2004 IEEE International Conference on Multimedia and Expo (ICME)

Bandwidth Estimation for Wireless Video Transmission

Ji-An Zhao[1], Bo Li[1]*, Ishfaq Ahmad[2]

[1]Department of Computer Science
Hong Kong University of Science & Technology
Clear Water Bay, Kowloon, Hong Kong, China
Email: {zhaojian,bli}@cs.ust.hk

[2]Department of Computer Science and Engineering
University of Texas at Arlington
Arlington, TX 76019-0015, USA
Email: iahmad@cse.uta.edu

Abstract

This paper deals with layered video transmission over wireless networks. We focus on deriving the required bandwidth provisioning for each layered video given their respective QoS target. We model the layered video traffic by a discrete time batch Markovian arrival process (DBMAP) with marked transitions. We assume the link level behavior of the wireless channel can be modeled by a hidden Markov model (HMM), and the network supports automatic repeat request (ARQ) operation. We show that the video data transmission buffer can be modeled as a $G/D/c$ queue with time dependent feedback, and propose to approximate the ARQ feedback traffic by an HMM modulated DBMAP process. Based on the effective bandwidth approach, we derive the required channel capacity for both the input video traffic and the ARQ feedback traffic in order to meet the given QoS targets.

1 Introduction

Transporting multimedia data, especially real-time video traffic over wireless networks is a challenging task. Prior studies reveal that layered video encoding is a key element to provide smooth video communication in wireless environment [6]. In this paper we take the effective bandwidth [4] approach to quantitatively estimate how much bandwidth should be allocated for layered video transmission in wireless networks, such that the given QoS targets can be satisfied. The paper is organized as follows, in section 2, we show that the video data transmission buffer can be modeled by a $G/D/c$ queue with time dependent feedback; in section 3, we derive the required channel capacity for the input video traffic; in section 4, we approximate the ARQ feedback traffic by an HMM modulated DBMAP; we present numerical results in section 5, and conclude the paper in section 6.

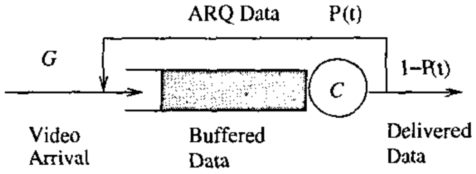

Figure 1: $G/D/c$ queue with feedback

2 The $G/D/c$ Queue with Feedback

The basic transmission unit for the wireless network is defined by a radio link control (RLC) data block, which can be transmitted within one time slot. Application data such as video packets are first segmented into RLC blocks, and then fed into the wireless transmission buffer waiting for transmission. Since VBR video traffic is highly dynamic, we expect the video arrival process is a general (G) random process.

We ignore the physical layer details of the wireless channel and assume that the link level behavior can be modeled by a hidden Markov model (HMM) [5]. We further assume that the network supports automatic repeat request (ARQ) operation, and that with time-dependent probability $P(t)$, the error data blocks will re-enter the queueing buffer for retransmission. We refer the retransmission data as the *feedback traffic* or the *ARQ traffic*. Since we consider real-time video applications, if the video data can not be successfully delivered within a certain time limit, they will be discarded. In wireless networks like cdma2000 or WCDMA, bandwidth is allocated in the unit of channels. If we assume $\mu(t) = c$ channels are allocated, then the link level video data transmission buffer can be modeled as a $G/D/c$ queue with time dependent feedback, as shown in Figure 1.

---

*Correspondent author.

3 Effective Bandwidth for the Video Traffic

For simplicity, we consider video source encoded in two layers, i.e., with the base layer and one enhancement layer. We assume the video traffic can be modeled by a discrete time batch Markovian arrival process (DBMAP) with *marked transitions*, which has nice property to capture the inherent bursty and correlated nature of the video traffic, as demonstrated in [7]. Assume the parameters for the marked DBMAP process are given by $D_{i_1 i_2}$, $0 \leq i_1 \leq k_1$, $0 \leq i_2 \leq k_2$, where k_1 and k_2 are the maximum arrival batch size for the base and enhancement layer video traffic, respectively. Obviously, each layer of the video traffic is a DBMAP [2], and the base and enhancement aggregation traffic is also a DBMAP.

