A Collision-Free MAC Scheme for Multimedia Wireless Mesh Backbone

Ping Wang, *Member, IEEE,* and Weihua Zhuang, *Fellow, IEEE*

***Abstract*—Wireless mesh networking is a promising wireless technology for future broadband Internet access. In this paper, a novel collision-free medium access control (MAC) scheme supporting multimedia applications is proposed for wireless mesh backbone. The proposed scheme is distributed, simple, and scalable. Benefiting from the fixed locations of wireless routers, the proposed MAC scheme reduces the control overhead greatly as compared with conventional contention-based MAC schemes (e.g., IEEE 802.11). In addition, the proposed scheme can provide guaranteed priority access to real-time traffic and, at the same time, ensure fair channel access to the routers with data traffic. Unlike most of the existing MAC schemes which focus on single-hop transmissions, the proposed MAC scheme takes the intra-flow correlations between up-stream and down-stream hops of a multi-hop flow into consideration. To avoid buffer overflow at bottleneck routers, a simple but effective congestion control mechanism is proposed. Simulation results demonstrate that the proposed scheme significantly improves the delay performance of real-time traffic and the end-to-end data throughput, as compared with IEEE 802.11 and distributed packet reservation multiple access (DPRMA). The performance analysis of the proposed scheme is also presented. The accuracy of the analytical results is verified by computer simulations.**

***Index Terms*—Wireless mesh backbone, multi-hop connection, QoS, multimedia applications, priority access, fairness, throughput, delay, congestion control.**

I. Introduction

WIRELESS mesh networking is a promising wireless technology for future broadband Internet access. A typical example of the network consists of wireline gateways, wireless routers, and mobile stations (MSs), organized in a three-tier architecture [2], as shown in Fig. 1. The third tier is the wireless access networks, through which users access the Internet. Wireless access networks includes WLANs, ad hoc networks, and cellular networks, among which the mobile users can seamlessly roam. The second tier is the wireless mesh backbone, consisting of a number of wireless routers at fixed sites. Each wireless router not only delivers traffic from the access networks in its coverage, but also forwards the traffic from and to its neighboring routers. The first tier is the mesh gateways, which connect the wireless mesh backbone to the Internet backbone. Normally a wireless mesh network covers a large geographical area. Thus, multi-hop communications are usually necessary, where a traffic flow from a source to its far away destination traverses multiple intermediate routers.

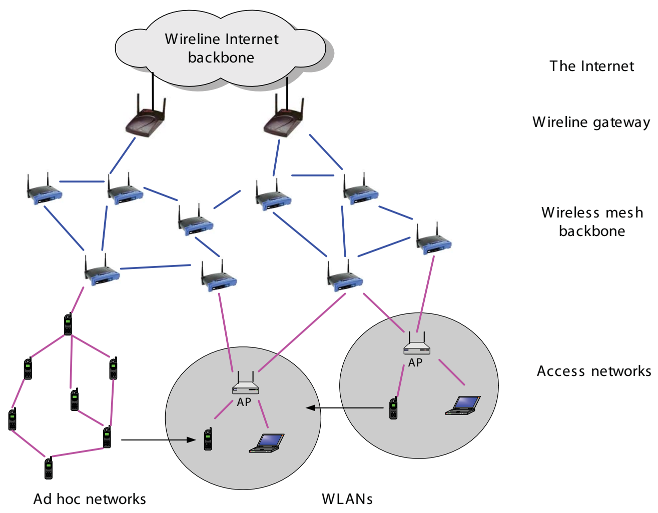

Fig. 1. An architecture of a broadband wireless mesh network.

With the growing demand for multimedia applications, wireless mesh networks are expected to support heterogeneous traffic types (e.g., voice, video, and data traffic) with various quality-of-service (QoS) requirements. Given a physical layer, a properly designed medium access control (MAC) scheme is the key to efficiently allocate radio resources and ensure QoS.

