Computer Communications 33 (2010) 1380–1390

Contents lists available at ScienceDirect

Computer Communications

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

A distortion-minimizing rate controller for wireless multimedia sensor networks ☆,☆☆

Scott Pudlewski, Tommaso Melodia *

Wireless Networks and Embedded Systems Laboratory, Department of Electrical Engineering, State University of New York (SUNY) at Buffalo, Buffalo, NY, USA

ARTICLE INFO

Article history:
Available online 20 February 2010

Keywords:
Multimedia sensor networks
Cross-layer design
Rate control

ABSTRACT

The availability of inexpensive CMOS cameras and microphones that can ubiquitously capture multimedia content from the environment is fostering the development of wireless multimedia sensor networks (WMSNs), i.e., distributed systems of wirelessly networked devices that can retrieve video and audio streams, still images, and scalar sensor data.

A new cross-layer rate control scheme for WMSNs is introduced in this paper with a twofold objective: (i) maximize the video quality of each individual video stream; (ii) maintain fairness in the domain of video quality between different video streams. The rate control scheme is based on analytical and empirical models of video distortion and consists of a new cross-layer control algorithm that jointly regulates the end-to-end data rate, the video quality, and the strength of the channel coding at the physical layer. The end-to-end data rate is regulated to avoid congestion while maintaining fairness in the domain of *video quality* rather than data rate. Once the end-to-end data rate has been determined, the sender adjusts the video encoder rate and the channel encoder rate based on the overall rate and the current channel quality, with the objective of minimizing the distortion of the received video. Simulations show that the proposed algorithm considerably improves the received video quality with respect to state-of-the art rate control algorithms, without sacrificing on fairness.

1. Introduction

The availability of inexpensive hardware such as CMOS cameras and microphones that can ubiquitously capture multimedia content from the environment has fostered the development of wireless multimedia sensor networks (WMSNs) [2], i.e., distributed systems of wirelessly networked devices deployed to retrieve video and audio streams, still images, and scalar sensor data from the environment. WMSNs will enable new applications such as multimedia surveillance, traffic enforcement and control systems, advanced health care delivery, structural health monitoring, and industrial process control [3]. Many of these applications require the sensor network paradigm to be re-thought in view of the need to deliver multimedia content with predefined levels of quality of service (QoS).

QoS-compliant delivery of multimedia content in sensor networks is a challenging, and largely unexplored task [4,5]. First, embedded sensors are constrained in terms of battery, memory, processing capability, and achievable overall rate [2], while delivery of multimedia flows may be a resource-intensive task. Second, in multi-hop wireless networks the attainable capacity of each wireless link depends on the interference level perceived at the receiver. Hence, capacity and delay attainable at each link are location dependent, vary continuously, and may be bursty in nature, thus making QoS provisioning a challenging task. Lastly, functionalities handled at different layers of the networking protocol stack are inherently and strictly coupled due to the shared nature of the communication channel. Hence, different functionalities aimed at QoS provisioning should not be treated separately when efficient solutions are sought, i.e., a cross-layer design approach is needed [6–10].

In this paper, we consider a multi-hop wireless network of video sensors deployed for surveillance applications and focus on reliable and real-time transport of video traffic. The objective is to design algorithms to *efficiently* and *fairly* share the common network infrastructure among the video streams generated by different video sensors, to deliver high-quality video on resource-constrained devices. To achieve this objective, we propose the distortion-minimizing rate control (DMRC) algorithm, a new decentralized cross-layer control algorithm that jointly regulates the end-to-end data rate, the video

☆ This work was supported by the US Air Force Research Laboratory under Grant FA8750-08-1-0063.

☆☆ A preliminary shorter version of this paper [1] appeared in the Proceedings of the IEEE International Conference on Mobile Ad hoc and Sensor Systems (MASS'2009).

* Corresponding author.
E-mail addresses: smp25@buffalo.edu (S. Pudlewski), tmelodia@eng.buffalo.edu (T. Melodia).

0140-3664/$ - see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.comcom.2010.02.016

quality, and the strength of the channel coding at the physical layer to minimize the *distortion* of the received video. The end-to-end data rate is chosen to avoid congestion while maintaining fairness in the domain of *video quality* (rather than data rate as in traditional rate control algorithms). Once the end-to-end data rate has been determined, the sender calculates the optimal proportion of video encoder rate and channel encoder rate based on the overall rate available and on the current quality of the wireless channel on the source-destination path, with the objective of minimizing the video distortion at the receiver.

Video distortion in wireless networks is mostly caused by lossy source coding, transmission errors originated by channel fading, buffer overflows and playout deadline misses. Intuitively, if the loss happens at a relay node due to congestion, then the video encoder rate should be decreased smoothly to reduce congestion. In case packets are being lost due to correlated fading on the wireless link, the video encoder rate should remain unchanged and the channel encoder rate can be reduced. The channel encoder rate should then be increased as the wireless channel errors decrease. In DMRC, the signal to noise ratio (SNR) and the round trip time (RTT) are used to determine what is causing the distortion at the receiver. By using feedback packets, the receiver updates the sender with the current forward channel information including SNR and RTT. This allows the sender to correctly react to the cause of packet errors.

