Multi-user Video Streaming over Multi-hop Wireless Networks: A Cross-layer Priority Queuing Scheme

Hsien-Po Shiang and Mihaela van der Schaar
Department of Electrical Engineering (EE), UCLA
{hpshiang, mihaela} @ee.ucla.edu

Abstract

In this paper, we propose a distributed, end-to-end, integrated cross-layer scheme to maximize the decoded video quality of multiple users engaged in simultaneous real-time streaming sessions over a multi-hop wireless network. Our algorithm explicitly considers the distortion impact and delay constraints in assigning priorities to the various packets and then relies on priority queuing to drive the optimization of the various users' transmission strategies across the multi-hop network. The proposed solution is enabled by the scalable coding of the video content and the design of cross-layer optimization strategies including a dynamic routing algorithm, which allow priority-based adaptation to varying channel conditions. Our proposed delay-driven, packet-based transmission is superior in terms of both network scalability and video quality to previous static flow-based solutions based on predetermined paths and rate requirements.

1. Introduction

Emerging multi-hop wireless networks provide a low-cost and flexible infrastructure that can be simultaneously utilized by multiple users for a variety of applications, including delay-sensitive multimedia transmission. However, this wireless infrastructure is often unreliable and provides dynamically varying resources with only limited Quality of Service (QoS) support for multimedia applications. Hence, efficient solutions for multimedia streaming must accommodate time-varying bandwidths and probabilities of error introduced by the shared nature of the wireless medium and quality of the physical connections. In the studied distributed transmission scenario, users need to proactively collaborate in sharing the available wireless resources, in order to ensure that the various multimedia applications are provided with the necessary QoS.

Prior research on multi-user multimedia transmission over multi-hop wireless networks has focused on centralized, flow-based resource allocation strategies based on a pre-determined rate-requirement [1][2]. Such an optimization ensures that the end-to-end utility function (benefit) is maximized while satisfying constraints on individual link capacities. However, the flow-based optimization does not guarantee that explicit packet-based delay constraints are met for video applications. Importantly, they do not take into account the loss tolerance provided by video applications, which can be exploited by the wireless network to support a larger number of users. Therefore, these solutions often lead to inferior network efficiency and suboptimal resulting qualities for the video users.

Alternatively, the majority of the video-centric research does not consider the protection techniques available at lower layers of the protocol stack (MAC, PHY) and/or optimizes the video transport using purely end-to-end metrics, thereby excluding the significant gains of cross-layer design [3][4]. In [5], an integrated cross-layer optimization framework was proposed that considers the video quality impact. However, the solution proposed in [5] considers only the single user case, where a set of paths and transmission opportunities are statically pre-allocated for each video application. In summary, while significant contributions have been made to enhance the separate performance of the various OSI layers, no framework exists that integrates distributed and adaptive routing and resource allocation with cross-layer optimization for efficient multi-user multimedia streaming over multi-hop wireless networks.

In this paper, we propose such an integrated cross-layer solution for multiple video users. In our setting, the importance of the packets is determined based on their contribution to the overall distortion of a particular video as well as their delay deadlines. This information is encapsulated in the header of each transmitted packet and is used by intermediate nodes to drive the cross-layer transmission strategies. The paper is organized as follows. Section II gives the problem formulation and introduces the queuing models required to describe the network system. In Section III, we propose a dynamic routing policy that maximizes the received video quality. Section IV presents our simulation results, and Section V concludes the paper.

2. Problem formulation

2.1. Video sub-flows and network parameters

We assume that there are V video users (with distinct source-destination nodes). We separate each scalable encoded video stream into a certain number of sub-flows (*quality layers*). We assume that the packets within each sub-flow have the same delay deadline, similar to [6]. If the number of sub-flows for video sequence v equals N_v, then the total number of sub-flows that need to be transmitted across all users in the network equals $K = \sum_{v=1}^{V} N_v$. Each packet of sub-flow f_k is associated with the following parameters – the sub-flow's rate requirement R_k, packet length L_k, delay deadline d_k, and its priority represented by the quality slope parameter λ_k. The quality slope parameters are determined at the video sources based on the rate-quality curves[1]. We label the K sub-flows (across all users) in descending order of their priorities, i.e. $\lambda_1 \geq \lambda_2 \geq .... \geq \lambda_k \geq \lambda_{k+1} .... \geq \lambda_K$. Let $\alpha_{k,v}$ be the indicator parameters (which take values of '0' or '1') indicating whether a sub-flow f_k belongs to a video stream v. We assume that the quality-rate curve of the video stream Q_v^{trans} can be modeled as a piecewise linear approximation:

