support the video playback consumption, a playback starvation occurs, which is also known as *buffer underflow*. Then video stops playing until sufficient data are gathered. This event of playback interruption is usually termed *playout jitter*, and the time duration for data buffering is termed *buffering delay*. Intuitively, the longer buffering delay is employed, the more packets will be received and the smaller jitter occurrence probability we can anticipate in the future, but at the same time the user viewing experience is correspondingly degraded due to the increased waiting time. Therefore, it is especially important to properly balance the tradeoff between buffering delay and jitter frequency for a media streaming system.

Guang Ji and Ben Liang are affiliated with the Department of Electrical and Computer Engineering, University of Toronto; email: {gji, liang}@comm.toronto.edu. Aladdin Saleh is affiliated with Bell Mobility; email: aladdin.saleh@bell.ca. This work is funded in part by Bell Canada through its Bell University Laboratories R&D program.

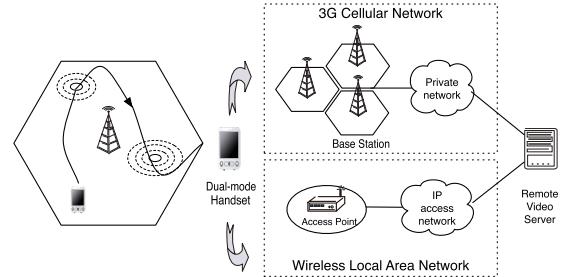


Fig. 1. A typical illustration of video streaming system for mobile users in two-tier wireless networks. The dual-mode handset user is able to switch between 3G cellular network and WLAN.

In this paper, we study the problem of streaming on-demand *variable-bit-rate* (VBR) video over heterogeneous wireless networks<sup>1</sup>. The video is pre-encoded with variable bit rate and stored in a remote *media server* that can be accessed through both tiers of the network, which are labeled “CELL” and “WLAN” for illustration purposes without loss of generality. Mobile end users view the videos streams while roaming in the two-tier network. A typical video streaming system in heterogeneous wireless networks is presented in Figure 1.

We initiate an analytical model for the mobile end user’s receiver buffer. Then the expected jitter frequency during the whole streaming session is derived. Furthermore, in order to evaluate the user-perceived streaming media quality, we adopt a cost function combining the jitter numbers and average buffer delay during the entire playback. We first examine the performance of *Fixed Buffering Schemes* which employ the same buffering parameters, such as the fixed buffering delay, fixed buffered playout data, and fixed playout time. Then we consider *Separate Buffering Schemes* which use different buffering parameters for CELL and WLAN. The parameters are obtained from the analysis of the wireless networks separately. We also study *Jointly Optimal Buffering Schemes*, which select optimal buffering parameters directly from the heterogenous networks under certain average buffering delay constrains. Through extensive analysis and simulation, we compare theses three families of schemes to find appropriate buffering methods for mobile devices with various level of storage memory and computation power.

The rest of the this paper is organized as follows. We discuss the related work in Section II. The system model is presented in Section III. We derive the analysis framework for video streaming process over heterogeneous works in Section IV. Section V presents our comparisons for different buffering

<sup>1</sup>VBR media provides better quality for the same average bit rate by adapting quantization and compression to the time-varying entropy of the media [3].

schemes. Simulation results and further discuss are provided in Section VI. We conclude the paper in Section VII.

## II. RELATED WORK

Several smoothing techniques deal with network link transfers of stored video streams [4] [5] [6], but they only consider a wired network which offers guaranteed bandwidth service and an intermediate smoothing node is required. Hence these schemes are not suitable for error-prone wireless network streaming systems. Varsa *et al.* [7] proposed a separation between a delay jitter buffer and a decoder buffer for VBR video. The delay jitter buffer is particular designed to compensate for delay jitters and bit rate variations caused by variable bit rate channel. But in [8], the authors compared the single receiver buffer with the separate buffer, and concluded that the single receiver buffer performs at least as good as the two separate buffers. In this work, we focus on a single-buffer design for heterogeneous wireless networks.

Studies in [9] show that the pattern of packet loss can be captured by Markov models. Kalman *et al.* used a Markov chain analysis method in [10] to examine the tradeoff between buffer underflow probability and latency for adaptive playout video streaming. Adaptive media playout allows the streaming client to control the data consumption rate, but can introduce noticeable artifacts in the displayed video. Xu *et al.* [11] considered the transmission of prerecorded media from a server to a client by using TCP-Friendly Rate Control (TFRC). The models focus on the impact of the TFRC rate changes to the probability of rebuffering events and analytically study its impact on media quality. This work does not consider buffering delay as a performance metric.

