Admission Control and QoS Provisioning in Multi-service MDA for IEEE 802.11s-Based Wireless Mesh Networks

Jihene Rezgui, Abdelhakim Hafid, and Michel Gendreau

Network Research Laboratory
University of Montreal, Canada
{rezguiji,ahafid}@iro.umontreal.ca,
{michel.gendreau}@cirrelt.ca

Abstract. In this paper, we design an efficient scheduler for Mesh Deterministic Access (MDA) in IEEE 802.11s-based Wireless Mesh Networks (WMNs), called Enhanced Differentiated MDA (ED-MDA). ED-MDA is combined with an Efficient Admission Control algorithm, called EAC, that guarantees QoS for different service classes and provides higher capacity in WMNs. EAC processes both new and HandOver (HO) calls to maintain a balance between two conflicting requirements: maximizing the resource utilization and minimizing the dropping rate. To establish priority between new and HO calls, especially to avoid the forced termination of an ongoing call, ED-MDA coupled with EAC reserves the minimum amount of necessary resources while maintaining an acceptable HO call dropping rate and high resource utilization. Particularly, our model provides an efficient adaptive adjustment of the Contention Free Period (CFP) duration to make efficient use of the scarce wireless resources while supporting different services with different QoS requirements, such as delay. Simulations show that ED-MDA together with EAC outperforms existing schemes.

Keywords: Wireless mesh networks, muti-services, admission control, medium access control and quality of service.

1 Introduction

Wireless Mesh Networks (WMNs) have recently emerged as a promising technology for the next-generation wireless networks [1]. A WMN consists of two types of nodes: Mesh Clients (MCs) and Mesh Routers (MRs). The MRs form a wireless mesh backbone infrastructure that forwards most of the traffic between MCs and Internet gateways. In general, MRs have very little mobility and operate just like stationary routers, except for being connected by wireless links using the very popular IEEE 802.11 WLAN standard or other wireless technologies, such as WiMax. Using more than one radio interface in each MR allows better channel diversity resulting in less interference and therefore more throughput and capacity. However, this only improves the performance for best effort traffic [2] since supporting Quality of Service (QoS) for real-time traffic in WMNs remains an open challenge. For example, real-time connections (RTC)

X. Zhang and D. Qiao (Eds.): QShine 2010, LNICST 74, pp. 204–221, 2011.

require low delays and low packet losses, which are difficult to guarantee because of the random contention used by the traditional CSMA/CA MAC in IEEE 802.11-based WMNs. Nevertheless, some studies solve this problem by enforcing contention-based MAC with complex admission control schemes [3].

In this paper, we are concerned with QoS provisioning as well as the mobility management in multi-service IEEE 802.11s WMNs. Thus, the allocation and efficient use of limited wireless resources are to be studied. The challenge is, to support multimedia applications in order to meet users's expectations in terms of delay and packet losses while maintaining a reasonable high use of radio resources. Generally, WMNs consist of MRs whose small transmission range may result in frequent handoffs of mobile MCs. This inherent property of the mobile MC may often lead to high packet delays and loss rates [4]. Thus, a suitable admission control protocol must be designed to avoid the forced termination of an ongoing call so that a mobile MC could be able to freely move across the WMN while maintaining its ongoing communications uninterrupted. In multiple radio WMNs, it has been proved that there are still channels that may interfere even if the highest quality channel assignment heuristics [5] are adopted. The concept of clique [3] is used in this paper to identify the conflicting links that need to be carefully managed to limit contentions. Recall that a clique is defined as a subset of logical links that interfere with each other.

Mesh Deterministic Access (MDA) aims to provide stringent medium control access (MAC) delay guarantees for real-time services such as voice over IP (VoIP) which is a condition that can hardly be satisfied in classical IEEE 802.11 standard. The MDA scheme studied in [6-9] extends the typical IEEE 802.11 medium instantaneous reservation procedure, also known as the virtual carrier sensing (V-CS), to a more advanced reservation procedure using scheduled MDA OPportunities (MDAOPs) within two-hop neighborhood. MDAOPs are first negotiated between neighboring mesh nodes by exchanging broadcast setup messages, then, MDAOPs reservations are performed in multiples of time-slot unit, during the Delivery Traffic Indication Message (DTIM) periodic interval. To limit the message broadcast signaling overhead, MDA-related messages are sent only within two-hop neighborhood. IEEE 802.11 V-CS [6] is performed by a four way handshake procedure in which request-to-send (RTS), clear-to-send (CTS), data and acknowledgment packets are exchanged between two communicating nodes, while a network-allocation vector (NAV) is set by the other nodes in the physical sensing range (P-CS). V-CS works flawlessly in single hop wireless networks but can cause severe interference in wireless links that are multiple hops away and share the same channel within an overlapped transmission or interference range. Therefore, in multi-hop wireless networks, such as WMNs, MDA, with less multi-hop interference, was adopted in the IEEE 802.11s draft amendment [6] to provide stringent delay bounds. Although, this adopted scheme of MDA has reduced, to a certain extent, the delay bounds it lacks the concept of differentiating frames with different priorities. Basically, MDA is supposed to provide a channel access with equal probabilities for all stations contending for the channel deterministic access in a distributed manner. However, equal access probabilities are not desirable among stations with different priority frames.

Hiertz et al. [6] studied the MAC in IEEE 802.11s by first describing the mandatory Enhanced Distributed Channel Access (EDCA) and then the optional feature MDA. In opposition to EDCA, MDA is particularly designed to support multi-hop

WMNs and is based on scheduled medium access. In addition, we note that QoS support in IEEE 802.11s is not sufficient to provide lower delays and losses required by multimedia applications. This being said, there exists a number of contributions [10-12] that propose techniques in the context of IEEE 802.11 to improve voice QoS and network capacity. However, these contributions neither take into account real and no real time traffic at the same time nor differentiate between new and HO calls.