There are extensive research results on MAC over mobile ad hoc networks in the literature (a literature survey is given in [3]). However, two unique characteristics of the wireless mesh backbone result in that it may not be effective or efficient to directly apply existing MAC schemes proposed for ad hoc networks to the wireless mesh networks [4]. First, most existing MAC schemes for ad hoc networks are designed to handle node mobility with power consumption constraints. For the wireless mesh backbone, the wireless routers are usually located at fixed sites with wired power supply. Thus, the node mobility and power consumption should not be the main consideration for the MAC design. Second, contention-based MAC schemes (e.g., IEEE 802.11 [5]) are one major stream for wireless ad hoc networks. However, the traffic volume in the wireless mesh backbone may be much higher than that in an ad hoc network due to traffic aggregating at each router. It is well known that, when traffic load is

Manuscript received December 14, 2007; revised July 2, 2008 and October 22, 2008; accepted February 16, 2009. The associate editor coordinating the review of this paper and approving it for publication was G. Mandyam.

P. Wang is with the School of Computer Engineering, Nanyang Technological University, Singapore, 639798 (e-mail: wangping@ntu.edu.sg).

W. Zhuang is with the Centre for Wireless Communications, Department of Electrical and Computer Engineering, University of Waterloo, 200 University Avenue West, Waterloo, Ontario, Canada N2L 3G1 (e-mail: wzhuang@bbcr.uwaterloo.ca).

This work was supported by a research grant from the Natural Science and Engineering Research Council (NSERC) of Canada. This paper is presented in part in a paper at IEEE ICC 2008 [1].

Digital Object Identifier 10.1109/TWC.2009.071406

1536-1276/09$25.00 © 2009 IEEE

heavy, contention-based MAC schemes suffer from serious collisions due to the severe contention, leading to dramatically decreased throughput and increased delay. As pointed out in [2], for application to wireless mesh networks, all existing MAC schemes need to be enhanced or re-invented. So far, very limited work has been done to enhance the existing MAC schemes or design a new MAC scheme specifically for wireless mesh networks. To enhance IEEE 802.11, in [6], an end-to-end reservation protocol is proposed to support QoS of real-time traffic. In [7], a new protocol named Wireless Channel-oriented Ad-hoc Multi-hop Broadband (W-CHAMB) is proposed based on time-division multiple access/time division duplex (TDMA/TDD) technology. In [8], with a cross-layer design principle, an interference aware MAC scheme is proposed for a code-division multiple access (CDMA)-based wireless mesh backbone.

In this paper, our objective is to propose a MAC scheme for a single-channel wireless mesh backbone to provide QoS support for multimedia applications. Different from the existing MAC schemes, our MAC scheme design benefits greatly from the fixed network topology of wireless mesh backbone. With the router location information, collision-free transmissions are scheduled in a deterministic way, without the request to send/clear to send (RTS/CTS) handshaking prior to every packet transmission. Thus, the overhead is greatly reduced, compared with contention-based MAC schemes. Meanwhile, the deterministic schedule in our scheme is adaptive to the traffic dynamic and can achieve maximal spatial reuse. By eliminating collisions, reducing overhead, and achieving maximal spatial reuse, the proposed scheme achieves much higher resource utilization than contention-based schemes. In addition, the proposed scheme can provide guaranteed priority access to real-time traffic and, at the same time, ensure fair channel access to data traffic. In contrast to contention-based MAC schemes, where the real-time traffic may suffer from performance degradation when the data traffic load increases [9], the performance of real-time traffic in the proposed scheme is not affected by the data traffic load.

The rest of this paper is organized as follows. The system model is described in Section II, and the proposed MAC scheme is presented in Section III. The performance of the proposed scheme is analyzed in Section IV. Section V is devoted to numerical results of the performance of the proposed scheme, followed by concluding remarks in Section VI.

II. The System Model

Consider a wireless mesh backbone, which consists of a number of routers located at fixed sites and covers a large geographical area. All routers are synchronized in time[1]. There is a single information channel in the network, through which all the routers send their packets. Two routers are one-hop neighbors with each other if they are within the transmission range of each other. Based on the fixed locations of routers, the transmission power and rate for each wireless link can be appropriately determined, so that the required transmission accuracy at each link can be achieved and two or more links (which are more than two-hop away[2]) can transmit simultaneously without corrupting each other's transmission. As there is no central controller in the wireless mesh backbone, distributed MAC is required. Distributed MAC is more challenging than centralized MAC, because each node does not have complete information of other contenders, and there is no efficient way to let one node control transmissions of other nodes.