Let X represent the buffer occupancy, b as size of the buffer, and p be the overflow probability. Definition of the effective bandwidth is based on performance criteria on the tail probability $P(X > b) \leq p$ for the related queueing system [4]. When the traffic arrival process is a DBMAP with parameters $\{D_0,\ D_1,\ \ldots\}$, the effective bandwidth is related to the probability matrix generating function $D(z) = \sum_{i=0}^{\infty} D_i z^i$. Let $\delta(z) = \lim_{t\to\infty}[\pi D(z)^t e]^{1/t}$, where π is the stationary probability vector of the DBMAP, and e is a column vector of 1's. It has been shown in [3] that $\delta(z)$ is equal to the Perron-Frobenius eigenvalue of $D(z)$, and the effective bandwidth function $\gamma(z)$ for the DBMAP traffic arrival process can be expressed by $\gamma(z) = \log(\delta(z))/\log(z)$. The *effective bandwidth* for the DBMAP traffic source is given by $\gamma(z^*)$, where $z^* = p^{-1/b}$.

We assume that logically each layer of the video data occupies one transmission queue, and that the network transmits the base layer video data with a higher priority. Denote d_i as the maximum queueing delay; p_i as the tolerable loss rate; X_i as the buffer occupancy; and b_i as the threshold buffer size; for layer i of the video traffic. We assume the QoS requirements of the real-time video application are given by $\{d_1,\ d_2,\ p_1,\ p_2\}$, and the QoS constraints are given by $P(X_1 > b_1) < p_1$ and $P(X_2 > b_2) < p_2$. We *approximate* b_1 and b_2 by $b_1 = \lambda_1(d_1+1)$ and $b_2 = \lambda_2(d_2+1)$, where λ_1 and λ_2 are the average data arrival rates, d_1+1 and d_2+1 are the maximum sojourn time since the service time for each data block is one time slot.

We now estimate the required channel capacity for the input video traffic by making use of the call admission control algorithm in [1]. Denote $z_1^* = p_1^{-1/b_1}$, $z_2^* = p_2^{-1/b_2}$, the required channel capacity c for the input video traffic has to comply with the following constrains:

$$\begin{aligned} \gamma_1(z_1^*) &\leq c \ \text{(for base layer only)}, \\ \gamma_1(z_2^*) + \gamma_2(z_2^*) &\leq c \ \text{(for base + enhance)}, \quad (1) \end{aligned}$$

where $\gamma_i(.)$ is defined for layer i of the video traffic. Here $\gamma_1(z_2^*)$ can be interpreted as the base layer effective bandwidth that is subject to the QoS criterion of the enhancement layer, or in other words $\gamma_1(z_2^*)$ is the effective bandwidth for the base layer traffic as seen by the enhancement layer traffic.

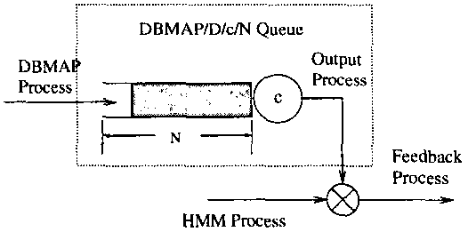

Figure 2: Approximation of the feedback traffic

4 Effective Bandwidth for the Feedback Traffic

In order to reserve certain capacity for ARQ operation, the overall allocated bandwidth should be somewhat larger than (1). We refer this extra channel capacity as *the ARQ bandwidth.* The exact analysis of the $G/D/c$ feedback traffic is cumbersome, as it depends not only on the input traffic, but also on the error behavior of the wireless channel. Therefore, we approximate the feedback traffic by the output process of the $DBMAP/D/c/N$ queue that is further modulated by the HMM channel process, as depicted in Figure 2. Here c denotes the channel rate, $N = b_1 + b_2$ denotes the maximum queue size.

Examining the $DBMAP/D/c/N$ queue at the end of each time slot, we obtain a Markov chain with transition matrix T as shown in Figure 3. Since the channel rate is c, the output process can generate i $(0 \leq i \leq c)$ data blocks during each time slot. When the $DBMAP/D/c/N$ queue is idle, there is no output from the queue; when the queue size is no less than c, the queue generates a batch departure with group size of c; otherwise the queue generates a batch departure with group size smaller than c and greater than 0. Therefore, the output process is an N state DBMAP with maximum batch size c. The parameters of the output DBMAP are given by $\{\tilde{D}_0,\ \tilde{D}_1,\ \tilde{D}_2,\ \ldots\ \tilde{D}_c\}$, Numerical results: effective bandwidth for *Foreman*, *Paris* and *Grandma*