We note that interferences outside the two-hop neighborhood can still occur with MDA; this may degrade WMN performance. The contributions in [8-9] address the problem of interferences outside the two-hop neighborhood in MDA. Cicconetti et al. [8] use dynamic relocation (called, MDA-DR) of conflicting MDA time-slots that are two hops away from each requesting node. However, MDA-DR limits its performance improvement to throughput which makes it suitable only for data applications such as file transfer and Web browsing and not for delay sensitive applications such as voice. Furthermore, it lacks an admission control scheme that limits packet losses. In [9], Medium Access through REservation (MARE) is proposed as a generalization of the RTS/CTS V-CS procedure to multi-hop wireless networks. In this scheme, rather than using the excessive beaconing broadcast of each node in MDA, MARE uses a more elaborated RTS/CTS-like scheme that is able to consistently notify the appropriate group of neighboring nodes to which a given set of time-slots has been reserved. However, no scheduling and admission control was developed for MARE which limits its efficiency as well as its QoS guarantees for voice traffic. It is worth noting that several research studies focused on defining admission control mechanisms in wireless networks to differentiate between HO and new calls [13] or between different types of traffic with different QoS requirements [14] and [15]. However, current MDA or improved MDA schemes ([8] and [9]) do not support differentiation between different classes of services and also between HO and new calls.

Our contributions, in this paper, can be summarized as follows: (1) We propose a novel scheduler algorithm for per-flow MDA, called ED-MDA, which takes into account traffic differentiation; (2) We elaborate an adaptive adjustment of CFP (Contention Free Period) for each service class by prioritizing RTC in order to provide QoS guarantees in terms of delay and blocking probability that does not exceed predefined thresholds; (3) We propose a new admission control algorithm, called EAC, in conjunction with ED-MDA. EAC reserves the minimum amount of necessary resources (i.e., time-slots) to maintain an acceptable HO call dropping rate and provide high resource utilization; it also establishes priority between new and HO calls to avoid interrupted communication for highly mobile users; and (4) We present a centralized static algorithm (Alg. I) which searches for an initial feasible Upper Bound of the number of Time-Slots for each service class based on Erlang-b formula [16], called UBTS. Thus, starting with this good initial solution rather than a randomly generated one, ED-MDA coupled with EAC provides a better final adjustment of CFP/time-slots while improving the initial solution according to the state of the blocking probability constraint (Section 4).

The remainder of the paper is organized as follows. Section 2 presents our notations, assumptions, and network model. Section 3 proposes an analytical formulation of the differentiated MDA (ED-MDA). Section 4 describes our admission control algorithm EAC. Section 5 evaluates the proposed solution via simulations. Finally, Section 6 concludes this paper.

2 Overview of ED-MDA

In this section, we present a network model for WMNs. In particular, we first define key concepts, namely, transmission and interferences cliques, and then present the assumptions/notations used in the rest of this paper. We also present an overview of the proposed scheme ED-MDA.

2. 1 Assumptions and Notations

We consider a multi-hop WMN as illustrated in Fig. 1. MRs (e.g., MR_3 and MR_{17}) aggregate and forward traffic from/to MCs. These MRs communicate with each other to form a multi-hop wireless backbone network. This backbone network forwards traffic to/from the gateway access points (e.g., P_1 and P_2) so that a packet traverses multiple hops (MRs) to reach a destination. We model the backbone of a WMN as an undirected graph called the connectivity graph G = (V, E), where V represents the set of mesh nodes, and E represents the set of edges between nodes (connections). Among these nodes, we notice the subset P such that, $P \subset V$ to be the gateway access points that connect clients to the Internet. In the rest of this paper, MRs and gateway access points are collectively referred to as mesh nodes. Finally, $\forall$ (u, v) $\in$ V, an edge e = (u, v) $\in$ E if the distance between u and v, which is denoted by d(u, v), is smaller than the transmission range (i.e., d(u, v) $\leq$ r), where r represents the radio transmission range of nodes u and v.

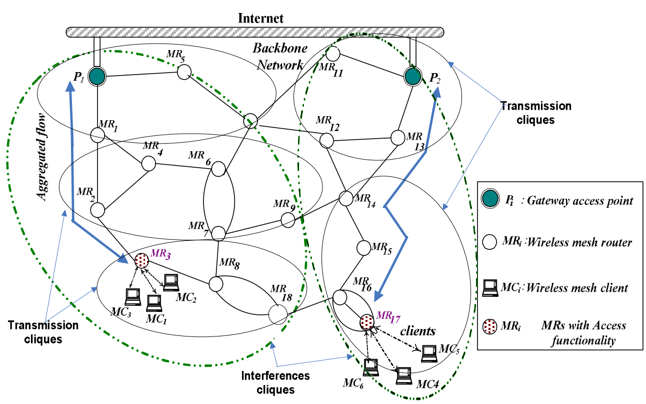


Fig. 1. A typical WMN where transmission cliques vs. interferences cliques are shown, topology of mesh routers mutually interfering within cliques

Generally, we can define in any topological network two types of cliques (see Fig. 1): A transmission clique and an interference clique. Transmission cliques are defined as sets of nodes, sharing the same channel and having a pairwise distance smaller than or equal to the radio transmission range of nodes in the pair (i.e., $d(u, v) \leq r$). Interference cliques are defined as sets of nodes that use the same channel and have a pairwise distance in the interval $]r, R]$, where R is the interference range. Interference cliques are used to identify nodes in the Carrier Sense Range (CSR). These interference cliques are briefly called cliques in the rest of the paper. Fig. 1 shows a sample MR-WMN that consists of two interference cliques and a number of transmission cliques inside each of the interference clique. Routers MR_{14}, MR_{15}, MR_{16} and MR_{17} compose the transmission clique where all routers links are in transmission range of each other. The logical topology is usually built using a Dijkstra shortest path based routing algorithm with the number of hops as the routing metric which is suitable for delay-sensitive RTC.

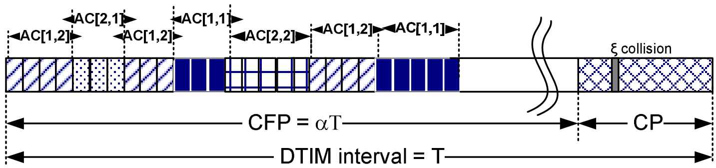
We note that the Erlang-b formula [16] describes the probability of call lost/blocked in a network (all time-slots are busy) until a time-slot becomes free. In Table 1, we assume that $i \in [1..4], j \in [1..2]$ where i represents the service class and j is equal to 1 if it is an HO call, 2 otherwise (i.e., new call).

