Computer Communications 77 (2016) 1–9

Contents lists available at ScienceDirect

Computer Communications

journal homepage: www.elsevier.com/locate/comcom

Cooperative routing with video complexity for heterogeneous motion-level streaming

Jeongtae Kim*, Soyoung Park, Hyowon Suh

Collaborative Engineering Laboratory, Department of Industrial and Systems Engineering, KAIST, 291 Daehak-ro, Yuseong-Gu, Daejeon 34141, Korea

ARTICLE INFO

Article history:
Received 11 June 2014
Revised 30 January 2015
Accepted 24 July 2015
Available online 30 July 2015

Keywords:
Video complexity
Rate-distortion model
Cooperative relay
Flow routing
Multi-hop relay

ABSTRACT

Cooperative relay is a promising technique that can improve network capacity in multi-hop wireless networks. However, conventional video streaming schemes on cooperative relay networks do not consider the video complexity of each video sequence. In this paper, we develop and experimentally evaluate a video streaming scheme that considers video complexity over a cooperative multi-hop relay network. We first develop a video distortion model taking into account the video complexity of the video sequences. Then, we propose a flow-routing algorithm for heterogeneous motion-level video streams in multi-hop cooperative networks. To evaluate the video routing performance for heterogeneous motion-level video sequences, we conduct experimental simulations with the proposed routing algorithm and a video distortion model. Numerical results show that network performance improves when video sequences are routed while considering heterogeneous motion levels to maximize the minimum peak signal-to-noise ratio (PSNR) value for the video sessions. The simulation results also show that the adoption of cooperative relays and a hop-count limitation can also improve the routing performance of video streams.

1. Introduction

The amount of video traffic over wireless channels has remarkably increased over the past decade as increasingly more applications pushing media traffic traverse wireless networks. However, reliable video streaming on wireless networks is very challenging because wireless channel conditions and mobile rate demands have significant time-varying characteristics. Substantial effort has been dedicated to developing wireless video streaming schemes for the improvement of video streaming performance over wireless networks. In particular, cooperative relay communication is being focused on as a promising technique to improve network capacity in multi-hop wireless networks. Cooperative relay communications has advantages in that it is possible to acquire spatial diversity without requiring multiple antennas equipped within a node.

Many researchers have proposed resource allocation strategies or routing frameworks to obtain increment or optimality in data transmission performance over cooperative wireless networks [1–6]. Ng and Yu [1] proposed a utility maximization framework for relay strategy and resource allocation taking user traffic demand into consideration in wireless cooperative cellular data networks. Liu et al. [2] dealt with cooperative communication techniques assuming that relaying nodes can forward information fully or in part. Awad and Shen [3] studied resource allocation for OFDMA-based two-hop relay networks and proposed a sub-optimal algorithm for an NP-complete problem. Other researchers [4–6] have also investigated relay and resource allocation strategies to improve network performance. However, recent studies have devoted more attention to video streaming and multi-hop relays in cooperative wireless networks [7–10]. Guan et al. [7] proposed a global optimization algorithm based on a branch and bound framework, and on convex relaxation of non-convex constraints to solve the cross-layer design problem for video streaming over cooperative networks. Sharma et al. [8] developed a mathematical model and proposed a solution procedure based on the branch and bound with cutting planes (BB-CP) to explore the behavior of cooperative communication in multi-hop wireless networks. Mastronarde et al. [9] proposed a solution based on cooperative coding that warrants a uniformly better experience to video users, and requires relatively modest changes to the multiple access cross-layer optimization framework. In other studies, Mastronarde et al. [10] investigated the impact of cooperative relaying on uplink multi-user wireless video transmissions. Although these studies show outstanding achievement in video streaming schemes over wireless relay networks, a significant portion of the attention is paid to resource allocation strategies and data transmission schemes, and the characteristics of video sequences are not considered.

In this paper, we develop a video distortion model that takes into account the temporal complexity of video sequences. The model

* Corresponding author. Tel.: +82428601172.
E-mail address: jeongtae@gmail.com (J. Kim).

http://dx.doi.org/10.1016/j.comcom.2015.07.021
0140-3664/

focuses on the relationship between distortion and video complexity, which is an important factor constituting the distortion function. Temporal video complexity can be defined as the average difference between a frame and the successive frames for a video sequence; thus, highly complex video sequences require higher data rates than those with low complexity. The complexity of each video sequence is assumed to be already calculated at the source node or another place. Video distortion is typically a representative factor that has a significant effect on the quality of video streaming services over wireless networks. There are two different types of distortion: loss distortion and source distortion. Loss distortion can be introduced by packet or frame losses caused by transmission packet errors. Source distortion arises from lossy video encoding at the transmission source node. Video sequences are typically highly compressed for volume reduction at the source node using lossy encoding methods, which necessarily causes video distortion. Those relations can be summarized as video complexity may have an effect on video distortion, and finally on the video streaming experience of the user. A number of researchers have investigated video complexity and distortion during video coding and transmission. Feng et al. [11] showed that the trade-off between distortion and delay can be dependent on the complexity of the video. Video distortion and motion complexity have also received attention in various other studies [12–14]. Although video complexity is a noticeable characteristic of video sequences and is highly relevant to video distortion, to the best of our knowledge, there is no study on video streaming schemes over wireless networks for heterogeneous motion-level video streams.

