

Cross-Layer Performance of a Distributed Real-Time MAC Protocol Supporting Variable Bit Rate Multiclass Services in WPANs

David Tung Chong Wong¹, Jon W. Mark², and Kee Chaing Chua³

¹ Institute for Infocomm Research, 21 Heng Mui Keng Terrace, Singapore 119613, Singapore
wongtc@i2r.a-star.edu.sg

² Center for Wireless Communications, University of Waterloo,
Waterloo, Ontario, Canada N2L 3G1
jwmark@bbcr.uwaterloo.ca

³ Electrical and Computer Engineering, National University of Singapore,
10 Kent Ridge Crescent, Singapore 119260, Singapore
chuakc@nus.edu.sg

Abstract. A cross-layer optimization problem to maximize utilization for a distributed real-time medium access control (MAC) protocol supporting variable bit rate (VBR) multiclass services is formulated. A complete sharing (CS) scheme is used as the admission control policy at the connection level to relate the maximum number of devices that can be admitted in the wireless personal area network (WPAN) to the grade of service (GoS) of blocking probability at the connection level, the quality of service (QoS) of packet loss probability at the packet level and the effective data transmission slots efficiency at the MAC layer. With this cross-layer analytical framework, the maximum number of devices that can be admitted into the system to achieve maximum utilization while maintaining prescribed GoS/QoS requirements under different total device mean connection arrival rate can be determined. Numerical results are presented to demonstrate the effectiveness of the proposed cross-layer coupling strategy.

1 Introduction

Cross-layer design, a hot research area [1], aims to optimize performance across different layers of the layered-communications model. By considering cross-layer design, the design system performance can be optimized. In this paper, we consider a basic distributed real-time medium access control (MAC) protocol like the multiband OFDM Alliance (MBOA) MAC [2] without the non-real-time contention MAC protocol for Ultra-Wideband (UWB) systems. A wireless mobile multimedia UWB network has to provide a reasonable user-transparent grade of service (GoS)/quality of service (QoS) for different service classes. To our knowledge, there is no cross-layer performance analysis among utilization, maximum number of devices, blocking probability, packet loss probability, effective data transmission slots efficiency and total device arrival rate in such a distributed real-time MAC protocol to date. The blocking probability is at the connection level in the network layer, while the packet loss probability is at the packet level in the network layer. The effective data transmission slots efficiency is at the MAC sublayer in the link layer. The goal is to find the maximum number of devices that can be supported such that utilization is maximized under reason-

able GoS/QoS specifications. The main contribution of this paper is the analytical formulation and evaluation of the cross-layer optimization problem for the support of variable bit rate multiclass services. The cross-layer optimization here is shown to have some system utilization improvement exceeding 100% compared with a simple admission scheme.

2 Distributed Real-Time MAC Protocol

The distributed real-time MAC protocol, e.g., [2], uses a frame format, as shown in Fig. 1. Each frame consists of two parts: a beacon subframe and a data transmission subframe. The former consists of beacon slots for existing devices in the system, which are packed at the beginning of the beacon period subframe, while N_{BC} beacon slots are available for new devices to get into the system through contention using, e.g., the slotted ALOHA protocol. After successful entry into the system, the new devices are packed together at the beginning of the beacon period. The number of contention beacon slots is assumed constant. When devices leave the system, a packing protocol is used to pack the remaining devices' beacon slots together at the beginning of the beacon slots period. Thus the beacon period is not fixed but varies according to the number of devices in the system. The beacon period subframe is assumed to be an integer number of packet slots.

The data transmission period is used to transmit data packets whose data reservations are announced in its device beacon slot. This is called the data reservation protocol (DRP) [2]. Transmission need not be in the same order as the devices in the beacon slots and the DRP packets for each device also need not be transmitted immediately after other DRP packets. The number of transmission data packets for each device is not fixed but can vary. Note that all devices announce their data reservations and each device beacon slot contains information on all other devices [4]. Since the beacon period varies, the available number of data packet slots, C , also varies, depending on the number of devices in the system and the number of contention beacon slots:

$$C = C(n_1, n_2, \dots, n_K) = \left\lceil N_{SF} - \left\lceil \left(\sum_{k=1}^K n_k + N_{BC} \right) / N_{BS} \right\rceil \right\rceil, \sum_{k=1}^K n_k \leq N_D, \quad (1)$$

where K is the number of traffic classes, N_{SF} is the number of packet slots in a super-frame, n_k is the number of class k devices in the system and N_{BS} is the number of beacon slots equivalent to one packet slot. $\lceil x \rceil$ is the nearest upper integer value of x .