Table 1. Notations

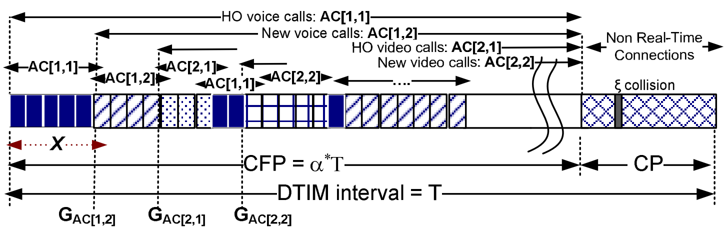
$AIFS_{AC[i.j]}$	Arbitrary Inter Frame Space of radio for service class $AC[i.j]$
τ	Time-slot duration
$L_{AC[i.j]}$.	Packet size (including PHY and above) for service class $AC[i.j]$
O	MDAOP offset for a flow in a clique
π	MDAOP periodicity for a flow in a clique
δ	MDAOP duration for a flow in a clique
σ	DTIM utilization at time t
α	Fraction of the DTIM period reserved for MDA
T	DTIM period
C	IEEE 802.11 transmission rate in a clique
$T_{AC[i.j]}$	Upper bound of the number of time-slots for service class $AC[i.j]$
CFP	Contention Free Period to serve *RTC*
CP	Contention period to serve no RTC (*NRTC*)
E	Total offered load in Erlang-b
$B_{AC[i.1]}$	Dropping probability threshold of *HO* voice (i=1)/video (i=2) calls
$B_{AC[i.2]}$	Blocking probability threshold of *new* voice (i=1)/video (i=2) calls
$A_{AC[i.j]}$	Total number of *HO* (j=1) *or New* (j=2) [voice (i=1)/video (i=2)] calls admitted
$R_{AC[i.j]}$	Total number of *HO* (j=1) *or New* (j=2) [voice (i=1)/video (i=2)] calls rejected
$D_{AC[i.j]}$	The required delay tolerance in a clique for service class $AC[i.j]$
$D^{\max}_{AC[i.j]}$	The maximal delay for maximum number of hops m in a path for service class $AC[i.j]$

2.2 MDA without Traffic Differentiation

In MDA, the time between consecutive DTIM beacon frames is divided into time-slots of length $32\mu s$. The periodic broadcast of beacon frames to all radios in a clique allows the synchronization of these DTIM intervals. Initially, MRs reserve the wireless medium for MDAOPs, which are reserved as multiples of time-slots during a given Contention Free Period (CFP) of a maximum access fraction (MAF=αT) of the DTIM interval T (see Fig.2-(a)). The remaining part of the interval as illustrated in Fig. 2-(a) is the contention period (CP) used for throughput-sensitive data applications. Current MDA does not support traffic differentiation and has same behavior for all service classes in the network.



(a) Per-flow MDAOP schedules for MR_3 and MR_{17} RTC illustrated in Fig.1 over **MDA**



(b) Per-flow MDAOP schedules for MR_3 and MR_{17} RTC illustrated in Fig.1 over **ED-MDA** where X is the reserved guard to AC [1,1]

Fig. 2. (a) MDA per-flow schedule; and (b) ED-MDA per-flow schedule

2.3 MDA with Traffic Differentiation: ED-MDA

In this Section, we define the service classes supported by ED-MDA (see Tables 2-3). Then, we describe the details of ED-MDA functionality in the presence of multi-service traffic and HO calls (see Fig. 2-(b)); particularly, we will show how HO calls are prioritized over new calls to maintain high resources utilization.

Table 2. QoS requirement: 4 service classes

Class-id	**Delay sensitivity**	**Throughput sensitivity**	**Example**
AC[1]	Very high (≤ 150ms)	No	Voice
AC[2]	High (≤ 300ms)	High	Videoconference
AC[3]	Medium	Medium	Streaming
AC[4]	No	No	Best effort

Service classes

Generally, we can identify four types of classes related to the traffic nature either real time connections or not (see Table 2).

a) Voice traffic (real-time connections: RTC)

Typically, voice traffic is delay-sensitive, but tolerates some frame losses. Acceptable upper bound limit of one way voice delay according to ITU-T G.114 must be smaller than 150ms [17]. Besides, voice packet loss (%) according to ITU-T P.862 (Perceptual Evaluation of Speech Quality, PESQ) should be kept below 10% [18]. Empirically, this value is found to reach 6.79% if the maximum allowed delay is 120ms [18].

b) Video traffic (real-time connections: RTC)

Video traffic is delay-sensitive, but it tolerates some frame losses. In fact, the acceptable upper bound limit of one way video delay according to ITU-T H.261 must be smaller than 300ms [17]. Additionally, video packet loss (%) should be kept below 12%. Typically, a video traffic source generates frames at a constant rate over its active period. A popular example of RTC is a videoconference connection.

c) Video traffic (no real-time connections: NRTC)

In this kind of video traffic, the requirement is to reserve enough time-slots in each of the MDAOP periodicity sub-intervals in order to transmit 1 video packet (one downstream) of size $L_{\text{streaming}}$ over a clique having a transmission rate of C. This service class is considered as unidirectional calls when compared to voice/videoconference connections, the latter being bidirectional calls. An example of NRTC is video streaming.

d) Data traffic (no real-time connections: NRTC)

Data traffic is delay tolerable, but data packet loss (%) should be kept below 10%.

Recall that NRTCs traffic (AC [3] and AC [4]) in the proposed ED-MDA will be treated during the contention period (CP) and will not be studied in the scope of this paper. Indeed, we are only interested in supporting RTCs (AC [1] and AC [2]) that require low delays, low packet losses, and traffic differentiation support which are difficult to guarantee with both IEEE.802.11 protocol [2] and current MDA [6].

2) Priority setup in the presence of handover/new Calls

In the ED-MDA model we are studying, we consider a WMN consisting of several cliques. A clique serves heterogeneous users who require different service classes in the form of new calls generated in the clique or HO calls coming from adjacent cliques. According to the requirements of each service class in terms of blocking and delay constraints, it seems natural to consider in the DTIM interval this setup:

$$\Pr(AC[1]) > \Pr(AC[2]) > \Pr(AC[3]) > \Pr(AC[4]) \tag{1}$$

where Pr(x) denotes the priority of service class x.

