Computer Networks 53 (2009) 2997–3010

Contents lists available at ScienceDirect

Computer Networks

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

An RL-based scheduling algorithm for video traffic in high-rate wireless personal area networks ☆

Shahab Moradi, A. Hamed Mohsenian-Rad, Vincent W.S. Wong *

Department of Electrical and Computer Engineering, The University of British Columbia, Vancouver, Canada

ARTICLE INFO

Article history:
Received 15 March 2008
Received in revised form 21 March 2009
Accepted 21 July 2009
Available online 7 August 2009
Responsible Editor: E. Knightly

Keywords:
Wireless personal area networks
Scheduling
QoS
Ultra-wide band (UWB)
Markov decision process (MDP)
Reinforcement learning

ABSTRACT

The emerging high-rate wireless personal area network (WPAN) technology is capable of supporting high-speed and high-quality real-time multimedia applications. In particular, video streams are deemed to be a dominant traffic type, and require quality of service (QoS) support. However, in the current IEEE 802.15.3 standard for MAC (media access control) of high-rate WPANs, the implementation details of some key issues such as scheduling and QoS provisioning have not been addressed. In this paper, we first propose a Markov decision process (MDP) model for optimal scheduling for video flows in high-rate WPANs. Using this model, we also propose a scheduler that incorporates compact state space representation, function approximation, and reinforcement learning (RL). Simulation results show that our proposed RL scheduler achieves nearly optimal performance and performs better than F-SRPT, EDD + SRPT, and PAP scheduling algorithms in terms of a lower decoding failure rate.

1. Introduction

In recent years, ultra-wide band (UWB) technology has received increasing attention in the wireless world. It provides short-range connectivity, low transmit power levels, and high-data rates. This makes UWB be the physical layer of choice for high-rate wireless personal area networks (WPANs). UWB-enabled WPANs can offer many new applications, such as home entertainment, real-time multimedia streaming, and wireless universal serial bus (USB).

In order to fully exploit UWB technology in high-rate WPANs, upper layers, including the media access control (MAC) layer, must be properly designed for high-rate applications. Video transmission is one such application for high-rate WPANs, which is predicted to constitute a major traffic load. Real-time video flows are delay-sensitive and require quality of service (QoS) guarantee. However, in the IEEE 802.15.3 standard for MAC [1], which is designed for WPANs, details of scheduling and QoS support are left to the developers. Consequently, in this paper, we aim to design an application-aware scheduling algorithm for MAC layer to provide the required QoS for video traffic.

Similar to other real-time traffic, video flow is delay-sensitive and its frames are dropped at the receiver if their delay exceeds the maximum tolerable delay. However, video stream has a few unique characteristics that make QoS support more challenging than other real-time traffic. It has large peak-to-average ratio of the frame sizes and hierarchical structure with dependency among its frames [2].

Recently, various MAC scheduling algorithms have been proposed for high-rate WPANs (e.g., [2–10]). For impulse-based UWB, scheduling problems can be formulated as rate and power allocation problems. These problems can be modeled as a joint optimization problem, so as to minimize the total power consumption [7] or maximize the total system throughput [9]. The concept of exclusion region is also used for such schedulers [10]. On the other hand, with no assumption on the type of physical layer, Mangharam

☆ This paper was presented in part in *IEEE Global Telecommunications Conference (Globecom)*, Washington, DC, December 2007.

* Corresponding author. Tel.: +1 604 827 5135.
E-mail addresses: shahab@ece.ubc.ca (S. Moradi), hamed@ece.ubc.ca (A.H. Mohsenian-Rad), vincentw@ece.ubc.ca (V.W.S. Wong).

1389-1286/$ - see front matter
doi:10.1016/j.comnet.2009.07.012

et al. proposed the fair shortest remaining processing time (F-SRPT) scheduler [5]. SRPT schedules different jobs in the system in the order of their remaining processing time, from the shortest to the longest. F-SRPT is a variation of SRPT that maintains fairness among flows with different data rates. In [6], the earliest due date (EDD) method is used along with SRPT. In [11], Rhee et al. used the application layer information at MAC layer. Each device informs the piconet controller of the maximum size of its *I*, *P*, and *B* frames for video flows. The channel time is then allocated based on these values. Kim and Cho proposed a scheduling algorithm designed for MPEG-4 flows [2]. Each MPEG-4 frame type is scheduled with a pre-assigned priority (PAP) in *I*, *P*, and *B* order. The scheme proposed in [8] focuses on reducing the average waiting time by using an $M/M/c$ queuing model for channel time allocation at the piconet coordinator. It also proposed a command aggregation scheme in order to reduce MAC overhead when sending short command frames. Energy efficiency is considered by the scheduler proposed in [12]. It defines different service categories in order to balance between energy efficiency and QoS. It also uses the application layer information to assign priorities to the buffered frames at the source of a flow. In [13], we proposed a frame-decodability aware (FDA) technique to improve the performance of scheduling of video flows.