The authors previously presented in [12] an analytical framework to study the frequency of jitters and buffering delays under the constraint of initial playback delay and receiver buffer size, using a Markov VBR channel model for a homogeneous wireless network. The family of fixed buffering schemes are examined. In this work, we investigate further into separate and jointly optimal buffering schemes for heterogeneous wireless networks. To the best of our knowledge, this paper represents the first attempt to analyze buffering mechanisms for media streaming over heterogeneous wireless networks.

## III. SYSTEM MODEL

### A. Network Channel Model

We consider the video streaming over a two-tier network, using CELL-WLAN integration as an example. In general, CELL provides universal coverage, with WLAN forming several hotspots. The mobile clients use dual-mode handset which enables the network access switching between CELL and WLAN when necessary. We assume the mobile users will automatically switch to WLAN service when traversing into the overlapping of the two networks, in order to obtain a potentially higher data throughput.

The streaming video process is considered to be time-discrete with equal time slots. In each time slot, multiple video packets are sent to the mobile user. The base stations of CELL

and access points in WLAN forward video sequences to the roaming users independently. Clearly, the number of packets transmitted per time slot in WLAN is much larger than that in CELL.

In this work, we assume the transport channel to be error free, possibly due to an ideal error control mechanism or concealment scheme, but the network transmission rate may change over time. Noting the delay-sensitivity of the video streaming technology, the fluctuations in transmission rate may possibly lead to late packet arrivals and significant playback interruptions. In each of CELL and WLAN, we can model the network transmission channel as a discrete-time Markov variable-bit-rate channel [9]. Following the common assumption of exponential network residence times in CELL and WLAN, the transitions between these two sub-networks are memoryless. Hence, we can characterize the overall channel status over time by a Markov chain  $(\mathcal{S}, \mathcal{T}, \mathcal{R})$ , where  $\mathcal{S}$  is the set of possible channel states,  $\mathcal{T}$  is the transition probability matrix of the channel states, and  $\mathcal{R}$  is the set of possible transmission rates associated with the state. We define  $P_{ij}$ , where  $i, j \in \{c, w\}$  as the probability that the user will be in subnetwork  $j$  in the next time slot given she is in subnetwork  $i$  in the current time slot. For example,  $P_{wc}$  is the transition probability from WLAN to CELL.

The following is an example on how to combine the channels states in CELL and WLAN. To characterize the error events in the wireless communication channel, a simple and widely used model is the *Gilbert-Elliott* model with states  $\Omega \in \{\text{good}, \text{bad}\}$  [13]. The network state can be transmitted from good to bad by losing one packet, or from bad to good by receiving one packet. Sanneck and Carle further proposed in [14] an extended Gilbert model for the wireless channels, which is able to provide better prediction of performance measures depending on longer-term correlation of errors. We start with CELL constructed as an  $M$ -state extended Gilbert model while WLAN an  $N$ -state extended Gilbert model (Figure 2). Thus, the channel states set becomes  $\mathcal{S} = \{S_1, S_1, S_2, \dots, S_{M+N}\}$ , where  $S_1$  and  $S_{M+1}$  are the *good* or *reception* states for CELL and WLAN respectively. Other states are the *bad* or *loss* states for the two subnetworks with different loss patterns. For example,  $S_2$  represents two consecutive packet losses. In combining the two extended Gilbert models, we add subnetwork transitions and adjust correspondingly the transition probabilities in the original models. Figure 2 shows an example of our channel model. Note that in this model, we further restrict the transitions between subnetworks, so that only the *reception* states can be the destination of such transitions. We emphasize that this is adopted only as a common-sense assumption. The general analytical model presented in Section IV is applicable to all transition patterns.

### B. Receiver Buffer Model

The video receiver of the mobile terminal consists of a playout buffer and a playout scheduler. The playout buffer is used to temporarily store the incoming video packets. We denote the total number of video packets as  $L$  and the duration