Since we take HO calls into account and the no-delay-sensitive service classes are treated in CP duration in our proposed ED-MDA scheme (see Fig. 2-(b)), the problem under investigation is transformed and will be subject to different set of priority rules. The newly setup priority and QoS requirements as stated in (1) and Table 2 during the CFP period will become, in our model as follows:

Table 3. QoS requirement: 4 new service classes

Sub-class-id	**Delay sensitivity**	**Throughput sensitivity**	**Service class**
AC[1.1]	Very high (≤ 150ms)	No	Handoff voice calls
AC[1.2]	Very high (≤ 150ms)	No	New voice calls
AC[2.1]	High (≤ 300ms)	High	Handoff video calls
AC[2.2]	High (≤ 300ms)	High	New video calls

$$\Pr(AC[1.1]) > \Pr(AC[1.2]) > \Pr(AC[2.1]) > \Pr(AC[2.2]) \tag{2}$$

where $\Pr(AC[i.j]$ denotes the priority of service class i, $i \in [1..2]$ and the sub-class-id j denotes whether the call is a HO or new call with $j \in [1..2]$.

We assume that each call request (i.e., call) of service class $AC[i.j]$ includes the corresponding blocking probability and delay thresholds, $B_{AC[i.j]}$ and $D_{AC[i.j]}$ for $i \in [1..2]$ and $j \in [1..2]$, respectively. In addition, the call request contains a field called the relative blocking probability, $RB_{AC[i.j]}$ (see (3)), for accumulated blocking probability in WMN. $RB_{AC[i.j]}$ is updated at each intermediate WMN node and it is defined to be the total number of admitted calls per the total number of both admitted and rejected calls in WMN (computed according to the service class). Fig. 3 shows the main fields of the call request. $RB_{AC[i.j]}$ will be compared to the $B_{AC[i.j]}$ threshold according to each service class in our proposed admission control algorithm (see Section 4) and it is expressed as shown by (3).

$$RB_{AC[i.j]} = \frac{A_{AC[i.j]}}{A_{AC[i.j]} + R_{AC[i.j]}} \quad , i \in [1..2], j \in [1..2] \quad RB_{AC[i.j]} \tag{3}$$

Source	Destination	AC [i.j]	$B_{AC[i,j]}$	$D_{AC[i,j]}$	$RB_{AC[i,j]}$

Fig. 3. Main fields of the call request packet in IEEE 802.11s (ED-MDA)

3 The Real Time Connections Analysis in the ED-MDA Per-Flow Scheme

In this section, we analytically derive the MDAOP duration in number of time-slots $\delta^k_{AC[i.j]}$ and the MDAOP periodicity $\pi^k_{AC[i.j]}$ parameters in ED-MDA scheme in order to support QoS of RTC.

3.1 Bounds Conditions for CFP Adaptive Adjustment

In this sub-section, we introduce our model along with the underlying assumptions, so as to statistically describe the characterization of the traffic being studied and the guard computation of service classes in ED-MDA.

3.1.1 Traffic Characterization

We assume that our model follows the $M/M/1$ queuing model, the call arrival is described by a Poisson process and the inter-arrival time is exponentially distributed. We denote by E, $E=\lambda h$, the total amount of traffic offered in erlangs where λ is the call arrival rate and h denotes the call-holding time. The Erlang-b formula [16] is expressed by the probability: $P_b^{AC[i.j]} = \psi_b[E, T_{AC[i.j]}]$; this means that a new arriving call is blocked (meaning that all time-slots are busy) until a time-slot becomes free. Since the $\psi_b[E, T_{AC[i.j]}]$ formula assumes that the number of resources is known in advance (which is not the case in our model), we bound the Erlang-b to find an upper limit of the time-slots according to each service class. In our model, blocking threshold is considered to be known, while the number of time-slots is unknown. This can be better illustrated in UBTS (see ALG. I) where the probability of blocking is bounded by the required blocking probability $B_{AC[i.j]}$ and 1. $T_{AC[i.j]}$ is the upper bound of the number of resources (time-slots) and it is determined based on Erlang-b computation (see Alg. I).

ALG. I: COMPUTING THE NUMBER OF TIME-SLOTS FOR EACH SERVICE CLASS: UBTS

Input: $AC[i.j]$: service class, E: total traffic, $B_{AC[i.j]}$: blocking probability threshold, *CFP*, τ : time-slot duration, *T:*DTIM interval.

Output: $T_{AC[i.j]}$: Time-slots upper bound for each service class .

Hypothesis: **No resources (*k*=0) condition** $\psi_b[E,0]=1$;

we assume that $\psi_b[E,k-1] > \psi_b[E,k]$ for any *k* resources;

Step0: **Bounded recurrence form of Erlang-b**

$$B_{AC[i.j]} \le P_b^{AC[i.j]} = \psi_b[E, T_{AC[i.j]}] = \frac{E^{T_{AC[i.j]}} / T_{AC[i.j]}!}{\sum_{k=0}^{T_{AC[i.j]}} E^k / k!} \le 1 ;$$

For fixed k, we can express $P_b^{AC[i.j]}(k)$ as follows:

$$B_{AC[i.j]} \le P_b^{AC[i.j]}(k) = \frac{E \times \psi_b[E,k-1]}{E \times \psi_b[E,k-1]+k} \le 1 \;;$$

then $$B_{AC[i.j]} \le P_b^{AC[i.j]}(k) = \frac{E \times P_b^{AC[i.j]}(k-1)}{E \times P_b^{AC[i.j]}(k-1)+k} \le 1 \;;$$

Step1: **Bounds conditions for *k***

We determine $0 \le k = P_b^{AC[i.j]}(k-1)[\frac{E}{P_b^{AC[i.j]}(k)} - E] \le E \times P_b^{AC[i.j]}(k-1)[\frac{1}{B_{AC[i.j]}} - 1]$

Step2: **Time-slots upper bound according to each service class**

if ($i == 1$ *and* $j == 1$) **then** $T_{AC[1.1]} = \left\lceil \frac{CFP}{\tau} \right\rceil = \left\lceil \frac{\alpha^* T}{\tau} \right\rceil$;

else for $k\{1,2,..,T_{AC[i.j]}\}$, $T_{AC[i.j]} = \left\lceil E \times P_b^{AC[i.j]}(k-1)[\frac{1}{B_{AC[i.j]}} - 1] \right\rceil$;

where $(i=2, j \in [1.2]) \cup (i=1, j=2)$

where $B_{AC[i.j]}$ is a given blocking probability threshold for each service class; for example for voice call, $B_{AC[1.2]}$ is equal to 6%. This constraint is used to guarantee a smaller rate of blocking probability inside the network in order to meet users's expectations and to subsequently increase the operator's benefits.