Unlike single-hop wireless networks (e.g., WLANs), the multi-hop wireless mesh backbone presents more challenges to the MAC scheme design. The hidden terminals bring more collisions which can seriously degrade the resource utilization. The locations of the contending flows can greatly affect the channel access opportunity of each flow, resulting in serious unfairness (starvation of some flows) and priority reversal problems (i.e., a high-priority flow gets a smaller chance to access the channel than its low-priority counterpart) [13]. The wireless mesh backbone has a large scale, requiring the desired MAC scheme to be scalable such that, when the network scale increases, the complexity and overhead of the MAC scheme do not increase dramatically, and the network performance does not degrade significantly.

Heterogeneous traffic types including voice, video, and data traffic are supported with different QoS requirements (e.g., delay constraints for voice and video traffic, throughput and fairness requirements for data traffic). We assume that a routing protocol is in place to choose the path from the source to the destination of each flow.

III. The Proposed Scheme

A. Distributed Time Slot Allocation

The time is partitioned into slots of constant duration, which are allocated to each router in a distributed manner. Once a router is allocated a slot, it can transmit one (or multiple) packets to one (or multiple) one-hop neighbor(s), and all its one-hop and two-hop neighbors are not allocated the same slot in order to avoid packet transmission corruption. The same slot can be allocated to the routers which do not interfere with each other to achieve spatial reuse. As shown in Fig. 2-(a), one slot consists of two portions: the first portion is the control part, occupying a very small fraction of the whole slot time. The control part is used to determine whether or not a router can transmit its packets in that slot; the second portion is the transmission part, dedicated to packet transmissions. The control part is further divided into several mini-slots, indexed sequentially with numbers 1,2,3, etc. Each router is assigned one mini-slot, but one mini-slot may be assigned to different routers. The mini-slot assignment algorithm is presented in subsection III-B.

When a router (say router A with mini-slot k) has packet(s) to transmit, it first monitors the mini-slots from 1 to $k-1$. If a jamming signal[3] is detected at any of the mini-slots, it gives

[1]Synchronization can be provided by Global Positioning System (GPS) or other advanced synchronization techniques [10], [11]. Synchronized transmissions have also been adopted by WiMAX (Worldwide Interoperability for Microwave Access), where a Time-Division Duplex (TDD) protocol is applied to coordinate simultaneous transmissions on multiple links [12].

[2]Two links are two-hop away when the receiver of each link is two-hop away from the source of the other link.

[3]A jamming signal is a busy-tone signal sent by a transmitter to indicate that the channel is busy. It does not carry any information bit sequence.

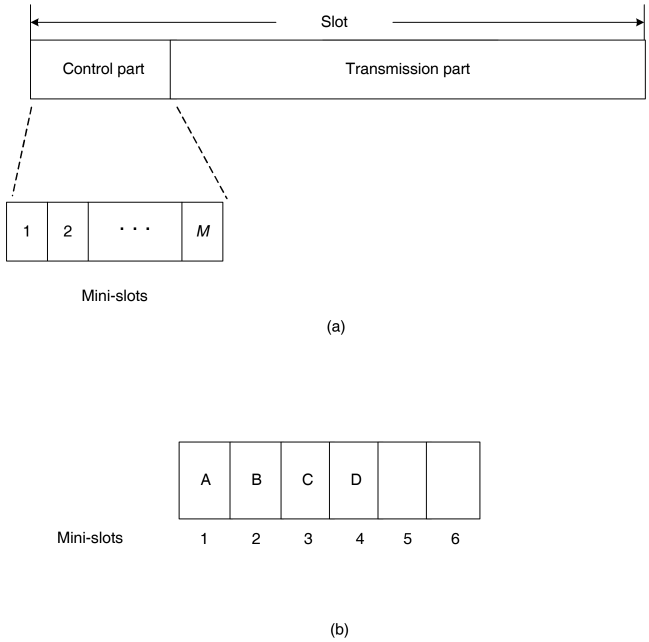

Fig. 2. The slot structure in the proposed scheme.

up the transmission at the current slot. Otherwise (i.e., the channel remains idle, which means that all the routers within two hops from router A and associated with mini-slot 1 to $k-1$ have no packet to transmit), router A sends a jamming signal at mini-slot k. By adjusting the transmission power of the jamming signals and the receivers' sensitivity, we can ensure that all the routers within two hops from router A hear the jamming signal[4]. Consequently, all of the one-hop and two-hop neighbors of router A will not transmit at the current slot to avoid corrupting router A's transmission. The router which sends a jamming signal at the control part will transmit its packets at the transmission part of the same slot.