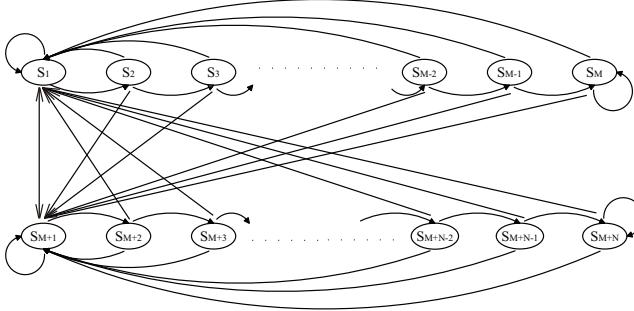


Fig. 2. Channel state transitions in two-tier wireless network

of the video as  $T$ . Let  $p(t)$  be the playback schedule which describes the total amount of packets which *should be* received at time  $t$ . Thus we have  $p(T) = L$ . Denote  $r(t)$  as the entire number of packets which are successfully arrived at the receiver at  $t$ . Then, if  $r(t) < p(t)$ , i.e., *buffer underflow*, a jitter occurs and further buffering is required. Playout is assumed only after enough packets are aggregated, which is termed *jitter recovery*. Furthermore, if the buffer size is finite, there may be instances that the incoming packets numbers exceeds the buffer limit. Then these packets will be lost due to *buffer overflow*. In this case, we assume that playout scheduler will send control signals to the video server requesting retransmission of the loss packets in the next time slot.

The playout scheduler is responsible of managing the buffering schemes. We consider three types of buffering schemes, based on the *buffering delay* (BD), the buffered *playout data* (PD), and the buffered *playout time* (PT) [12]. One common setting of the playout scheduler is to use a fixed BD, PD, or PT after each jitter. However, in heterogeneous networks, a fixed scheme would give no consideration of the user mobility or the present network conditions. In this work, we investigate into appropriate buffering schemes for video streaming in a two-tier wireless network. We first present the analytical model for video streaming in the next section.

#### IV. JITTER AND DELAY ANALYSIS

In this section, we extend the analysis framework proposed in [12] to a two-tier network, in order to derive the expected number of jitters given different jitter recovery buffering schemes. Due to page limitation, some details are omitted and can be found in [12].

We index the incoming video packets with  $i$ , where  $i$  is an integer ranging from 1 to the total number of packets of the video source  $L$ . Let  $J_n$  denote the time index of the video when the  $n$ th jitter occurs, and  $X_n$  the channel state when the  $n$ th jitter occurs. We use  $R^c$  and  $R^w$  to denote the maximum numbers of packets transmitted per time slot in CELL and WLAN, respectively. For the extend Gilbert model, they simply are the transmission rates in the good states.

We define  $p_k^{(n)}(i) = \Pr\{J_n = i, X_n = S_k\}$  as the probability that the  $n$ th jitter occurs at packet  $i$  with channel state  $S_k$ . Then, the expected number of jitters  $E\{J\}$  can be

expressed as [15]

$$\begin{aligned} E\{J\} &= \sum_{n=0}^{\infty} \sum_{i=1}^n (\sum_{k=1}^{M+N} p_k^{(n)}(i)) \\ &= \underbrace{\sum_{n=0}^{\infty} \sum_{i=1}^n \sum_{k=1}^M p_k^{(n)}(i)}_{\text{Number of jitters in CELL}} + \underbrace{\sum_{n=0}^{\infty} \sum_{i=1}^n \sum_{k=M+1}^{M+N} p_k^{(n)}(i)}_{\text{Number of jitters in WLAN}} . \end{aligned} \quad (1)$$

In order to obtain  $p_k^{(n)}(i)$ , we specify  $Q_{k,l}(i,j)$ , the probability that the  $(n+1)$ th jitter takes place at packet  $j$  with channel state  $S_l$ , given that the  $n$ th jitter occurs at packet  $i$  with channel state  $S_k$ . Therefore, we have

$$Q_{k,l}(i,j) = \Pr\{J_{n+1} = j, X_{n+1} = S_l | J_n = i, X_n = S_k\} . \quad (2)$$

Applying the total probability theorem, we have

$$p_l^{(n+1)}(j) = \sum_{i=1}^j \sum_{k=1}^{M+N} Q_{k,l}(i,j) p_k^{(n)}(i) . \quad (3)$$

In this way, with the first jitter probability  $p_k^{(1)}(i)$ , and the next jitter probabilities  $Q_{k,l}(i,j)$ , we are able to obtain the entire statistics of  $p_k^{(n)}(i)$  to calculate the expected number of jitters.

