DISTRIBUTED RATE ALLOCATION FOR MULTI-STREAM VIDEO TRANSMISSION OVER AD HOC NETWORKS

Xiaoqing Zhu and Bernd Girod

Information Systems Laboratory, Stanford University, CA, U.S.A.
{zhuxq,bgirod}@Stanford.EDU

ABSTRACT

When multiple video streams are present in an ad hoc network, they share and compete for the common network resources. A rate allocation algorithm must balance the available resources among the video streams in a fair and efficient manner. In addition, it is desirable to have a distributed solution so that nodes do not need to collect global information of the network and that the computational burden can be shared.

We propose a distributed rate allocation algorithm which minimizes the total distortion of all video streams. Based on the subgradient method, the proposed scheme only requires link price updates at each relay node based on local observations and rate adaptations at each source node derived from rate-distortion (RD) models of the video. Simulation results show that the proposed scheme can achieve the same optimal rate allocation as that obtained from exhaustive search.

1. INTRODUCTION

An ad hoc wireless network is a collection of wireless nodes which communicate with each other without the help of any fixed infrastructure. Each node can be a source, a destination, or a relay for traffic. Since deployment of such a network is fast and flexible, it is attractive to support real-time media streaming over the ad hoc network. Potential applications range from multimedia home entertainment systems to audiovisual communications in disaster areas. However, the limited power and bandwidth of the wireless nodes, combined with the demanding rate and delay requirements of video streaming, give rise to many technical challenges. In particular, rate allocation, scheduling and routing of the video packets need to be designed for efficient utilization of the network resources without overwhelming any individual link.

When multiple video streams are present in the network, they share and compete for the common resources such as transmission power and media access opportunities. In this case, rate allocation implicitly serves the purpose of resource allocation among the streams. Typically, a sender at a higher video source rate utilizes more network resources and achieves better quality of the transmitted video. Moreover, the rate-distortion tradeoff for different video streams are usually different, hence it is preferable to allocate rates to the video streams with a steeper slope in rate-distortion tradeoff. In general, the rate allocation algorithm should achieve fair and efficient resource allocation among the streams. It is also desirable to have a distributed algorithm so that the computational burden is shared among all participating nodes, that there is no overhead of collecting global information of the network, and that the solution can easily adjust to fluctuations in the wireless network conditions.

For data transmission over conventional networks such as the Internet, the problem of distributed multi-user rate allocation has been well studied. Solutions such as the TCP congestion control [1] and TCP-Friendly Rate Control (TFRC) [2] allow independent rate adjustment at each sender based on end-to-end observations. However, as the transport protocol is unaware of the different rate-distortion tradeoff of the video applications, the allocated rates are usually sub-optimal in terms of video distortion minimization. Another category of algorithms use the subgradient method to optimize some utility function at each user [3, 4, 5, 6]. They assume a network with fixed links and focus more on the numerical performance of the algorithms than on the actual implementation issues. For multi-user video streaming over ad hoc wireless networks, previous work in [7] proposed a centralized scheme for minimizing the total video distortion under a total rate constraint from the network. While this centralized approach suits well the surveillance scenario with a common receiver, its generalization to multiple sender-receiver pairs would require additional overhead in collecting information for the central decision node.

The goal of this paper is to achieve the same objective using a distributed algorithm. We also use the subgradient method for solving the optimization, but make several modifications to address the characteristics of ad hoc wireless networks. Using a simple pricing mechanism on each link, each relay node only needs to update the link price based on local observations, whereas each source node only needs to adapt the video transmission rate according to the video RD model and the accumulated price along the route.

The rest of the paper is organized as follows. The multi-stream rate allocation problem is presented formally in Section 2. The two major components, source rate adaptation and link price update are explained in details in Sections 3 and 4. In Section 5, we present simulation results for video streaming over an ad hoc wireless network with stationary nodes.

This work is partially supported by NSF Grant CCR-0325639.

2. PROBLEM FORMULATION

Consider N video streams in a network with L links. Each video stream is designated a route from source to destination, containing a set of links $\mathcal{P}_n = \{\ l$: link l belongs to route for stream $n\ \}$. Some of the streams may travel over a common link in the network. The source rate R_n for each video stream needs to be chosen so that the total distortion of all the streams are minimized, without overwhelming any individual link in the network. Given the capacity C_l at each link, the multi-user rate allocation problem can be cast as the following:

$$\min_{R_n} \quad \sum_{n=1}^{N} w_n D_n(R_n) \tag{1}$$

$$\text{s.t.} \quad \sum_{n:l\in\mathcal{P}_n} R_n \leq \alpha C_l,\ \ l = 1, \cdots, L \tag{2}$$

In (1), R_n is the allocated rate and D_n is the corresponding video distortion for stream n; w_n denotes the relative importance of that stream. The constraints in (2) dictate that the total video rate on each link l should be below the capacity C_l by a certain margin, indicated by the scaling factor $\alpha < 1$.