Table 1
List of nomenclature.

Notation	Parameter definition
$\mathbf{A}$	Action set in MDP model
$\mathbf{a}$	Action vector in MDP model
a_i	Binary action for flow i
B	Bidirectional frame
D^{max}	maximum frame deadline in units of superframe size η
d_i	Number of full superframes left until the arrival of the next frame
e_i	A binary variable to denote the eligibility of flow i being scheduled
F	Number of video flows
$\mathscr{F}$	Set of video flows
g_i	Offset of the frame in the queue of flow i, with respect to the beginning of the frame's GOP
$\mathbf{h}_k$	The kth prototype in Kanerva coding
I	Intra-coded frame
L_i^{max}	Maximum frame size of flow i
l_i	Channel time required by the frame in the queue of flow i, in milliseconds
$l^{partial}$	Remaining channel time
M_i	I-to-P frame distance of the ith video flow
N_i	I-to-I frame distance of the ith video flow
P	Predictive frame
p	The flow which can only transmit parts of its frame during a scheduling period
$\mathbf{S}$	State space in MDP model
$\mathbf{s}$	State vector in MDP model
$\mathbf{\Omega}$	Set of feature vectors
β	Activation radius in Kanerva coding
δ_i	Binary variable for flow i
η	Size of a superframe
γ	Inter-arrival time
$\lambda(\,)$	Similarity function
ρ^{π}	Policy gain given the use of policy π
$\mathrm{tx}_i(x)$	Channel time that flow i requires to send x bits of data

The objective of our work is to design a systematic scheduling algorithm for video flows in high-rate WPANs. The contributions of in this paper are as follows:

- We provide a mathematical framework based on Markov decision process (MDP) to determine the optimal scheduler for video flows in high-rate WPANs. This framework takes into account the number and pattern of video flows, and their hierarchical structure.
- Using the MDP framework, along with compact state space representation, function approximation, and reinforcement learning (RL), we propose a practical scheduling algorithm which can provide significantly better QoS to video flows when compared to some other schedulers [14].
- Simulation results show that our proposed RL scheduler achieves nearly optimal performance and performs 53%, 42%, and 49% better than F-SRPT, EDD + SRPT, and PAP schedulers, respectively [15].

The rest of this paper is organized as follows. In Section 2, we provide an overview of the IEEE 802.15.3 standard for MAC, the hierarchical structure of video streams, related scheduling algorithms, the performance metrics for video flows, and reinforcement learning. In Section 3, we describe the mathematical formulation for the optimal video scheduler. The implementation mechanisms of the RL algorithm are described in Section 4. Section 5 describes an extension of our proposed model. Results for performance evaluations and comparisons are presented in Section 6. Conclusions and future work are discussed in Section 7. A list of nomenclature is in Table 1.

2. Background

2.1. IEEE 802.15.3 MAC standard

The IEEE 802.15.3 is designed for WPANs and aims to provide low cost, low power consumption, and high-data rate communications, within its area of operation called the piconet [1]. A piconet consists of a number of independent devices (DEVs) that communicate with each other under the control of the piconet coordinator (PNC). The PNC provides basic timing, performs scheduling, and manages the QoS requirements of the piconet.

Within a piconet, the timing is based on superframes, which consists of three parts (see Fig. 1). The first part, *beacon*, announces timing allocations, superframe duration, and other piconet synchronization parameters. The second part, the contention access period (CAP), is used to communicate commands and asynchronous data. The third part, channel time allocation period (CTAP), is composed of channel time allocations (CTAs) and management CTAs (MCTAs). A DEV can use its CTA, which is assigned to it by the PNC, for isochronous streams, asynchronous data transfer, or sending commands. Channel access during CTAP is TDMA and there is no contention in this period. By sending beacon at the beginning of superframe, the PNC announces the start time and duration of each CTA as well as the DEVs that are allowed to use it.