We model the streaming system states with two tuples  $(g, s)$ , where  $g$  denotes the total number of received packets till current time and  $s \in \mathcal{S}$  specifies the channel state. We define  $P_{l,k,r}$  as the transition probability from state  $S_l$  to state  $S_k$  with  $r$  packets successfully transmitted in the time slot. We construct the following transition matrix:

$$\Psi = \begin{bmatrix} A_0 & A_1 & A_2 & \dots & A_{R^c} & A_{R^c+1} & \dots & A_{R^w} & 0 & 0 \\ 0 & A_0 & A_1 & A_2 & \dots & A_{R^c} & A_{R^c+1} & \dots & A_{R^w} & 0 \\ \vdots & \vdots & & & \ddots & & & & & \\ 0 & 0 & 0 & \dots & \dots & 0 & A_0 & A_1 & A_2 & \dots \\ 0 & 0 & 0 & 0 & \dots & \dots & 0 & A_0 & A_1 & A_2 \\ \vdots & \vdots \end{bmatrix} ,$$

where

$$A_r = \begin{bmatrix} P_{1,1,r} & \dots & \dots & P_{1,M+N,r} \\ \vdots & \ddots & & \vdots \\ \vdots & \ddots & & \vdots \\ P_{M+N,1,r} & \dots & \dots & P_{M+N,M+N,r} \end{bmatrix} .$$

We assume the video streaming starts at time  $-\Delta$ , where  $\Delta$  is the initial delay. Denote  $\pi_0$  as the initial system state distribution and  $\pi_t$  the system state distribution at time  $t$ . With the transition matrix  $\Psi$ , we can easily obtain  $\pi_t = \pi_0 \Psi^{t+\Delta}$ . However, in order to calculate the first jitter probability, what we are interested is to find the probability that the system reaches to a state without any jitter by time  $t$ . Instead, at each time  $t$ , we only consider the probabilities of the states which do not violate the playout constraints. In other words, we set the distribution probability  $\pi_t[(M+N)g + l]$  to 0 for  $g < p(t)$ ,  $l = 1, \dots, M+N$ . Furthermore, considering the possible limitation of receiver buffer size  $B$ , we can have received at most  $p(t) + B$  packets at time  $t$ . In this case, we merge the transitions into states with  $g \geq p(t) + B$  to the states with  $g = p(t) + B + 1$ . Therefore, we modifying  $\Psi$  by  $\Psi U_t$ , where

$$U_t = \begin{bmatrix} 0_{(M+N)p(t) \times (M+N)p(t)} & 0 & 0 \\ 0 & I_{(M+N)(B-1) \times (M+N)(B-1)} & 0 \\ 0 & 0 & I' \\ 0 & 0 & 0 \end{bmatrix},$$

and  $I' = [\underbrace{I_{(M+N) \times (M+N)} \cdots I_{(M+N) \times (M+N)}}_{R^w+1}]^T$ .

Then we have

$$\pi_t = \pi_0 \left( \prod_{s=-\Delta}^{t-1} \Psi U_t \right) \Psi. \quad (4)$$

(1) *Buffering Delay.* After each jitter, the stream stops and data is buffered for a certain buffering delay. We denote the buffering delay by  $D^c$  for CELL and  $D^w$  for WLAN. In order to find  $Q_{k,l}(j,i)$ , we imagine the video starts playing out from the jitter occurring time with an empty buffer. Denote this virtual initial state distribution, after a jitter occurs at  $t_j$  and the channel state is  $S_l$ , by  $\pi_{j,l} = [0 \cdots 0 1 0 \cdots 0]$ , where 1 is the  $((M+N)(j-1)+l)$ th element. Then, the state probability distribution at time  $t$  of having no jitter by  $t-1$  is given by

$$\begin{cases} \pi_t^c = \pi_{j,l} (\prod_{s=-D^c}^{t-1} \Psi U_t) \Psi, & \text{if } l \in [1, M] \\ \pi_t^w = \pi_{j,l} (\prod_{s=-D^w}^{t-1} \Psi U_t) \Psi, & \text{if } l \in [M+1, M+N]. \end{cases}$$

Finally, the  $Q_{l,k}(j,i)$  is obtained by

$$Q_{l,k}(j,i) = \begin{cases} \pi_t^c [(M+N)(i-1)+k], & \text{if } j \in [1, M] \\ \pi_t^w [(M+N)(i-1)+k], & \text{if } j \in [M+1, M+N]. \end{cases}$$