Unlike other current cross-layer optimization techniques [8,9,11–14], the proposed scheme minimizes the video distortion by finding the optimal ratio of video encoder rate to channel encoder rate. Furthermore, the control algorithm finds the best possible transmission rate for a network sending primarily video. Differently from previously proposed schemes such as TCP-Friendly Rate Control (TFRC) [15], we will not take fairness towards TCP as a key design principle. In a resource-constrained WMSN, priority must be given to the delay-sensitive flows, possibly at the expense of other delay-tolerant data. Furthermore, TCP traffic is unlikely to be simultaneously transmitted in a sensor network. By regulating the end-to-end data rate based on the video compression rate and by jointly optimizing the video coding and channel coding rates, DMRC uses network resources more efficiently than state-of-the-art rate control protocols like the TCP-Friendly Rate Control (TFRC) [15], thus resulting in higher video quality at the receiver.

The remainder of this paper is structured as follows. In Section 2, we discuss previous work on related topics. In Section 3, we introduce the considered system model. In Section 4, we describe our solution to the determination of the video encoder rate and the channel encoder rate, and in Section 5 we extend the analysis to the case of predictive video encoders. In Section 6, we describe the DMRC rate controller in detail. In Section 8 we discuss performance evaluation results, while in Section 9 we draw the main conclusions.

2. Related work

The most common form of rate control is the well-known transmission control protocol (TCP) [16]. Because of the additive increase/multiplicative-decrease algorithm used in TCP, the rate transitions that it determines are not smooth enough for high quality video transfer [17]. In addition, TCP assumes that the main cause of packet loss is network congestion. Although this assumption is reasonable for wired networks, for wireless networks channel errors must be taken into account if an accurate prediction of the network congestion is needed. For example, in [18] it was experimentally shown how in sensor networks packets are frequently dropped because of channel errors even on short-distance links. Because of this, very few links can be considered error-free and the effect of packet drops due to channel loss has a large impact on the video quality.

These considerations have led to a number of equation-based rate control schemes. Equation-based rate control analytically regulates the transmission rate of a node based on measured parameters such as the number of lost packets and the round trip time (RTT) of the data packets. Two examples of this are the TCP-Friendly Rate Control [15], which uses

$$X = \frac{S}{R\sqrt{\frac{2\pi}{3}} + \left(4R\left(3\sqrt{\frac{2\pi}{8}}\cdot\pi\cdot(1+32\cdot\pi^2)\right)\right)}, \tag{1}$$

i.e., the throughput equation of TCP Reno [16], and the Analytical Rate Control (ARC) [19], which uses

$$X = \frac{S}{4\cdot R}\cdot\left(3+\sqrt{25+24\left(\frac{1-\omega}{\pi-\omega}\right)}\right). \tag{2}$$

In (1) and (2), X [bit/s] represents the transmission rate, S [bit] is the packet size, R [s] is the round trip time (RTT), π is the loss event rate, and ω is the probability of being in a lossy state. Both of these schemes attempt to determine a source rate that is fair to any TCP streams that are concurrently being transmitted in the network. However, in a WMSN, priority must be given to the delay-sensitive flows at the expense of other delay-tolerant data. Therefore, both TCP and ARC result in a transmission rate that is more conservative than the optimal rate. For this reason, in an effort to optimize resource utilization in resource-constrained WMSNs, our scheme does not take TCP fairness into account.

Recent work has investigated the effects of packet loss and compression on video quality. In [20], the authors analyze the video distortion over lossy channels of MPEG encoded video with both inter-frame coding and intra-frame coding. A factor β is defined as the percentage of frames that are an intra-frame, or I frame, i.e., a frame which is independently coded. The authors then derive the value β that optimizes distortion at the receiver. Similar to our work, [20] investigates optimal strategies to transmit video with minimal distortion. However, the authors assume that the I frames are received correctly, and that the only loss is caused by the inter-coded frames. We take the idea a step further and assume that any packet can be lost. Also, we jointly optimize the video coding and the channel coding, which will lead to a better overall performance.

Cross layer design techniques to transmit video over wireless networks are also addressed in [21], where the authors minimize the video distortion by optimizing the code division multiple access (CDMA) coding parameters, the video encoder rate, and the channel encoder rate. This paper focuses specifically on CDMA channels and uses the *operational rate-distortion functions* (ORDF) for each scalable layer to determine the distortion. Conversely, we focus on a solution that is independent of the underlying MAC protocol and of the specific video source encoding scheme, and consider a multi-hop network. To address the multiple rates at the source, the originating node will alter the number of bits per pixel used in the video encoder, thereby changing the rate at the expense of the received video distortion. If a more specific transmission technology were to be considered, our approach could be extended to include characteristics of the receiver as in [21]. Finally, our previous work has investigated cross-layer joint routing, scheduling, and channel coding for WMSNs based on the time-hopping impulse radio ultrawide band (UWB) transmission technique [22]. However, transport-layer issues and video quality related metrics were not addressed.