We also provide a flow-routing algorithm for heterogeneous motion-level video streams in multi-hop cooperative networks. The algorithm, which aims to maximize the minimum peak signal-to-noise ratio (PSNR) among multiple concurrent video sessions, consists of three phases: path determination between the source and destination nodes, relay selection, and rate assignment. The proposed algorithm is more appropriate for application to video on demand (VoD) than real-time video streams because the video complexity of video sequences needs to be calculated prior to transmission.

The main contributions of this paper are as follows. We develop a video distortion model that takes into account temporal video complexity, one of the main characteristics of video sequences. We also propose a flow-routing algorithm for heterogeneous motion-level video sequences over multi-hop cooperative networks applying the proposed video distortion model, which considers video complexity. In the proposed algorithm, each video sequence has a different priority in determining the routing path according to the video complexity, pursuing higher user fairness. Finally, we provide experimental results indicating that the video streaming quality, in terms of the minimum PSNR, can be improved by taking into account video complexity. Additional analysis demonstrates the effect of cooperative relay and hop-count limitation.

2. Video distortion model

PSNR is commonly used to measure the video reconstruction performance of lossy compression codecs, and is computed using the ratio between the maximum possible video distortion and the current distortion, as described in (1) [15]:

$$PSNR = 10 \cdot \log\left(\frac{D_{max}}{D}\right) \tag{1}$$

where D_{max} is the maximum distortion possible.

In the following subsections, equations for loss distortion and source distortion are introduced to compute the PSNR value for each video session.

2.1. Loss distortion

Video frames are decodable at the destination node when more than a minimum number of required packets arrive successfully. When a frame is not decodable owing to a high rate of packet errors during transmission, a frame recovery method can be applied to recover the information contained in the non-decodable frame. An error concealment method is used in this paper for frame recovery, in which the recent correctly decoded frame is copied and duplicated into the first lost frame and all its successors in a group of pictures (GOP). Because the recently decoded frames are copied to the subsequent frames, the expected distortion of a GOP depends on the location of the first frame loss. The expected loss distortion of video stream s when the ith frame is the first lost frame in a GOP is computed using (2) [14].

$$D_{i,S} = (G-i)\frac{G \cdot i \cdot D_{min,s} + (G-i-1)\cdot D_{max,s}}{(G-1)\cdot G} \tag{2}$$

The values of D_{min} and D_{max} depend on each video sequence and can be obtained through experimental measurements. A GOP is assumed to be composed of a total of G frames, i.e., one I-frame and $(G-1)$ P-frames. As described above, some of the packets composing a frame can be lost during a video transmission from one node to the next. The transmission success rate of a packet for session s, denoted as $r_{pt,s}$, which indicates the probability that a packet of video sequence s is delivered from the source to destination successfully, can be calculated as

$$r_{pt,s} = (1-PEP)^{h_s} \tag{3}$$

where variable h_s is the number of hops on the streaming path of corresponding session s, and the packet error probability (PEP) is the probability that a packet error occurs during a one-hop transmission. In this paper, for performance calculation simplicity, the PEP is assumed to be constant for all links.

A frame can be decoded successfully at the destination node when more than a minimum number of required packets, including the first packet, successfully arrive at the destination without error. The first packet is necessarily required for decoding because it has the required information to decode the video stream and also information on its successors. The minimum number of packets, s_s, required to decode a frame shows how sensitive the decoder is to packet errors. When the error sensitivity values of the I-frame and P-frame of session s are $s_{I,s}$ and $s_{P,s}$, the probability that an I-frame or a P-frame of a GOP is successfully decoded at the destination can be computed as

$$p_{I,s} = r_{pt,s} \cdot \sum_{i=s_{I,s}}^{n_{I,s}-1} \binom{n_{I,s}-1}{i} r_{pt,s}^{i} \, (1-r_{pt,s})^{n_{I,s}-1-i} \tag{4}$$

$$p_{P,s} = r_{pt,s} \cdot \sum_{i=s_{P,s}}^{n_{P,s}-1} \binom{n_{P,s}-1}{i} r_{pt,s}^{i} \, (1-r_{pt,s})^{n_{P,s}-1-i} \tag{5}$$

where variables $n_{I,s}$ and $n_{P,s}$ are the number of packets composing an I-frame and a P-frame for session s, respectively. The size of an I-frame is typically much larger than that of a P-frame because an I-frame contains all the information required to describe a frame, whereas a P-frame contains only information on the difference between the corresponding and previous frames. The ratio of the number of packets of the I- and P-frames in a GOP in session s, i.e., $n_{I,s}$ divided by $n_{P,s}$, is set as 10. The fraction of P-frames for various video sequences can be found in [16]. The values of $s_{I,s}$ and $s_{P,s}$ vary according to the video complexity [12]. More specifically, the values need to be high with a fast-motion video sequence because the loss of one frame of a fast-motion video causes high distortion compared to a slow-motion video sequence. The $s_{I,s}$ and $s_{P,s}$ are set proportional to the number of packets of each frame with a constant β and variable α, which reflects