(2) *Buffered Playout Data.* After each jitter, the stream stops and data are buffered until the number of packets in the buffer reaches a certain predetermined amount. We denote the buffered playout data by  $B^c$  for CELL and  $B^w$  for WLAN. We first find the probability distribution of the states  $(g,s)$  when the playout restarts. Suppose the jitter occurs in CELL, we construct a Markov chain of this buffering state with the transition probability matrix:

$$\Phi = \begin{bmatrix} j-1 & A_0 A_1 \cdots A_{R^w} & 0 & \cdots & 0 \\ j & 0 & A_0 A_1 \cdots A_{R^w} & 0 & \cdots \\ \vdots & \vdots & \ddots & & \vdots \\ j+B^c-2 & 0 & \cdots & 0 & A_0 A_1 \cdots A_{R^w} \\ j+B^c-1 & 0 & \cdots & 0 & I & 0 & \cdots \\ \vdots & \vdots & & \ddots & \ddots & 0 \\ j+B^c+R^c-2 & 0 & & \cdots & 0 & I \end{bmatrix}. \quad (5)$$

Once the system enters into any one of state in  $[j+B^c-1, j+B^c+R^c-2]$ , it exits the jitter recover buffering state. Hence, these states are modeled as absorption states. We can obtain the distribution of these states by solving the absorption probabilities of the Markov chain [15], which leads to the state distribution  $\pi_{j,l}^c$ . In the same way, we can obtain the state distribution  $\pi_{j,l}^w$  for the case when a jitter occurs in WLAN. Thus

$$\begin{cases} \pi_t^c = \pi_{j,l}^c (\prod_{s=1}^{t-1} \Psi U_t) \Psi, & \text{if } l \in [1, M] \\ \pi_t^w = \pi_{j,l}^w (\prod_{s=1}^{t-1} \Psi U_t) \Psi, & \text{if } l \in [M+1, M+N]. \end{cases}$$

Then,  $Q_{l,k}(j,i)$  can be calculated in the same way as the BD scheme.

(3) *Buffered Playout Time.* After each jitter, the stream stops and data is buffered until the amount of buffered data can sustain a certain amount of playout time. We denote the buffered playout time by  $T^c$  for CELL and  $T^w$  for WLAN. The process to find  $Q_{l,k}(j,i)$  is similar to that in the PD scheme, except the number of packets to buffer is  $p(d_j + T^c) - (j-1)$  for jitters happening in CELL, and the number of packets to buffer is  $p(d_j + T^w) - (j-1)$  for jitters happening in WLAN.

## V. BUFFERING SCHEMES IN HETEROGENEOUS NETWORKS

Ideally, video frames should be displayed continuously with each successive frame displayed immediately after its predecessor. However, due to the unstable network situations, continuous playout is not always possible, especially for streaming over wireless networks. Clearly, there exists a tradeoff between the jitter occurrences and the average buffering delay after each jitter. A superior buffering scheme should strike a balance between the two factors that leads to an overall optimized user satisfactory. Thus, we introduce a cost function  $C$  as the weighted sum of the expected number of jitters  $E\{J\}$  and the average jitter-recovery buffering delay  $\bar{D}$ :

$$C = (1-\alpha)\bar{D} + \alpha E\{J\}. \quad (6)$$

where  $\alpha$  is the weight parameter ranging from 0 to 1, indicating the video viewer's preference.

(1) *Fixed Buffering Schemes.* In the fixed buffering schemes, the same buffering parameter value is used in both subnetworks. For each  $\alpha$ , the parameter that minimizes (6) is chosen. Thus, we have Fixed Buffering Delay (FBD), where  $D^c = D^w$ , Fixed Buffered Playout Data (FPD), where  $B^c = B^w$ , and Fixed Buffered Playout Time (FPT), where  $T^c = T^w$ .

The fixed buffering schemes can be easily implemented, as they don't have to take into consideration the network conditions or the mobile user location. However, it cannot provide optimal performance the heterogeneous wireless networks. It is possible that designed buffering amount is too large for users in WLAN or too small for users in CELL.

(2) *Separate Buffering Schemes.* In the separate buffering schemes, we first find, independently for each type of subnetwork, the optimal buffering parameter value that minimizes (6). We then use them in the two-tier network. Thus, we have Separate Buffering Delay (SBD), Separate Buffered Playout Data (SPD), and Separate Buffered Playout Time (SPT).

The separate buffering schemes consider the different subnetworks separately. Surprisingly, our numerical results in Section VI show that the performances of these schemes generally do not improve over the fixed buffering schemes.