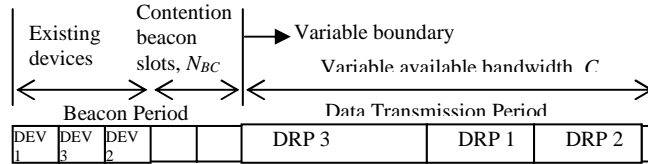


Fig. 1. Frame format of a distributed real-time MAC protocol

3 Analytical Model

At the connection level, the GoS is the blocking probability. In general, this GoS decreases with increase in the maximum number of devices, N_D . At the packet level, the QoS is the packet loss probability. In general, this QoS is zero as long as the maximum total packet transmissions do not exceed the available number of data transmissions in a frame. However, when it starts to increase with the increase in the maximum number of devices, N_D , its initial increase can be at a very sharp rate. This is the limiting constraint in the numerical example in Section 4. At the MAC layer, the QoS is the effective data slots transmission efficiency. It decreases with the increase in the maximum number of devices, N_D . The cross-layer optimization here is shown to have some system utilization improvement exceeding 100% compared with a simple admission criterion in Section 4.

The complete sharing (CS) scheme is used as the resource admission policy. Fig. 2 shows the CS admission Markov chain for 2 classes of devices. λ_k is the class k device mean connection arrival rate, while μ_k is the class k device mean departure rate. $1/\mu_k$ is the class k mean connection holding time. These parameters determine the blocking probability GoS.

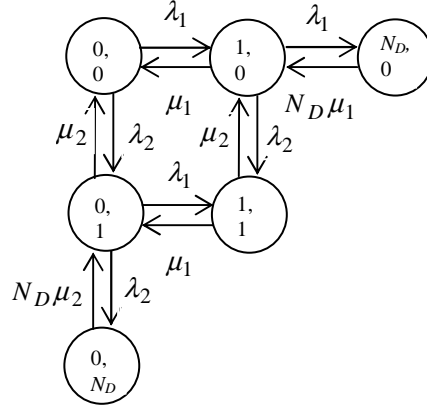


Fig. 2. Complete sharing resource allocation scheme for two classes

The steady state probability of a complete sharing resource allocation scheme for K classes is given by

$$P(n_1, n_2, \dots, n_K) = \frac{(\lambda_1/\mu_1)^{n_1} (\lambda_2/\mu_2)^{n_2} \dots (\lambda_K/\mu_K)^{n_K}}{\sum_{n_1=0}^{N_1} \sum_{n_2=0}^{N_2} \dots \sum_{n_K=0}^{N_K} \frac{(\lambda_1/\mu_1)^{n_1} (\lambda_2/\mu_2)^{n_2} \dots (\lambda_K/\mu_K)^{n_K}}{n_1! n_2! \dots n_K!}}, \quad (2)$$

where $N_k = N_D - \sum_{i=1}^{k-1} n_i$, $k = 1, 2, \dots, K$. The N_k 's together with the summations determine the state space of the K -dimensional Markov chain. The blocking probability, P_B , is given by

$$P_B = \sum_{n_1=0}^{N_1} \sum_{n_2=0}^{N_2} \dots \sum_{n_K=0}^{N_K} P(n_1, n_2, \dots, n_K), \text{ if } \sum_{k=1}^K n_k = N_D. \quad (3)$$

A class k variable bit rate source can be modeled by a continuous-time Markov chain with finite states [3] in Fig. 3.