$$Q_v^{trans} = \sum_{k=1}^{K} \alpha_{k,v} \lambda_k R_k \, . \qquad (1)$$

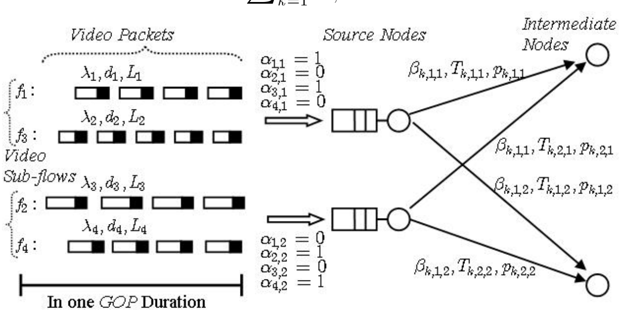

Figure 1. The video multi-flows and the parameters at the first hop.

The network consists of H hops with M_h intermediate nodes at each *h*-th hop ($0 \leq h \leq H-1$). Each node will be tagged with a distinct number m_h ($1 \leq m_h \leq M_h$), and the number of source and destination nodes are the same, i.e. $M_0 = M_H = V$. We define β_{k,h,m_h} as the probability for a packet of sub-flow f_k to select the node m_h as its relay at the *h*-th hop. We refer to this term as the *relay selecting parameter*. (A dynamic approach for determining the relay selecting parameters is proposed in Section III.) Multiple paths could be selected for a sub-flow, i.e. $0 \leq \beta_{k,h,m_h} \leq 1$. Whenever an intermediate node m_h is not reachable for sub-flow f_k, then $\beta_{k,h,m_h} = 0$. Thus, we have $\sum_{m_h=1}^{M_h} \beta_{k,h,m_h} = 1$. Note that since each sub-flow f_k has a pre-determined destination (i.e. $m_H = v$), the relay selecting parameter at the last hop (β_{k,H,m_H}) is equal to '1', if m_H is the destination of the sub-flow, and '0', otherwise. $T_{k,m_h,m_{h+1}}$ and $p_{k,m_h,m_{h+1}}$ are the transmission rate and packet error rate of the link between node m_h and m_{h+1} for sub-flow f_k. The MAC retransmission protection is performed when packet error occurs, and based on our prior results [6], the optimal retransmission strategy is to send the highest priority packet until its successful arrival at the receiver or until its delay deadline expires at an intermediate node.

At each intermediate node, the packets are scheduled according to the quality slope parameter. A video packet is dropped at the intermediate nodes as soon as its delay deadline expires. Assuming that the average end-to-end packet loss probability for the packets belonging to sub-flow f_k is P_k, the expected received video quality of video v can be modeled as:

$$Q_v^{receive} = \sum_{k=1}^{K} \alpha_{k,v} \lambda_k R_k (1 - P_k) \, . \qquad (2)$$

The total received quality across all users is:

$$Q_{tot}^{receive} = \sum_{v=1}^{V} Q_v^{receive} = \sum_{k=1}^{K} \lambda_k R_k (1 - P_k) \, . \qquad (3)$$

2.2. Problem Formulation

The problem of multi-user resource allocation and cross-layer adaptation are usually regarded as a static, centralized optimization that maximizes the total quality of the various users given pre-determined channel (capacity) constraints and video rate requirements [2][4]. These solutions have several limitations. First, the rate requirements of the video streams are time-varying. Second, the delay constraints of the various packets are not explicitly considered. Third, the complexity of the centralized approach grows exponentially with the size of the network and number of video flows. Finally, the capacity region of the network needs to be known for this oracle-based optimization. This is not practical as channel conditions are time-varying and having accurate information about the status of all the network links is not realistic.