This convex optimization formulation can be solved by a distributed algorithm using the subgradient method [8], which decomposes the optimization process into the iteration between two separate procedures of source rate adaptation and link price update:

$$R_n = \arg\min_{R_n} w_n D_n(R_n) + (\sum_{l:l\in\mathcal{P}_n} \lambda_l) R_n \tag{3}$$

$$\lambda_l = \lambda_l - \mu G_l \tag{4}$$

where λ_l indicates price for link l, μ is a scaling factor for updating λ_l, and the subgradient G_l calculated as the margin in the rate constraint on that link:

$$G_l = \alpha C_l - \sum_{n:l\in\mathcal{P}_n} R_n. \tag{5}$$

Note that the update of the source video rate only depends on the distortion-rate function $D_n(R_n)$ and the accumulated price $\Delta\lambda_n = \sum_{l:l\in\mathcal{P}_n} \lambda_l$ along the route for video stream n; the update of each link price only depends on the observed total video rate and capacity on that link.

Unlike in a wired network, the capacity C_l's in the ad hoc wireless network do not have a fixed value. Instead, they are often affected by other factors in the network, such as interference in transmission or contention for media access opportunities from other nodes. The values of C_l's also depend on the traffic pattern in the network and the average size of the transmitted packets. It is therefore necessary to estimate the C_l's dynamically, as the allocated rate to each stream changes over time. In other words, the optimal rate allocation cannot be achieved in one shot, but rather has to be evolved gradually with link state information updates. The iteration of the two procedures and the exchanges of information between different entities are illustrated in Fig. 1.

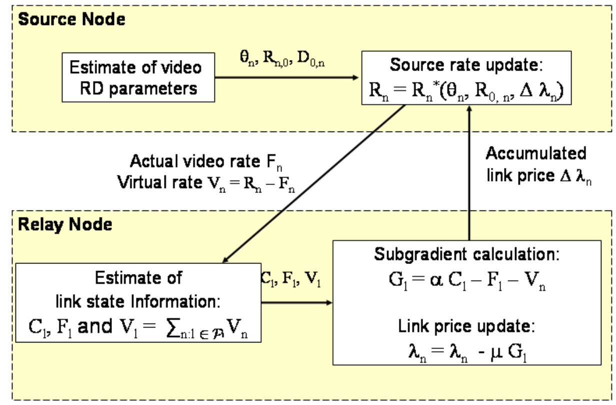

Fig. 1. The rate allocation scheme iterates between link state estimation at each relay node and video rate adaptation at each source node.

3. SOURCE RATE ADAPTATION

The parametric model in [9] is used to represent the rate-distortion tradeoff of each video stream:

$$D_n(R_n) = D_{0,n} + \frac{\theta_n}{R_n - R_{0,n}} \tag{6}$$

The parameters $D_{0,n}$, $R_{0,n}$ and θ_n can be estimated from trial encodings. Given the accumulated link price $\Delta\lambda_n$ over the route, the optimized rate in (4) can now be expressed analytically:

$$R_n = R_{0,n} + \sqrt{\frac{w_n\theta_n}{\Delta\lambda_n}}. \tag{7}$$

Note that for the chosen RD model , both the objective function and its dual are differentiable, therefore the optimization process is equivalent to the gradient method with a fixed update step size. The subgradient scheme also works for non-differentiable RD models such as piecewise linear functions.

Rate adaptation is achieved by switching between different versions of the video stream, pre-encoded with different quantization parameters. Consequently, the available transmission rates are discrete and fall within a certain range. In contrast, the allocated rate R_n is continuous and may vary significantly from iteration to iteration during the initial stage of the optimization. In order to avoid unnecessary fluctuation of the video source rates while still allowing further link updates, we record the difference of allocated rate R_n and the actual video rate as *virtual rate* V_n. It is included in the video packet header and accessed by intermediate nodes for link price update, as explained in Section. 4. Only after the allocated rate stabilizes for some time do we switch the video stream according to the allocated rate. Note that in a live encoding situation, the source rate adjustment can be achieved by a rate control algorithm at the encoder and the field of virtual rate is no longer needed.

4. LINK PRICE UPDATE

When a video packet is transmitted, the relaying node reads the virtual rate V_n from the video packet header, updates its estimation of the capacity and traffic flow rate, calculates the new subgradient G_l, updates its own link price and adds this price to the *accumulated price* in another field of the video packet header. Denoting V_l as the total virtual rate from different streams and F_l as the estimated total actual rate of traffic on link l, the subgradient calculation in (5) can re-written as:

$$G_l = \alpha C_l - F_l - V_l; \tag{8}$$

Intuitively, the subgradient can be interpreted as the difference between requested traffic rate and available capacity on that link, and affects the change in price as in (3).