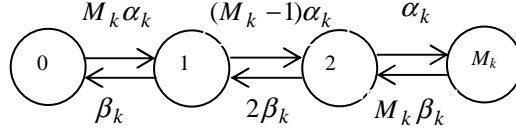


Fig. 3. Continuous-time Markov chain for a single variable bit rate source

Each state represents the discrete level of bit rate generated by a single source. State 1 requires r_k number of packet slots for transmission. The highest state is state M_k . This state requires $M_k r_k$ number of packet slots for transmission. Thus the packet slots variations are in the set of $\{0, r_k, 2r_k, 3r_k, \dots, M_k r_k\}$. α_k is the increase rate of one two-state mini-source, while β_k is the decrease rate of one two-state mini-source. The steady-state probability of being in state m_k , denoted by P_{m_k} , is given by

$$P_{m_k} = \binom{M_k}{m_k} (p_k)^{m_k} (1-p_k)^{M_k-m_k}, \quad m_k = 0, 1, 2, \dots, M_k, \quad (4)$$

where $p_k = \alpha_k / (\alpha_k + \beta_k)$. The probability that l_k levels of bit rate for class k traffic given that there are n_k sources, denoted by $\Pr[l_k/n_k]$, is given by

$$\Pr[l_k | n_k] = \binom{n_k M_k}{l_k} (p_k)^{l_k} (1-p_k)^{n_k M_k - l_k}, \quad l_k = 0, 1, 2, \dots, n_k M_k, \quad (5)$$

Assuming real-time traffic with no storing of packets for retransmission, the packet loss probability, P_L , is given by

$$P_L = \frac{\left[\sum_{n_1=0}^{N_1} \sum_{n_2=0}^{N_2} \dots \sum_{n_K=0}^{N_K} P(n_1, n_2, \dots, n_K) \sum_{l_1=0}^{n_1 M_1} \sum_{l_2=0}^{n_2 M_2} \dots \sum_{l_K=0}^{n_K M_K} \Pr(l_1 | n_1) \times \Pr(l_2 | n_2) \dots \Pr(l_K | n_K) \left[\sum_{k=0}^K l_k r_k - C(n_1, n_2, \dots, n_K) \right]^+ \right]}{\left[\sum_{n_1=0}^{N_1} \sum_{n_2=0}^{N_2} \dots \sum_{n_K=0}^{N_K} P(n_1, n_2, \dots, n_K) \sum_{l_1=0}^{n_1 M_1} \sum_{l_2=0}^{n_2 M_2} \dots \sum_{l_K=0}^{n_K M_K} \Pr(l_1 | n_1) \times \Pr(l_2 | n_2) \dots \Pr(l_K | n_K) \left[\sum_{k=0}^K l_k r_k \right] \right]}, \quad (6)$$

where $[x]^+ = \max[0, x]$. The utilization, N_u , is given by

$$N_u = \sum_{n_1=0}^{N_1} \sum_{n_2=0}^{N_2} \dots \sum_{n_K=0}^{N_K} P(n_1, n_2, \dots, n_K) \sum_{l_1=0}^{n_1 M_1} \sum_{l_2=0}^{n_2 M_2} \dots \sum_{l_K=0}^{n_K M_K} \Pr(l_1 | n_1) \times \Pr(l_2 | n_2) \dots \Pr(l_K | n_K) U, \quad (7)$$

where

$$U = \begin{cases} \sum_{k=1}^K l_k r_k, & \text{if } \sum_{k=1}^K l_k r_k \leq C(n_1, n_2, \dots, n_K) \\ C(n_1, n_2, \dots, n_K), & \text{if } \sum_{k=1}^K l_k r_k > C(n_1, n_2, \dots, n_K) \end{cases}. \quad (8)$$

The probability of n devices in the system, P_n , is given by

$$P_n = \Pr\left[\sum_{k=1}^K n_k = n\right] = \sum_{n_1=0}^{N_1} \sum_{n_2=0}^{N_2} \dots \sum_{n_K=0}^{N_K} P(n_1, n_2, \dots, n_K), \text{ if } \sum_{k=1}^K n_k = n, n = 0, 1, \dots, N_D. \quad (9)$$

Thus the effective available capacity, C_e , is given by

$$C_e = \sum_{n=0}^{N_D} P_n C(n_1, n_2, \dots, n_K), \sum_{k=1}^K n_k = n. \quad (10)$$

The data slots transmission efficiency, ϕ , is given by

$$\phi = C(n_1, n_2, \dots, n_K) / N_{SF}, 0 < \phi < 1. \quad (11)$$

Similarly, the effective data slots transmission efficiency, ϕ_e , is given by

$$\phi_e = C_e / N_{SF}, 0 < \phi_e < 1. \quad (12)$$

The system utilization can be maximized by solving the following constraint optimization problem of maximizing N_u subject to the constraints of $P_B \leq P_B^*$, $P_L \leq P_L^*$, and $\phi_e \geq \phi_e^*$, where the superscript $*$ denotes the requirement values of the corresponding parameters. The results here can be extended for the blocking and loss probabilities of each traffic class.