To solve the above-mentioned limitations, we formulate the multi-user wireless video transmission problem as a delay-driven cross-layer optimization:

[1] A vast literature on Rate-Distortion models exists for video coding schemes including empirical and information-theoretic based solutions. Here, we use a similar method to that employed in [7].

$$[\beta_{k,h,m_h}]_{0\le h\le H-1} = \arg\max_{\beta} \sum_{k=1}^{K} \lambda_k R_k (1 - P_k(\beta)), \quad (4)$$
$$\text{s.t. } D_k(\beta) < d_k,\ k = 1,...,K,\ \beta \in NET$$

where NET is the set of all feasible choices of relay selecting parameters. All received packets must have their end-to-end delay D_k smaller than their corresponding delay deadline d_k. In an H-hop directed acyclic multi-hop network, P_k can be decomposed based on the hop-by-hop packet loss probability $P_{k,h}$:

$$P_k = 1 - \left(\prod_{h=0}^{H-1} (1 - P_{k,h})\right). \quad (5)$$

2.3. Multi-hop priority queuing model

All the queues in the intermediate nodes are assumed to perform a preemptive-repeat priority M/G/1 model. Let η_{k,h,m_h} be the average arrival rate for the queue at node m_h, and $P_{k,h-1}$ be the packet loss due to delay expiration from the previous hop. $R_{k,h}$ is the updated arrival rate of sub-flow f_k for all the intermediate nodes between the h-th hop and (h+1)-th hop ($1 \le h \le H-1$), and we set $R_{k,0} = R_k$ for the beginning source nodes. Then, the average arrival rates η_{k,h,m_h} have the following recursive relationship:

$$R_{k,h} = (1 - P_{k,h-1}) R_{k,h-1}, \quad (6)$$
$$\eta_{k,h,m} = \beta_{k,h,m} R_{k,h}. \quad (7)$$

Eq.(6) illustrates that the video rate was reduced from hop to hop due to the packet deadline expiration. Eq.(7) shows that the average input rate is distributed based on the relay selecting parameters at the h-th hop.

Assume that X_{k,h,m_h} is the service time of the priority M/G/1 queue at node m_h between the h-th hop and (h+1)-th hop. The first two moment of X_{k,h,m_h} can be obtained from the geometric distribution due to the retransmission nature of the MAC protocol, given the relay selecting parameters β_{k,h,m_h}, the goodput of the associated link. Denote W_{k,h,m_h} as the waiting time of this queue at node m_h for sub-flow f_k. Then, the expected average value can be calculated as:

$$E[W_{k,h,m_h}] = \frac{\sum_{i=1}^{k} \eta_{i,h,m_h} E[X^2_{i,h,m_h}]}{2\left(1 - \sum_{i=1}^{k-1} \eta_{i,h,m_h} E[X_{i,h,m_h}]\right)\left(1 - \sum_{i=1}^{k} \eta_{i,h,m_h} E[X_{i,h,m_h}]\right)}. \quad (8)$$

Therefore, the expectation of the waiting time $E[W_{k,h}]$ over the h-th hop for packets of sub-flow f_k is:

$$E[W_{k,h}] = \sum_{m_h=1}^{M_h} \beta_{k,h,m_h} E[W_{k,h,m_h}]. \quad (9)$$

The probability of packet loss becomes:

$$P_{k,h,m_h} = \text{Prob}\left(W_{k,h,m_h} > d_k - \sum_{j=0}^{h-1} E[W_{k,j}]\right)$$
$$\approx \left(\sum_{i=1}^{K} \eta_{i,h,m_h} E[X_{i,h,m_h}]\right) \exp\left(-\frac{\left(d_k - \sum_{j=0}^{h-1} E[W_{k,j}]\right)\left(\sum_{i=1}^{K} \eta_{i,h,m_h} E[X_{i,h,m_h}]\right)}{E[W_{k,h,m_h}]}\right). \quad (10)$$