The capacity and flow rates are estimated from the observed packet arrival and departure times, averaged over many packets. Since the traffic pattern on each link affects the amount of transmission interference and media access contention in the network, the observed capacity changes with the allocated source rate over time as well. As an illustration, Fig. 2 shows the estimated C_l and F_l over time for one link in the experiments of Section 5. Since the average packet sizes with lower video rates are smaller, the inefficiency from overhead packets and headers in the link layer transmissions are more pronounced. Therefore, as the actual rate of traffic increases over time, the estimated capacity also increases slightly, yet *not proportionally*.

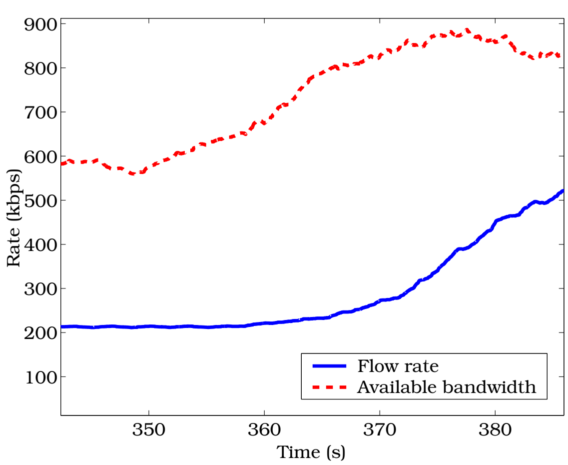

Fig. 2. Estimated capacity and flow rate over a single link in the experiments in Section 5. The estimated capacity increases slightly with increased traffic rate.

5. EXPERIMENTAL RESULTS

Network simulations are performed in ns-2 [10] for video streaming over an ad hoc wireless network with 10 stationary nodes. Two video streaming sessions are set up between two arbitrary node pairs. Routes are manually selected and remain fixed throughout the simulation. Each stream travels over a path of 3 hops; there is one common link for both streams.

	Rate(kbps)		PSNR(dB)
	Proposed	Exhaustive Search	
(A)	763.4	652.9	37.9
(B)[1]	573.7, 252.2	490.0, 213.0	36.7, 40.5
(C)	598.9	518.8	43.9

Table 1. Average allocated rates from the proposed scheme and from exhaustive search, as well as the corresponding video quality in PSNR. Note that the PSNR values are the same for both schemes, since the actual rates in the proposed scheme are the same as those from exhaustive search. The three different scenarios are: (A) first stream (*Foreman*)is active; (B) both are active; (C) second stream (*Mother and Daughter*) is active.

The video sequences *Foreman* and *Mother and Daughter* in CIF format are transmitted over the two routes respectively. They are both encoded with the H.264 codec at different quantization parameters, with GOP length of 15 and the IBBP... structure. The frame rate is 30 Hz. For better efficiency in the link layer transmissions, the smaller B and P frames are first grouped into video packets greater than 1500 bytes. During the actual transmission, all packets exceeding the UDP packet size of 1000 bytes are segmented at the sender and reassembled at the receiver by the transport agent.

The network simulation lasts for 500 seconds. The first video session (*Foreman*) is active for the first 350 seconds and the other (*Mother and Daughter*) is active for the last 300 seconds. Therefore the simulation covers three different scenarios: either one of the streams or both are active in the network. The weights (w_n in (1)) of both streams are set to be equal, and the scaling factor α in (2) is chosen to be 0.9. Each video stream starts at the lowest available transmission rate and gradually ramps up the encoding rate according to the rate allocation results. For comparison, we also exhaustively search the optimal rate or rate pairs, by trial fixed-rate transmissions at all available rates of the video streams for each scenario.

Figure 3 shows the link price on the common link shared by the two streams. Notes that when both streams are active from 200s to 350s, the link price increases significantly so that the source rates can be reduced according to (4).

The allocated rate and result video quality from the proposed scheme and from exhaustive search are listed in Table 1. The actual rates are the same as the ones from exhaustive search, since the allocated rates are rounded down to the closest available rates from encoding. Also note that when both streams are present, more rate is allocated to *Foreman*, as it contains more active contents.

Figure 4 plots the allocated and actual rates for both streams over time. Convergence time for each transition of the scenarios are 9.5, 36.3 and 21.9 seconds respectively. Therefore, the algorithm is mainly suitable for rate allocation in a network with stationary nodes.