4 Numerical Results

In this section, we illustrate the system performance by presenting results for a two-class traffic example. The parameter values used in the numerical examples are tabulated in Table 1.

Table 1. Parameter Values Used

Symbol	Value	Symbol	Value
N_{SF}	64	$1/\mu_1$	1 minute
N_{BS}	3	$1/\mu_2$	2 minutes
N_{BC}	3	α_1	0.352
M_1	1	β_1	0.650
M_2	2	α_2	0.9
r_1	1	β_2	0.1
r_2	2	P_B^*	10^{-3}
λ_1	λ_2	P_L^*	10^{-3}
λ	$\lambda_1 + \lambda_2$	ϕ_e^*	0.8

Due to lack of space, graphical results are not shown here but can be obtained using the analysis in section 3. From these graphical results, the data transmission efficiency, ϕ , decreases as the maximum number of devices increases. It does not depend on the total device mean connection arrival rate but only on the total number of de-

vices in the system. If we choose the data transmission efficiency requirement, ϕ^* , to be greater than 0.8, then the maximum number of devices, N_D , that can be supported is only 33. The admission criterion is simply limited to no more than 33 devices. The utilization $N_u = \{22.3, 23.3, 22.6, 21.8, 21.0\}$ when the total device mean connection arrival rate, $\lambda (= \lambda_1 + \lambda_2)$, is $\{10, 20, 30, 40, 50\}$ arrivals per minute.

Similarly from the graphical results, the effective data transmission efficiency, ϕ_e , decreases at a slower rate compared to the data transmission efficiency, ϕ , as the maximum number of devices increases. However, it decreases as the total device mean connection arrival rate increases. Solving the cross-layer optimization in Section 3, we have the maximum utilization, $N_u = \{36.1, 46.1, 46.8, 46.2, 45.3\}$ packet transmission slots at the maximum number of devices, $N_D = \{65, 63, 63, 63, 63\}$ for the total device mean connection arrival rate, $\lambda = \{10, 20, 30, 40, 50\}$ arrivals per minute. The improvement in utilization are respectively $\{62\%, 98\%, 107\%, 112\%, 116\%\}$. Thus this optimized solution results in much higher utilization compared to the case where the data transmission efficiency requirement, ϕ^* , is chosen to be greater than 0.8. The limiting constraint in this numerical example is caused by the packet loss probability requirement, P_L^* , at 10^{-3} . Note that the maximum number of devices, N_D , to achieve maximum utilization, N_u , is quite insensitive to the total device mean connection arrival rate. Thus, the maximum number of devices, N_D , can be set at 63, for example, where the utilization is maximized for most of the total device mean connection arrival rate under consideration. This is a simple admission criterion. In practical systems, measured average device arrival rates are needed.

5 Concluding Remarks

A cross-layer optimization problem has been formulated to maximize utilization in a distributed real-time MAC protocol for WPANs. The GoS/QoS performance metrics in the connection level in the network layer, the packet level in the network layer and the MAC layer have been coupled together to optimize system performance. Numerical results show that this cross-layer optimization approach results in much higher utilization (62% to 116%) than the approach that simply considering the data transmission efficiency requirement. The analysis here can be used to determine the optimal maximum number of devices that can be admitted into the system such that the utilization is maximized under different total device mean connection arrival rate in a distributed real-time MAC protocol for WPANs.

References

1. Wijting, C., Prasad, R.: A Generic Framework for Cross-Layer Optimisation in Wireless Personal Area Networks. *Wireless Personal Communications*, Vol 29, (2004) 135-149
2. O'Connor, J., Brown, R.: MBOA Technical Specification: Distributed Medium Access Control (MAC) for Wireless Networks. MBOA Draft MAC standard version 0.95, (11 April 2005)
3. Maglaris, B., Anastassiou, D., Sen, P., Karlsson, G., Robbins, J.D.: Performance Models of Statistical Multiplexing in Packet Video Communications. *IEEE Transactions on Communications*, Vol. 36, No. 7, (July 1988) 834-844