The probability of packet loss at the node m_h is the waiting time tail distribution when the accumulated waiting time exceeds the delay deadline. Then, the expected hop-by-hop packet loss rate of the hop h is:

$$P_{k,h} = \sum_{m_h=1}^{M_h} \beta_{k,h,m_h} P_{k,h,m_h}. \quad (11)$$

3. Self-learning policy for dynamic routing

The relay selecting parameters provide a routing description across the network with multi-path capability. We propose a dynamic routing policy for the relay selecting parameters in this section. Each node maintains and feeds back the information of the expected delay from the current node itself to the destination, $E[Delay_{k,m_h}]$, for each class of traffic (sub-flow f_k). Note that each sub-flow has exactly one destination node v predetermined from the indicator parameters $\alpha_{k,v}$. The decisions are made by the following policy:

$$\beta_{k,h,m_h} = \frac{Coeff_k}{1 + \kappa E[Delay_{k,m_h}]^{\varphi}}. \quad (12)$$

$Coeff_k$ are normalized coefficients to make sure that β_{k,h,m_h} still holds as a probability:

$$Coeff_k = \left[\sum_{m_h \text{ that feedback}} \frac{1}{1 + \kappa E[Delay_{k,m_h}]^{\varphi}}\right]^{-1}, \quad (13)$$

where κ and φ are constants. Eq.(12) is inspired from the balking arrival probability in queuing theory. The term φ weights the average delay $E[Delay_{k,m_h}]$ such that the routing policy favors paths leading to a significant lower delay to the destination. We set $\beta_{k,h,m_h} = 0$ for the nodes whose information feedback is not received to avoid sending packets through this undesired region. We refer to this relay selecting policy as *self-learning policy*, since the decision of β_{k,h,m_h} will influence the future information feedback.

The expected delay to the destination of each sub-flow is constantly updated at each node using the information feedback from the next hop. For example, if the current node is node m_h at the h-th hop, the expectation of delay to the destination of each sub-flow is as follows:

$$E[Delay_{k,m_h}] = E[W_{k,m_h}] + \sum_{m_{h+1}=1}^{M_{h+1}} \beta_{k,h+1,m_{h+1}} E[Delay_{k,m_{h+1}}] \quad (14)$$

where $E[Delay_{k,m_{h+1}}]$ is provided from the information feedback of the nodes of the next hop, and the relay selecting parameter $\beta_{k,h+1,m_{h+1}}$ is calculated based on Eq.(12). The self-learning policy will dynamically adapt the relay selection to minimize the delay through the network.

This method is inspired by the Bellman-Ford shortest path (delay) routing algorithm that minimizes the end-to-end delay across the network. Our routing algorithm reduces to the well-known Bellman-Ford algorithm when $\beta_{k,h,m_{h+1}} = 1$ to the node m_{h+1} that feedbacks the smallest $E[Delay_{k,m_{h+1}}]$. Note that since the packet losses result from the violation of the delay constraint, the minimum end-to-end delay provided by our algorithm leads to the minimum end-to-end packet loss. The delay of sub-flow f_k is influenced only by the same or higher priority traffic. The overall received video quality is optimized in the sense that the packets with higher priority (larger λ_k) have minimum end-to-end packet loss P_k.

4. Simulation results

Two video sequences "Mobile", and "Coastguard" are compressed using a scalable video codec. Each bitstream is separated into 4 sub-flows ($N_v = 4, K = 8$). The PSNR-rate curves of the two video streams are modeled by Eq.(2). Based on this, λ_k and R_k are determined for each sub-flow. The delay deadline d_k is set to 0.533 seconds for all sub-flows, which is the interval of one GOP (16 frames per GOP at a frame rate of 30 Hz). Figure 2 shows the analytical average queue waiting time of each sub-flow in the six-hop wireless network. The various Tm represents different level of transmission rates over the links. (ranging from 0.3 Mbps to 0.6Mbps.) The packet error rates for the links are set to 10%. The results show the end-to-end probability of packet loss for each traffic class using our priority queuing approach. Once the bottleneck of a sub-flow (from P_{k,h,m_h}) is formed in the network, the packets of the sub-flow are dropped

In Table 1, we compare the proposed self-learning policy with a state-of-the-art routing algorithm [8] – "Optimal Fixed Path", which statically selects the links for transmission such that the goodput is maximized. Using the same network as in Figure 2, the simulation results show that the proposed dynamic routing approach significantly outperforms the static optimization algorithm, since it provides the ability to alleviate congestion and interference.