By periodically applying steps (0-2) in UBTS, ED-MDA can compute an adequate upper bound of the number of time-slots to be reserved for each service class. Although this reservation scheme guarantees the finding of suitable time-slots, it may cause interferences among time-slots that correspond to different users. Thus, a guard interval mechanism must be set to ensure that distinct transmissions belonging to different service classes do not interfere with each other.

3.1.2 Guard Computation

The class with highest priority in our case corresponds to HO voice calls; these calls could use up the CFP duration for voice packet transmission in the absence of other service classes, as shown in Fig.2 (b). Additionally, a HO voice call takes up a minimum portion of the whole DTIM ($CFP - G_{AC[1.2]}$ denoted by X) in order to transmit its packets without sharing them with other classes as shown in Fig.2-(b). ED-MDA manages the priorities among new and HO calls according to inequality (2); it operates under a time-slot reservation scheme expressed by the adjustment of CFP to maintain the HO priority (see Fig.2 (b)) for video or voice calls.

The parameters $G_{AC[1.2]}, G_{AC[2.1]}, G_{AC[2.2]}$ are the respective guard intervals (computed according to service class). The guard interval threshold in ED-MDA is defined to be the maximum number of slots, during DTIM interval, that traffic of a given service class can use. The purpose of using guard intervals is to give higher priority to handoff calls over new calls and to protect against interferences among service classes. This is expressed in the following equations:

$$G_{AC[1.1]} = [0, \lceil CFP = \alpha^{*} T \rceil] \tag{4}$$

$$G_{AC[1.2]} = [\lceil X \rceil + 1, \frac{T_{AC[1.2]}}{K_{AC[1.2]}}] \tag{5}$$

where $K_{AC[1.2]}$ is the number of slots for voice

$$G_{AC[2.j]} = [\frac{T_{AC[i.2]}}{K_{AC[i.2]}} + 1, \frac{T_{AC[2.j]}}{K_{AC[2.j]}}] \tag{6}$$

where $K_{AC[2.j]}$ is the number of slots for video. If $i = 1$ then $j=1$ and if $i=2$ then $j=2$ in (6).

3.2 ED-MDA Formulation

In per-flow ED-MDA scheduling, $2\delta^{k}_{AC[i.j]}$ is the number of time-slots reserved for each service classes (see Eq. 7) in each of the $\pi^{k}_{AC[i.j]}$ (see Eq. 8) sub-intervals that satisfies a hard constraint on a maximal delay $D^{\max}_{AC[i.j]}$ for maximum number of hops m in a path. In fact, we need to reserve enough time-slots during the sub-intervals for transmitting 2 voice or video conference (RTC) packets of size $L_{AC[i.j]}$, one upstream and the other downstream (bidirectional calls) over a clique having a transmission rate C. We note that this transmission occurs after duration $AIFS_{AC[i.j]}$. To prevent exceeding the one-hop delay, the periodicity $\pi^{k}_{AC[i.j]}$ in the ED-MDA reservation request has to be sufficiently upper and lower bounded by: $\pi^{k}_{AC[i.j]} \geq T / D^{\max}_{AC[i.j]}$. For the sake of simplicity, we consider a uniform distribution of $D^{\max}_{AC[i.j]}$ over interfering links even though a better repartition may take into account the non uniformity of traffic load over these links. Thus, the MDAOP duration (Eq. 7) and periodicity (Eq.8) are expressed as follows:

$$2\delta^{k}_{AC[i.j]} = 2\left\lceil \frac{AIFS_{AC[i.j]} + \frac{L_{AC[i.j]}}{C}}{\tau} \right\rceil \quad k \in N \text{ and} \tag{7}$$

$$\pi^{k}_{AC[i.j]} = \frac{DTIM}{D^{\max}_{AC[i.j]}} = \frac{T}{D^{\max}_{AC[i.j]}} \quad \text{where} \tag{8}$$

$$D^{\max}_{AC[i.j]} = \frac{D_{AC[i.j]}}{m} \tag{9}$$

The maximal delay is denoted by $D^{\max}_{AC[i.j]}$, i.e., the hard constraint on maximal delay for maximum number of hops m in a path and $D_{AC[i.j]}$ is the required delay by the service class $AC[i.j]$.

4 Admission Control Scheme EAC

The Call Admission Control (CAC) is an important task to ensure QoS requirements in multi-service WMNs. Therefore, in this study, ED-MDA is combined with a CAC for multimedia service networks, called EAC. The main objectives of EAC are: (1) differentiating between traffic with different QoS requirements; (2) providing an efficient adjustment of the CFP period to maintain high utilization of radio resources; and (3) providing higher priority to HO calls.

4.1 CFP Values' Update Based on EAC Operation

To guarantee QoS to RTC in terms of delay and blocking probability, we need an efficient fine-tuning of the CFP duration according to service classes. EAC defines three variables: (1) α: holds the current fraction of the DTIM period reserved for MDA; (2) β: a multiplicative factor which is used to update CFP; and (3) VCS (Violation Constraint State): assumes 0 when the state of the required blocking threshold is violated; otherwise assumes 1. VCS-v is defined as a vector of VCSs (e.g., VCS_v="01").

The first variable is common in IEEE 802.11s whereas the latter two variables are EAC specific.

The parameter β is a multiplicative factor used to update the value of α and subsequently CFP according to the service class and VCS. A similar multiplicative factor (equal to 2) is implicitly defined in the original IEEE 802.11 binary exponential backoff; Contention Window (CW) is doubled upon each transmission failure. However, β in EAC can be either assigned statically, i.e. before runtime and remains constant or dynamically (adaptive).

EAC first stores the most recent S states of VCSs upon each call request where S indicates the length of VCS_v. For example, in the case of S=2 and VCS_v is equal to "10" means that VCS is equal to 1 followed by VCS is equal to 0. EAC re-adjusts α accordingly to the rules shown in Alg.II. Besides, EAC checks the value of VCS_v; if the latter is equal to "11" and the current VCS is equal to 1 then α is divided by β. In other words, a slow decrease of α /CFP is preferred in this case since a QoS violation (for at least two states) is unlikely to be encountered. In the case where VCS_v is equal to "00" indicating a QoS violation and that the current VCS is equal to 1, α will revert to its current value. Furthermore, EAC multiplies the value of α by β when the current VCS is equal to 0

4.2 Time-Slots Reservation over ED-MDA by EAC Algorithm

EAC algorithm performs the necessary verification of the incoming traffic in terms of the DTIM utilization and decides whether the DTIM parameters (e.g., fractions of the DTIM parameter) should be modified according to the service class. If a violation in the blocking probability constraint ($B_{AC[i.j]}$) is encountered during a period T, EAC adjusts all the DTIM parameters in a way that may lead to a better utilization of the available resources. To realize DTIM parameters variation, we assume the existence of an online measurement mechanism that measures DTIM parameters variations and computes a suitable time interval, denoted by T, to capture periods with minimal DTIM parameters variation.