Video Sequence[1] & Mean Rate		Per-Channel Rate[2]	Channel Model	Effective BW[3]	(in kbps)
Foreman B	5.19	$\frac{181bits}{20ms} = 9.05kbps$	Error free	5.64	(51.0)
Foreman B+E	5.19 + 2.62	-	-	5.55+4.32	(89.3)
Foreman B	5.19	-	Geometric	5.64+[0.28]	(53.6)
Foreman B+E	5.19 + 2.62	-	-	5.55+4.32+[0.48]	(93.7)
Foreman B	5.19	-	HMM	n/c	n/c
Foreman B+E	5.19 + 2.62	-	-	5.55+4.32+[0.58]	(94.6)
Paris B	5.10	$\frac{592bits}{20ms} = 29.6kbps$	Error free	5.14	(152.1)
Paris B+E	5.10 + 5.64	-	-	5.13+10.34	(457.9)
Paris B	5.10	-	Geometric	5.14+[0.26]	(159.8)
Paris B+E	5.10 + 5.64	-	-	5.13+10.34+[0.78]	(481.0)
Paris B	5.10	-	HMM	n/c	n/c
Paris B+E	5.10 + 5.64	-	-	5.13+10.34+[0.96]	(486.3)
Grandma B	5.20	$\frac{181bits}{20ms} = 9.05kbps$	Error free	5.27	(47.7)
Grandma B+E	5.20 + 4.14	-	-	5.26+8.24	(122.2)
Grandma B	5.20	-	Geometric	5.27+[0.26]	(50.0)
Grandma B+E	5.20 + 4.14	-	-	5.26+8.24+[0.67]	(128.2)
Grandma B	5.20	-	HMM	n/c	n/c
Grandma B+E	5.20 + 4.14	-	-	5.26+8.24+[0.83]	(129.7)

Notes: "-" means having the same value as the above entry; "n/c" means not computed.

[1] B = base layer only, B+E = base layer + enhancement layer.

[2] Per-Channel Rate=RLC Block Size/RLC Block Duration.

[2] Effective BW is given as a form of sum: base layer + enhancement layer + [ARQ traffic].

$$T = \begin{bmatrix} B_{0,0} & B_{0,1} & B_{0,2} & \cdots & B_{0,N-3} & B_{0,N-2} & B_{0,N-1} & B_{0,N} \\ B_{1,0} & B_{1,1} & B_{1,2} & \cdots & B_{1,N-3} & B_{1,N-2} & B_{1,N-1} & B_{1,N} \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \vdots \\ B_{c,0} & B_{c,1} & B_{c,2} & \cdots & B_{c,N-3} & B_{c,N-2} & B_{c,N-1} & B_{c,N} \\ 0 & B_{c+1,0} & B_{c+1,1} & \cdots & B_{c+1,N-4} & B_{c+1,N-3} & B_{c+1,N-2} & B_{c+1,N-1} \\ 0 & 0 & B_{c+2,0} & \cdots & \vdots & B_{c+2,N-4} & B_{c+2,N-3} & B_{c+2,N-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & B_{N,0} & \cdots & B_{N,c-1} & B_{N,c} \end{bmatrix},$$

$$\tilde{D}_0 = \begin{bmatrix} B_{0,0} & B_{0,1} & B_{0,2} & \cdots & B_{0,N-3} & B_{0,N-2} & B_{0,N-1} & B_{0,N} \\ 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 \end{bmatrix},$$

$$\tilde{D}_1 = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 \\ B_{1,0} & B_{1,1} & B_{1,2} & \cdots & B_{1,N-3} & B_{1,N-2} & B_{1,N-1} & B_{1,N} \\ 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 \end{bmatrix}, \ldots$$

$$\tilde{D}_{c-1} = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \vdots \\ B_{c-1,0} & B_{c-1,1} & B_{c-1,2} & \cdots & B_{c-1,N-3} & B_{c-1,N-2} & B_{c-1,N-1} & B_{c-1,N} \\ 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 \end{bmatrix},$$

$$\tilde{D}_c = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \vdots \\ B_{c,0} & B_{c,1} & B_{c,2} & \cdots & B_{c,N-3} & B_{c,N-2} & B_{c,N-1} & B_{c,N} \\ 0 & B_{c+1,0} & B_{c+1,1} & \cdots & B_{c+1,N-4} & B_{c+1,N-3} & B_{c+1,N-2} & B_{c+1,N-1} \\ 0 & 0 & B_{c+2,0} & \cdots & \vdots & B_{c+2,N-4} & B_{c+2,N-3} & B_{c+2,N-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & B_{N,0} & \cdots & B_{N,c-1} & B_{N,c} \end{bmatrix}.$$

Figure 3: Output process parameters (below) and Numerical results (above)

as shown in Figure 3. Assume the HMM channel contains m states and the parameters are given by $H = H_0 + H_1$. Evidently, the HMM modulated $DBMAP/D/c/N$ output process is an $N \times m$ state DBMAP. The transition probability matrix for the underlying Markov chain is given by $T \otimes (H_0 + H_1)$, where $\otimes$ denotes the Kronecker product. Further, the parameters of the HMM modulated DBMAP are given by $\{\hat{D}_0, \hat{D}_1, \hat{D}_2, \dots \hat{D}_c\}$, where $\hat{D}_i = \tilde{D}_i \otimes H_1$ for $i = 1, 2, \dots c$, and $\hat{D}_0 = T \otimes (H_0 + H_1) - \sum_{i=1}^{c} \hat{D}_i$. Since the feedback process can be approximated by a DBMAP, we can estimate the ARQ bandwidth in a similar way as that of the input video traffic.