5. Conclusions

In this paper, we present a novel distributed cross-layer streaming algorithm for the transmission of multiple videos over a multi-hop wireless network. The essential feature behind our approach is the priority queuing, based on which, the most important video packet is selected and transmitted, at each intermediate node, over the most reliable link, until the transmission success or the deadline expiration. Importantly, we introduce a self-learning policy for dynamic routing that minimizes the end-to-end packet loss for each sub-flow of the video streams. The proposed distributed cross-layer algorithm is also fully adaptive to changes in the network, number of users, priorities of the users.

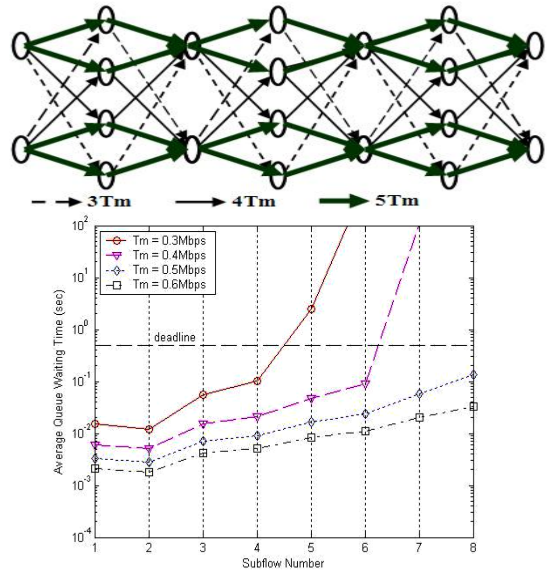

Figure 2. The six-hop network settings and the analytical average queue waiting time of the eight sub-flows.

Table 1. Comparison of video quality

Method	**Tm** = 0.3 (Mbps) Low Network Efficiency		**Tm** = 0.6 (Mbps) Medium Network Efficiency	
	"Mobile" Y-PSNR (dB)	"Coastguard" Y-PSNR (dB)	"Mobile" Y-PSNR (dB)	"Coastguard" Y-PSNR (dB)
Optimal Fixed Path	24.93	30.50	31.07	34.18
Self-Learning Policy	30.48	33.43	33.10	35.61

References

[1] W. Wei, and A. Zakhor, "Multipath unicast and multicat video communication over wireless ad hoc networks," *Proc. Int. Conf. Broadband Networks*, Broadnets, pp. 496-505, 2002.

[2] E. Setton, T. Yoo, X. Zhu, A. Goldsmith, and B. Girod, "Cross-layer design of Ad hoc Networks for real-time video streaming," *IEEE Wireless Communications Mag.*, pp. 59-65, Aug 2005.

[3] A. Butala, L. Tong, "Cross-layer Design for Medium Access Control in CDMA Ad-hoc Networks," *EURASIP J. on Applied Signal Processing*, to appear.

[4] Y. Wu, P. A. Chou, Q. Zhang, K. Jain, W. Zhu, S.Y. Kung, "Network Planning in Wireless Ad Hoc Networks: A Cross-Layer Approach", *IEEE Journal on Selected Areas in Communications*, vol. 23, no. 1, pp. 136-150, Jan. 2005

[5] Y. Andreopoulos, N. Mastronarde, and M. van der Schaar, "Cross-layer Optimized video Streaming over wireless multi-hop Mesh Networks," *IEEE Journal on Selected Areas in Communications*. special issue on multi-hop wireless mesh networks, to appear.

[6] M. van der Schaar, D. S. Turaga, "Cross-layer Packetization and Retransmission Strategies for Delay-sensitive wireless Multimedia Transmission," *IEEE Transactions on Multimedia*, to appear.

[7] A. Ortega, and K. Ramchandran, "Rate-distortion Methods for Image and Video Compression," *IEEE Signal Processing Mag.*, vol. 15, issue 6, pp. 23-50, Nov, 1998

[8] D. B. Johnson, and D. A. Maltz, "Dynamic source routing in ad hoc wireless networks," Chapter in *Mobile Computing*, Kluwer Acad. Pub., 1996.