The algorithm starts by looking for RTC traffic (Step 0-1 of Alg. II); once found, EAC proceeds with the initialization phase (see step 2 of Alg. II). We apply UBTS algorithm to search for initial feasible CFP adjustments (i.e., computes adequately the upper bound of the number of time-slots for each $AC[i.j]$). EAC initializes this feasible solution rather than a randomly generated one. Then, according to our ED-MDA classification (see Table 3), EAC periodically checks whether DTIM exceeds a certain threshold ($\frac{CFP}{\tau}$) as well as the ($\frac{\alpha_{AC[i.j]}T}{\tau}$) threshold for the handoff voice $AC[1.1]$ and for the others service classes, respectively. Based on the above, it then rejects or accepts the incoming calls (see Alg. II, line-7).

In case of a rejected call, EAC increments the rejection parameter (Alg. II, line-8) that can be used later for eventual adjustment of CFP parameters. Otherwise, (a call is accepted), EAC recalculates the relative blocking parameter $RB_{AC[i.j]}$ (see Eq.3) and compares it to the given threshold blocking parameter; if $RB_{AC[i.j]}$ exceeds the threshold ($B_{AC[i.j]}$), then a violation of the constraint is detected; in this case, EAC re-adjusts the parameters accordingly as depicted in line-13 of Alg. II. Otherwise, it uses the information from line-11 of Alg. II.

In fact, EAC makes use of blocking probability statistics, VCS and VCS_v that are collected during the period of time T , to dynamically trigger or not CFP adjustment at the end of the time period (see Section 4.1).

ALG. II: EFFICIENT ADMISSION CONTROL: EAC

Input: Call request: k ($AC[i.j]$, $B_{AC[i.j]}, D_{AC[i.j]}$) in clique q_j where $B_{AC[i.j]}$ and $D_{AC[i.j]}$ are the blocking probability and the delay required for the service class AC[i.j], respectively.
Output: Admission decision: accept or reject and CFP adjustment

1 Receive the call request from user with QoS requirements ($B_{AC[i.j]}, D_{AC[i.j]}$)
2 Step 0: **Throughput-sensitive traffic**
if ($k \in NRTC$) **then** the call is served in CP duration;
3 Step 1: **Delay and throughput-sensitive traffic**
if ($k \in RTC$) {
4 Step 2: **Initialization**
if ($k \in AC[1.1]$) **then** i=Voice, j=HO;
if ($k \in AC[1.2]$) **then** i=Voice, j=New;
if ($k \in AC[2.1]$) **then** i=Video, j=HO;
if ($k \in AC[2.2]$) **then** i=Video, j=New;
5 Step 3: Upper bound **Time-slots computation**
Call UBTS /* Initial feasible CFP adjustment based time-slots computation for each $AC[i.j]$ */
6 Step 4: **DTIM utilization test**
$U = \delta_{AC[i.j]} + \sigma$;
7 **if** ($U > (\frac{CFP}{\tau})$ && $U > \frac{\alpha_{AC[i.j]}T}{\tau}$)
8 **then** $R_{AC[i.j]}{+}{+}$; exit; // The call k is rejected
9 $A_{AC[i.j]}{+}{+}$; // The call k is admitted in CFP duration
10 $\sigma = \sigma + \delta_{AC[i.j]}$; // Update of DTIM utilization
11 Step 5: **CFP adjustment**
if ($RB_{AC[i.j]} < B_{AC[i.j]}$) **then** VCS =1;
// No violation of the constraint: currently state
12 **then if** (VCS_v=="00")
$\alpha_{AC[i.j]} = \alpha_{AC[i.j]}$; **else** $\alpha_{AC[i.j]} = \alpha_{AC[i.j]} / \beta$;
13 **else if** (VCS==0) **then** $\alpha_{AC[i.j]} = \alpha_{AC[i.j]} \times \beta$;
// Violation of the constraint: currently state
Repeat Step 4 Every period of time T }
14 Update of the number of time-slots for each $AC[i.j]$;

5 Simulation Results

In this section, we conduct a simulation study using ns-2 to evaluate and compare the performance of ED-MDA with other existing schemes. We evaluate several metrics: 1) end-to-end delay; 2) packet loss rate; and 3) call blocking/dropping probability. Note that the end-to-end delay is the sum of the access delays (the queuing delay, which is negligible in low load, and the contention delay) experienced in hops along the path's call from the MC to the gateway.

5.1 Simulation Configurations

The WMN topology used in simulations is arranged as a regular grid of 5x5 802.11 stations acting as MRs. More precisely, the node spacing is about 100 m and the position of each node is deviating from the regular grid by choosing a random perturbation of an angle in [0, 2π] and a radius in [0 m, 25 m]. This kind of disturbance has been adopted in several existing WMN performance studies (e.g., [3, 8]). We notice that in tests with MDA, all MRs apply the DCF and its MAC parameters after the scheduling of MDAOPs. In the first scenario (scenario 1), simulations are performed under combination of two service classes (voice and video calls) while in the second scenario (scenario 2) four service classes (HO/new voice and HO/new video calls) are considered as described in sub-section 3.3.2. In the simulation results, we compare our proposed ED-MDA performance to: (1) WLAN 802.11 CSMA/CA for DCF [2]; (2) WLAN 802.11s MDA [8]; (3) WLAN 802.11s EDCA [6]; and (4) An improved version of MDA that considers dynamic relocation [8] of reserved MDAOPs two-hop away; we call this scheme MDA-DR.

5.2 Results Analysis

MAC Access Method Analysis—**DCF/MDA/MDA-DR/EDCA/ED-MDA Using Scenario 1:** As presented in Section 1, the IEEE 802.11s draft 2.0 allows an optional contention-free MDA access method besides the well-known CSMA/CA contention-based access DCF method. We study the performance of these access methods when transmitting data on the channel of each clique compared to our proposed scheme, ED-MDA, which is based on MDA while taking into account multi-service traffic. Moreover, ED-MDA prioritizes HO calls over new calls; this feature will be simulated in scenario 2.