(3) *Jointly Optimal Buffering Schemes.* In the jointly optimal buffering schemes, we find the optimal pair of buffering parameter values:

$$\begin{aligned} \text{Minimize} \quad & C = (1-\alpha)\bar{D} + \alpha E\{J\} \\ \text{subject to} \quad & \alpha \in [0, 1], \bar{D} \in \mathcal{D} \end{aligned} \quad (7)$$

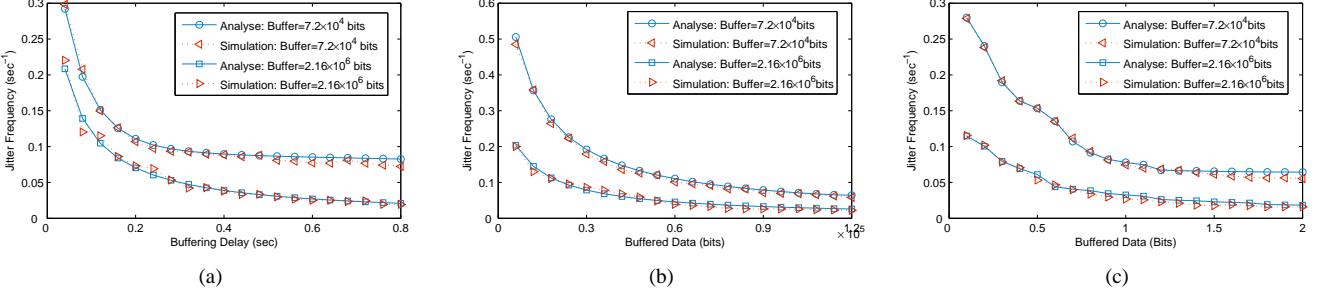


Fig. 3. Analysis and simulation results for fixed buffering schemes with different buffer sizes: (a) FBD (b) FPD (c) FPT

Thus, we consider the combination of  $(D^c, D^w)$  for Jointly Optimal Buffering Delay (JBD),  $(B^c, B^w)$  for Jointly Optimal Buffered Playout Data (JPD), and  $(T^c, T^w)$  for Jointly Optimal Buffered Playout Time (JPT).

The jointly optimal buffering schemes provide improved performance over fixed and separate buffering schemes. However, they involve much higher computational complexity.

## VI. SIMULATION RESULTS

In this section, we validate our analytical results and evaluate the proposed buffering schemes via detailed simulations.

### A. Simulation Setup

We use the “Alpin ski” MPEG-4 variable-bit-rate video trace provided by [16]. The video sequences were encoded at a constant frame rate of 25 frames/s in the Quarter Common Intermediate Format (QCIF) resolution. Table I summarizes the main parameters of the video trace in the simulation. The packet size is set to 1800 bytes and the transmission time slot duration is 80 ms.

Parameter	Value
Sequence Length	89998 frames
Video Size	$7.2e + 06$ bytes
Format	QCIF ( $176 \times 144$ pixels)
Video Run Time	$1.6e + 06$ msec
Mean Bit Rate	$1.9e + 05$ bps
Peak Bit Rate	$1.8e + 06$ bps

TABLE I  
VIDEO PARAMETERS IN SIMULATION

We assume that the mobile user is initially located in the cellular network. We use a two-state Gilbert model for each of the wireless networks, so that for the two-tier system, we have totally four states:  $S_1$ ,  $S_2$ ,  $S_3$ , and  $S_4$ , where  $S_1$  and  $S_3$  are the good states in CELL and WLAN respectively, and  $S_2$  and  $S_4$  are the corresponding bad states. For CELL, the transition probabilities from “good” to “bad” and reverse are 0.2 and 0.5, respectively, before alterations due to user movement. For WLAN, they are 0.05 and 0.4. The subnetwork transition probabilities are  $P_{cw} = 0.005$  and  $P_{wc} = 0.01$ . The data rate in  $S(1)$  180 kbps, i.e. 1 packets per unit time slot, while data rate in  $S(3)$  is 1.8 Mbps, i.e. 10 packets per time slot. Hence, the average rate in CELL is 90 kbps, and that in WLAN is about 1.2 Mbps.

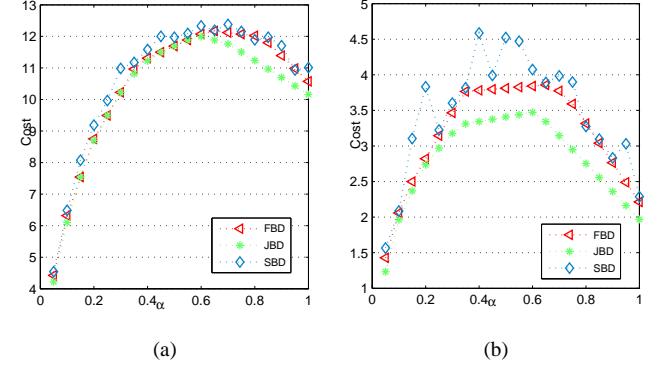


Fig. 4. Comparison of FBD, SBD, and JBD: (a) buffer=  $7.2 \times 10^4$  bits. (b) Infinite buffer.