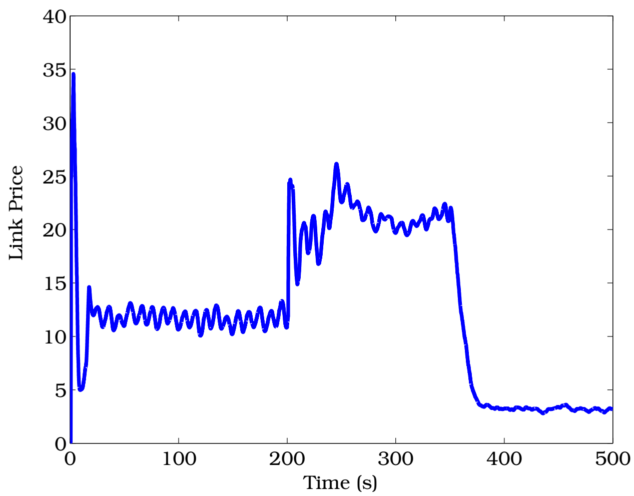

Fig. 3. Price of the common link shared by the two video streams. The link price is higher during the period when both streams are present in the network

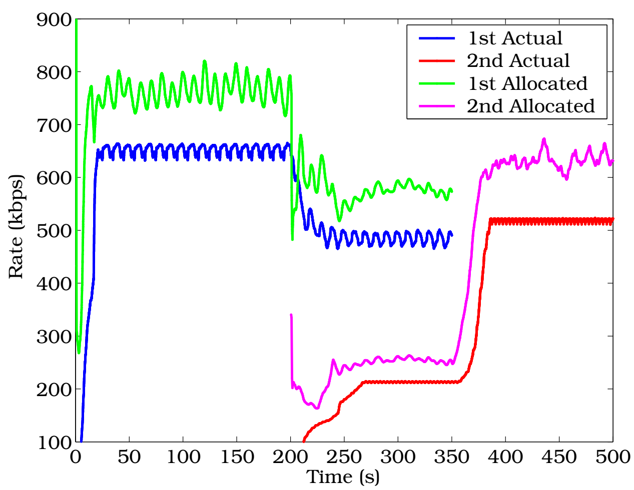

Fig. 4. Allocated and actual rate for each video stream. The three transitions of scenarios occur at 0.1s, 200s and 350 respectively.

6. CONCLUSIONS

We propose a distributed algorithm for rate allocation among multiple video streams over the ad hoc wireless network. Based on the subgradient method, minimization of the total video distortion with rate constraints from each link can be solved by iteratively updating the link prices and video source rates. Simulation results show that in spite of the issues from a realistic implementation, such as discrete set of available transmission rates and varying capacity on each link, the proposed scheme can achieve the same optimal rate allocation as that from exhaustive search. Convergence time of the scheme from the simulation is on the order of tens of seconds, therefore the current implementation is mainly suitable for networks with stationary nodes. It is the interest of future work to improve the convergence performance and to investigate the behavior of the proposed scheme in conjunction with rate control for live coding scenarios.

[1]The first rate is for *Foreman* and the second for *Mother and Daughter*.

7. ACKNOWLEDGMENTS

The authors would like to thank Houda Kamoun at Munich Technical University for programming part of the implementations in Section 4.

8. REFERENCES

[1] M. Allman, V. Paxson, and W. R. Stevens, "TCP Congestion Control," *RFC 2581*, Apr. 1999.

[2] M. Handley, S. Floyd, J. Pahdye, and J. Widmer, "TCP Friendly Rate Control (TFRC): Protocol Specification," *RFC 3448*, Jan. 2003.

[3] F. P. Kelly, "Charging and Rate Control for Elastic Traffic," *European Transactions on Telecommunications*, vol. 8, no. 1, pp. 33–37, 1997.

[4] F. Kelly, A. Maulloo, and D. Tan, "Rate Control for Communication Networks: Shadow Prices, Proportional Fairness and Stability," *Journal of Operations Research Society*, vol. 49, no. 3, pp. 237–252, 1998.

[5] R. J. La and V. Anantharam, "Utility-Based Rate Control in the Internet for Elastic Traffic," *IEEE Transactions on Networking*, vol. 10, no. 2, pp. 272–285, 2002.

[6] K. Kar, S. Sarkar, and L. Tassiulas, "A Simple Rate Control Algorithm for Maximizing Total User Utility," *Proceedings Twentieth Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM-01)*, April 2001.

[7] X. Zhu, E. Setton, and B. Girod, "Rate Allocation For Multi-Camera Surveillance Over An Ad Hoc Wireless Network," *Picture Coding Symposium (PCS-04)*, December 2004.

[8] N. Z. Shor, "Minimization Methods for Non-differentiable Functions," *Springer Series in Computational Mathematics. Springer*, 1985.

[9] K. Stuhlmüller, N. Färber, M. Link, and B. Girod, "Analysis of video transmission over lossy channels," *IEEE Journal on Selected Areas in Communications*, vol. 18, no. 6, pp. 1012–32, June 2000.

[10] "NS-2," *http://www.isi.edu/nsnam/ns/*.