Fig. 4 presents the average end-to-end delay experienced by network voice calls when using the various simulated schemes. In the case of low offered load (1 to 4

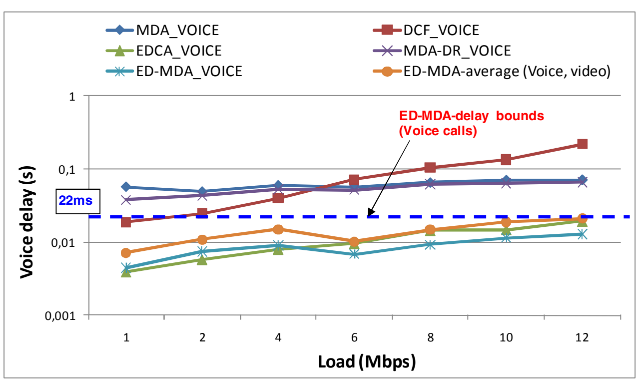


Fig. 4. Delay for voice calls using different access methods

Mbps), where collisions are very rare, the DCF method provides lower delays than MDA and MDA-DR. We note that DCF, MDA and MDA-DR do not grant higher priority to voice calls compared with video calls; this explains their poor performance with respect to voice delay. Besides, due to the scheduling of MDAOPs although the time-slots are available in a low offered load condition, the average delay is higher with MDA and MDA-DR compared to DCF. In highly offered load situation (e.g., 10 Mbps), the average delay with MDA or MDA-DR does not exceed 224 ms; it is bounded by the DTIM interval, which is equal to 32ms, multiplied by the maximum number of hops, which is equal to 7 in our topology. Fig. 4 shows that EDCA and ED-MDA schemes present lower delays, which is expected since these two schemes consider traffic differentiation and give priority to voice calls.

In the case of low offered load, the methods based on MDA wait for longer periods of time before transmitting in specific reserved contiguous time-slots even when collisions are rare and time-slots are available. However, in high offered load situation the delay in these schemes is bounded as explained above. Nevertheless, the delay provided by DCF and EDCA increases boundlessly with the increase of the offered load. We observe that the link delay with ED-MDA experienced by voice packets never exceeds the hard constraint on the maximal delay of $D_{AC[1]}/7=150/7=21.42$ ms.

Fig. 5 shows that ED-MDA delay, experienced by video packets never exceeds the hard constraint on the maximal delay of $D_{AC[2]}/7=300/7=42.85$ ms.

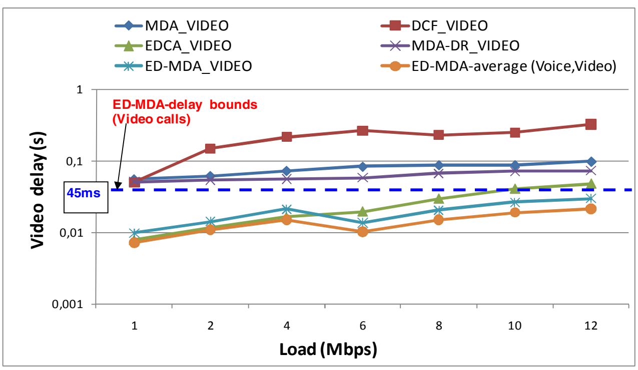


Fig. 5. Delay for video calls using different access methods

MAC Access Method Analysis—**MDA/ED-MDA Using Scenario 2:** in addition to video and voice calls, we take into account new and HO calls in this scenario.

In order to study the impact of the reserved guard $CFP\text{-}G_{AC[1.2]}$ on the HO voice calls dropping probability (see Fig.6), we simulated the network behavior with the same parameters as in ED-MDA (called W.ED-MDA) and without ED-MDA (called W/O.ED-MDA). Also, similar tests are done to investigate the impact of the guard threshold for HO video calls. We note that the admission control policy of EAC, with guard thresholds, offers HO calls dropping rate much lower than without guard thresholds (i.e., MDA). This dropping rate is 2 to 12 times lower for an offered load ranging from 2 to 10 Mbps. Indeed, up to 2 Mbps, all HO voice or video calls with ED-MDA are admitted but over 4Mbps, the dropping probability ranges from

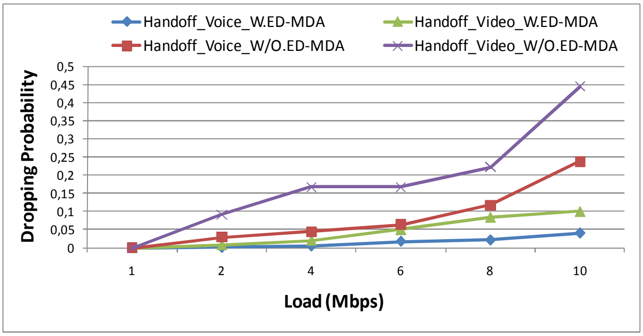


Fig. 6. HO calls dropping probability with and without ED-MDA

0.4% to 2% for HO voice and HO video, respectively, which is very acceptable for users. However, when offered load is equal to 10 Mbps, the dropping probability does not exceed 4% which is still acceptable for voice traffic.

Fig. 7 shows that new connections blocking probability increases with the offered load. We observe that for offered loads varying between 1 and 8Mbps using ED-MDA scheme, we get an acceptable blocking rate (0 to 5% for new voice calls and 0 to 9% for new video calls); beyond these values, the blocking rate for video calls exceeds 11% when the offered load exceeds 8Mbps. Fig. 7 also shows that the lowest blocking rate is offered when using W/O.ED-MDA; this is expected since the new and HO calls are treated similarly; as long as time-slots are available, these calls are admitted. However, when a guard threshold is used for new calls using ED-MDA, a part of the CFP (i.e., a number of time-slots) is reserved exclusively for HO voice calls and the guard threshold for HO video calls is higher than the guard threshold for new video calls; the new calls have less chance to be admitted which causes the increase of call blocking probability.