We simulate the transmission and playback for the target sequence in Matlab for over 500 realizations of the random VBR channel and obtain the average jitter numbers and average buffering delay. The initial delay  $\Delta$  is set to 0.4 seconds for all cases.

### B. Model Validation

We compare the analytical and simulation results of the expected number of jitters for different average buffering delay values. Fig. 3 shows the comparison for fixed buffering schemes with different buffer sizes. We observe a good match between the simulation and analysis results. Moreover, as expected, the mean number of jitters decreases as the buffering delay increases. Note that the variations in the analysis curve of Fig. 3(c) is due to the VBR nature of the video. The comparison for other schemes are similar and is omitted to reduce redundancy.

### C. Comparison of Buffering Schemes

We compare the fixed, separate, and jointly optimal buffering schemes. Fig. 4, Fig. 5, and Fig. 6 show the results for BD, PD, and PT, respectively. In each case, both finite and infinite buffers are studied.

All three figures show an approximately convex shape for the cost function over  $\alpha$ . Recall that a small  $\alpha$  favors the cost of jitters over the cost of buffering delay. This suggests that in general it is easier to reduce only one of either the number of jitters or the buffering delay, and harder to strike a balance between the two. Indeed, if we allow the buffer delay

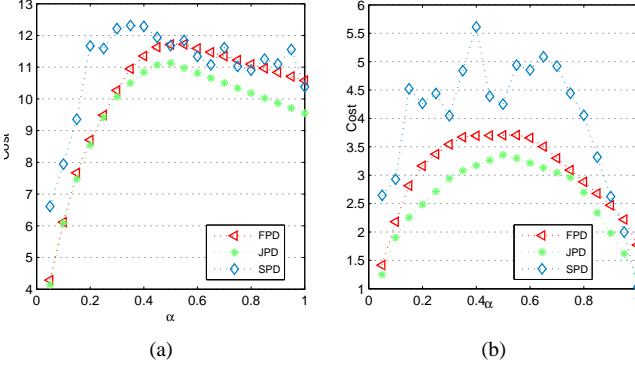


Fig. 5. Comparison of FPD, SPD, and JPD: (a) buffer =  $7.2 \times 10^4$  bits. (b) Infinite buffer.

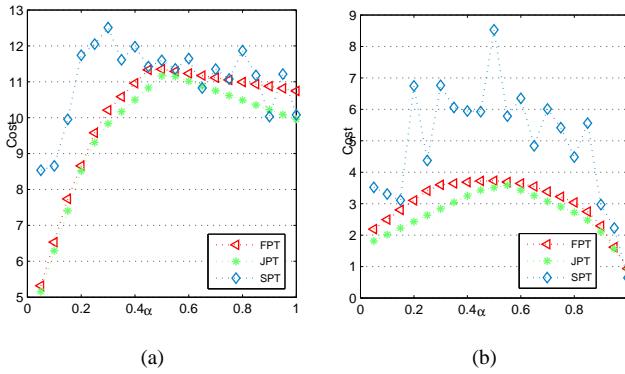


Fig. 6. Comparison of FPT, SPT, and JPT: (a) buffer =  $7.2 \times 10^4$  bits. (b) Infinite buffer.

to become very large to maximize the amount of data buffered, we can significantly reduce jitters.

These figures also show that the jointly optimal buffering schemes indeed perform the best. They also show an interesting phenomenon. Even though the parameters chosen for the separate buffering schemes are individually optimal in each subnetwork, their application to the two-tier network actually degrades the system performance, often to a degree worse than the fixed buffering schemes. This suggests that there exists strong correlation between streaming performance and the user mobility dynamics between the subnetworks, so that the subnetworks should not be considered separately in optimizing the performance of streaming in a heterogeneous network.

Comparing the three figures, we see that, when there is no limit on the buffer size and the buffering parameters are optimally chosen, JBD, JPD, and JPT all give similar performance. Furthermore, in this case, FBD, FPD, and FPT all give similar degradation from the optimal performance. However, if buffer size is limited, and the fixed buffering scheme is used, then FPT outperforms FPD and FBD. This result is unique to heterogeneous networks and is in contrast to [12], which shows that in homogeneous networks FPT, FPD, and FBD perform similarly in terms of jitter frequency and buffering delay. It suggests that buffering based on playout time can be more adaptive to the streaming client's movement between the subnetworks. Hence, mobile devices with limited storage memory and insufficient computation power to

produce network-aware optimal buffering parameters should adopt FPT in heterogeneous wireless networks.