5 Numerical Results

We compute the effective bandwidth for three video sequences *Foreman*, *Paris* and *Grandma*. For all the video sequences, the base layer is encoded with TM5 rate control, while the enhancement layer is encoded with MPEG-4 fine grained scalability. We assume the parameters of the HMM channel are given by

$$H_0 = \begin{bmatrix} 0.9 & 0.07 \\ 0.5 & 0.35 \end{bmatrix}, \; H_1 = \begin{bmatrix} 0.0 & 0.03 \\ 0.0 & 0.15 \end{bmatrix}.$$

By computation, the average error rate for each RLC block in the HMM channel is 0.05. For comparison we also consider a memoryless Geometric channel with equal error probability of 0.05, which in fact is a special case of the HMM channel with a single state.

Assume the QoS parameters are given by $\{p_1 = 0.001,\ p_2 = 0.005,\ d_1 = d_2 = 25\}$, which ensures the base layer benefit from a lower loss rate. We compute all the effective bandwidth values and present the results in Figure 3. For simplicity, we don't compute the feedback effective bandwidth for the case with base layer only traffic. It can be observed that for all the three video sequences, the effective bandwidth is larger than the corresponding average arrival rate of the video data. In particular, the base layer effective bandwidth is rather near from the mean data rate. However, the enhancement layer effective bandwidth is much larger than the mean data rate. The reason is that the base layer is under rate control and has less variation, while the enhancement layer is coded without rate control and has larger variation. For such highly dynamic VBR traffic, the required bandwidth for QoS guarantee is much larger than the mean arrival rate. Therefore, if the bandwidth is stringent, rate control in the base layer is critical to maintain QoS. Also note that the base layer effective bandwidth for the base layer only case is larger than the base plus enhancement layer case. The reason behind this discrepancy is that $\gamma_1(z_1^*) > \gamma_1(z_2^*)$, which is consistent with the observation in [1]. Finally, we note that the effective bandwidth for an HMM channel is larger than an Geometric channel, which confirms the anticipation that channel with correlated errors can downgrade the network performance.

6 Conclusions

In this paper we quantitatively study the bandwidth provisioning problem for wireless video transmission, specifically, given the QoS targets for each layer of video data, how much bandwidth should be provided to satisfy the QoS requirements? Such QoS oriented bandwidth estimation can be used in channel allocation or admission control for real-time video applications in wireless networks. Numerical results show that the effective bandwidth is higher than the mean data rate. The difference is rather small for rate controlled base layer video traffic, but is substantial for the enhancement layer traffic. The results also reveal that an HMM channel requires more bandwidth than a Geometric channel in handling the ARQ traffic.

References

[1] A. W. Berger and W. Whitt, "Effective bandwidths with priorities," IEEE/ACM Transactions on Networking, vol. 6, no. 4, pp. 447-460, August 1998.

[2] C. Blondia, "A discrete time batch Markovian arrival process as B-ISDN traffic model," Belgian Journal of Operations Research, Statistics and Computer Science, vol. 32, no. 3-4, 1993.

[3] F. Ishizaki, "Study on reduction of total bandwidth requirement by traffic dispersion," Proceedings of International Conference on ATM and High Speed Internet 2001, pp. 285-289, Seoul, Korea, 2001.

[4] V. G. Kulkarni, "Effective bandwidths for Markov regenerative sources," Queueing Systems: Theory and Applciations, vol. 24, pp. 137-153, 1996.

[5] W. Turin, Digital transmission systems: performance analysis and modeling, McGrawHill, 1998.

[6] D. Wu, T. Hou, Y.-Q. Zhang, "Scalable video coding and transport over broadband wireless networks," Proceedings of the IEEE, vol. 89, no. 1, pp. 6-20, Jan. 2001.

[7] J.-A. Zhao, B. Li and I. Ahmad, "Traffic model for layered video: an approach on Markovian arrival process," Packet Video 2003, Universite de Nantes, Nantes, France, April 28-30, 2003.