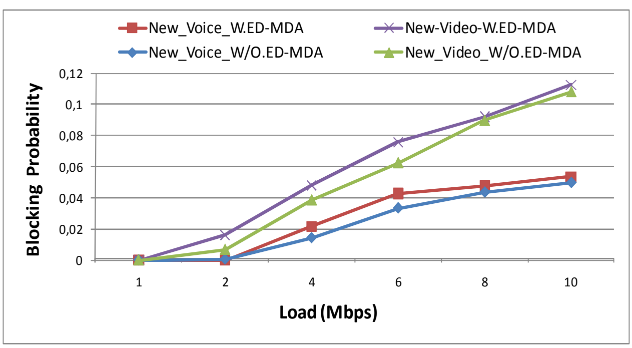


Fig. 7. New calls blocking probability with and without ED-MDA

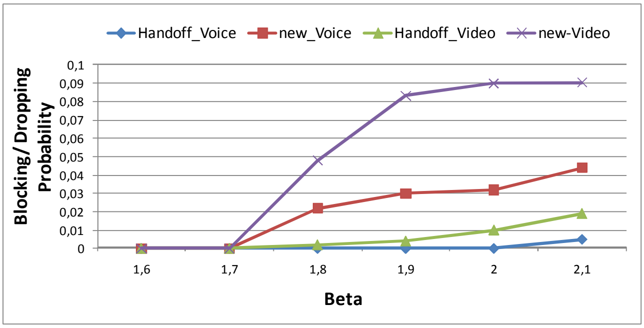


Fig. 8. Impact of β on the HO and new calls Blocking/Dropping probability for an offered load equal to 10 Mbps with ED-MDA (S=2)

Fig. 8 shows the blocking and dropping probability while varying β. We observe that the dropping of handoff calls is much lower (starting from β equal to 1.8) than new calls regardless of the nature of the call (voice or video). An interrupted communication is a very frustrating phenomenon that may happen to MCs. Thus, our proposed admission control protocol EAC coupled with ED-MDA avoids the forced termination of an ongoing call at the expense of slightly higher blocking of new calls. This is a price worth paying to provide low delays and loss rates while satisfying QoS requirements of multi-service WMNs.

6 Conclusion

In this paper, we have shown that MDA that carefully reserves dimensioned CFP according to each service class (ED-MDA) provides better voice/video QoS in terms of packet loss, delay and dropping rate. Furthermore, we proposed an admission control algorithm EAC that takes the HO and multiple service classes into account. Our proposed scheme ED-MDA combined with EAC is based on the guard interval principle which reserves necessary time-slots (i.e., resources) to meet and grants high priority to HO calls. Compared to MDA mechanism without guard interval, the simulation results show that our proposed solution outperforms the current MDA and many others existing schemes regarding the management of HO calls. Indeed, it offers a lower dropping probability and a lower delay for voice/video HO calls at the cost of a slightly higher blocking rate for new calls.

References

[1] Akyildiz, I.F., Wang, X., Wang, W.: Wireless mesh networks: a survey. Computer Networks 47(4), 445–487 (2005)

[2] Ergin, M., Gruteser, M., Luo, L., Raychaudhuri, D., Liu, H.: Available bandwidth estimation and admission control for QoS routing in wireless mesh networks. Computer Communications 31, 1301–1317 (2008)

[3] Rezgui, J., Hafid, A., Gendreau, M.: Distributed Admission Control in Wireless Mesh Networks: Models, Algorithms, and Evaluation. IEEE Transactions on Vehicular Technology 59(3), 1459–1473 (2010)

[4] Rezgui, J., Hafid, A., Gendreau, M., Rong, B.: Congestion-aware Clique-based Handoff in Wireless Mesh Networks. In: Proc of IEEE MSN, pp. 388–395 (2009)

[5] Raniwala, A., Gopalan, K., Chiueh, T.: Centralized Channel Assignment and Routing Algorithms for Multichannel Wireless Mesh Networks. ACM of Mobile Computing and Communications Review 8, 50–65 (2004)

[6] Hiertz, G.R., Max, S., Zang, Y., Junge, T., Denteneer, D.: IEEE 802.11s MAC fundamentals. In: Proc. of IEEE MeshTech, pp. 1–8 (2007)

[7] Hiertz, G.R., Max, S., Junge, T., Denteneert, D., Berlemann, L.: IEEE 802.11s-Mesh deterministic access. In: Proc. of EW, pp. 1–8 (2008)

[8] Cicconetti, C., Lenzini, L., Mingozzi, E.: Scheduling and dynamic relocation for IEEE 802.11s mesh deterministic access. In: Proc. of IEEE Secon, pp. 19–27 (2008)

[9] Gallardo, J., Makrakis, D., Mouftah, H.: MARE: an Efficient Reservation-Based MAC Protocol for IEEE 802.11 s Mesh Networks. In: Proc. of IEEE MESH, pp. 97–102 (2009)

[10] Niculescu, D., Ganguly, S., Kim, K., Izmailov, R.: Performance of VoIP in a 802.11 wireless mesh network. In: Proc. of IEEE Infocom, pp. 1–11 (2006)

[11] Wei, H., Kim, K., Kashyap, A., Ganguly, S.: On admission of VoIP calls over wireless mesh network. In: Proc. of IEEE ICC, pp. 1990–1995 (2006)

[12] Bayer, N., Xu, B., Rakocevic, V., Habermann, J.: Application-aware scheduling for VoIP in Wireless Mesh Networks. Journal of Computer Networks 54(2), 257–277 (2010)

[13] Zhang, Y., Liu, D.: An adaptive algorithm for call admission control in wireless networks. In: Proc. of IEEE Globecom, pp. 3628–3632 (2001)

[14] Wei, Y., Lin, C., Ren, F., Raad, R., Dutkiewicz, E.: Dynamic handoff scheme in differentiated QoS wireless multimedia networks. Computer and Communications Journal 27(1), 1001–1011 (2004)

[15] Belbekkouche, A., Rezgui, J., Hafid, A.: QoS provisioning for Wireless Mesh and Optical Burst Switching Convergence. In: Proc. of IEEE WCNC, pp. 1–6 (2010)

[16] Kleinrock, L.: Queueing Systems. Theory, vol. 1, 448 pages. Wiley Interscience, New York (1975)

[17] Siegel, E.D.: Designing Quality of Service Solutions for the Enterprise. John Wiley & Sons, Inc., Chichester (1999) ISBN: 0471333131

[18] Shim, C., Xie, L., Zhang, B., Sloane, C.J.: How delay and packet loss impact voice quality in Voip. Qovia, Inc. (2003), `http://www.qovia.com`