## VII. CONCLUSION

We have studied a wide variety of buffering schemes for VBR video streaming over heterogeneous networks, including fixed, separate, and jointly optimal schemes. These schemes can be based on buffering delay, buffered playout data, or buffered playout time. We model the video transmission process for mobile clients roaming within the network using a two-tier Markov variable-bit-rate channel model and analyze the jitter and delay characteristics of such systems. Our analytical and simulation results suggest that the streaming performance can be significantly improved by utilizing the location information of a mobile client, but separate optimization within the subnetworks is unsuitable. Furthermore, we show that buffering based on playout time is more appropriate for simple mobile devices that have limited storage memory and uses a constant buffering parameter regardless of location.

## REFERENCES

- [1] M. Etoh and T. Yoshimura, "Advances in Wireless Video Delivery," *Proceedings of the IEEE*, vol. 93, no. 1, pp. 111–122, Jan. 2005.
- [2] D. Axiotis, T. Al-Gizawi, E. Protonotarios, F. Lazarakis, C. Papadias, and P. Philippopoulos, "Services in interworking 3G and WLAN environments," *IEEE Wireless Commun. Mag.*, vol. 11, no. 5, pp. 14–20, Oct. 2004.
- [3] T. V. Lakshman, A. Ortega, and A. R. Reibman, "Variable bit rate (VBR) video: Tradeoffs and potentials," *Proceedings of the IEEE*, vol. 86, no. 5, pp. 952–973, May 1998.
- [4] S. Sen, J. L. Rexford, J. K. Dey, J. F. Kurose, and D. F. Towsley, "Online smoothing of variable-bit-rate streaming video," *IEEE Trans. Multimedia*, vol. 2, no. 1, pp. 37–48, Mar. 2000.
- [5] W. E. Naylor and L. Kleinrock, "Stream traffic communication in packet switched networks: Destination buffering considerations," *IEEE Trans. Communication*, vol. 30, no. 12, pp. 2527–2534, Dec. 1982.
- [6] P. Thiran, J. Y. L. Boudec, and F. Worm, "Network calculus applied to optimal multimedia smoothing," in *Proc. of IEEE INFOCOM*, 2001.
- [7] V. Varsa and I. Curcio, "Transparent end-to-end packet switched streaming service (pss); rtp usage model (release 5)," *3GPP TR 26.937 V1.4.0*, 2003.
- [8] T. Stockhammer, H. Jenkac, and G. Kuhn, "Streaming video over variable bit-rate wireless channels," *IEEE Trans. Multimedia*, vol. 6, no. 2, pp. 268–277, Apr. 2002.
- [9] M. Yajnik, S. Moon, J. Kurose, and D. Towsley, "Measurement and modelling of the temporal dependence in packet loss," in *Proc. of IEEE INFOCOM*, 1999.
- [10] M. Kalman, E. Steinbach, and B. Girod, "Adaptive media playout for low-delay video streaming over error-prone channels," *IEEE Trans. Circuits and System for Video Technology*, vol. 14, no. 6, pp. 841–851, Jun. 2004.
- [11] L. Xu and J. Helzer, "Media streaming via tfrc: An analytical study of the impact of tfrc on user-perceived media quality," in *Proc. of IEEE INFOCOM*, Apr. 2006.
- [12] G. Liang and B. Liang, "Balancing interruption frequency and buffering penalties in vbr video streaming," in *Proc. of IEEE INFOCOM*, May 2007.
- [13] E. N. Gilbert, "Capacity of a burst-noise channel," *Bell Syst. Tech. J.*, vol. 39, no. 5, pp. 1253–1265, Oct. 1960.
- [14] H. Sanneck, G. Carle, and R. Koodli, "A framework model for packet loss metrics based on loss run length," in *Proceedings of SPIE/ACM SIGMM Multimedia Computing and Networking Conference*, Nashville, USA, Jan. 2000.
- [15] A. Papoulis and S. U. Pillai, *Probability, Random Variables, and Stochastic Processes*, 4th ed. McGraw-Hill, 2002.
- [16] F. H. P. Fitzek and M. Reisslein, "Mpeg-4 and h.263 video traces for network performance evaluation," *IEEE Network*, vol. 15, no. 6, pp. 40–54, Nov. 2001.