Note that in our scheme, the time slots are dynamically allocated to each router according to the traffic load. If a router has packet(s) to transmit, it needs to send a jamming signal in its own mini-slot; otherwise, it keeps silent in its own mini-slot. A router can transmit packets in a time slot when all the mini-slots prior to its own mini-slot in that slot are idle. When we have unbalanced traffic load among routers, the routers with light traffic load may not always have packets to transmit. Therefore, they may keep silent in their mini-slots, and leave the chance for the routers with heavy traffic load to transmit. As a result, the routers with heavy traffic load will be allocated more time slots.

B. Mini-Slot Assignment

The mini-slot assignment has the following requirements: 1) Any two routers which are within the two-hop neighborhood of each other should not be assigned the same mini-slot; 2) A minimum number of mini-slots should be assigned. In other words, the number of mini-slots cannot be reduced without violating requirement 1). The first requirement is to ensure that the routers which send jamming signals at the same mini-slot can transmit simultaneously without interfering with each other. The second requirement is to reduce the control overhead as much as possible. A mini-slot assignment algorithm which satisfies these two requirements is proposed in the following. Note that our scheme is not restricted to the proposed mini-slot assignment algorithm. Other methods (e.g., graph coloring) [14], [15] may also be used for the mini-slot assignment. Since the routers are located at fixed sites, the mini-slot assignment can be determined based on the whole network topology at the initialization of the network.

The overhead of the proposed scheme is dependent on the maximal number of routers in a two-hop neighborhood but not the total number of routers in the network, making the proposed scheme scalable for large networks. Since the overhead caused by mini-slots in our scheme is much smaller than that caused by the backoff and RTS/CTS control message exchanging in contention-based schemes, the control overhead in the proposed scheme is expected to be greatly reduced.

Algorithm 1 Mini-Slot Assignment

$N_m = 1$; //N_m denotes the number of mini-slots. At the beginning of the algorithm, it is set to 1.
$S = \{$all the routers in the networks$\}$, $S_1 = NULL$; //S_i denotes the set of routers which are assigned mini-slot i.
while $S \neq NULL$ **do**
    Randomly choose a router (denoted by A) from S
    assign_flag = FALSE
    for $i = 1, .., N_m$ **do**
        if none of one-hop and two-hop neighbors of router A belongs to S_i **then**
            Assign mini-slot i to router A, and add router A into S_i;
            Delete router A from S;
            assign_flag = TRUE;
            break;
        end if
    end for
    if assign_flag = FALSE **then**
        $N_m = N_m + 1$;
        Assign mini-slot N_m to router A, $S_{N_m} = \{$A$\}$;
        Delete router A from S;
    end if
end while

C. Maximal Spatial Reuse

The proposed scheme can achieve maximal spatial reuse. By maximal spatial reuse we mean that the set of routers which transmit simultaneously (without interfering with each other) in each slot is a maximal set. That is, there does not exist any router which does not belong to this set but can transmit simultaneously (without interfering with each other) with all the routers in the set.

Proof: Consider a slot T. Let *S* denote the set of routers which transmit at slot T. Suppose there exists one router A which does not belong to *S* (i.e., does not transmit at slot T), and whose potential transmission at slot T does not interfere with the transmissions of all the routers in *S*. Router A does not

[4] Here we consider a good propagation environment. When router A sends a jamming signal, it is possible that some of its two-hop neighbors may not hear the jamming signal if there are obstacles in between. In this case, we let each router send jamming signals to its one-hop neighbors (with lower power), and split one mini-slot into two parts. In the first part, router A sends a jamming signal to its one-hop neighbors. Upon hearing the jamming signal, all its one-hop neighbors relay the jamming signal in the second part. Therefore, all the two-hop neighbors of router A can hear the jamming signal.

transmit at slot T means that router A hears the jamming signal from one router (say B) within its two-hop neighborhood. Thus, router B must be in *S*. Since router B is within the two-hop neighborhood of router A, a collision can happen if both A and B transmit to a same neighbor. This conflicts with the supposition. ■

D. Per-router Fairness and Per-flow Fairness

There are various measurements for fairness. Here we consider two fairness models: per-router fairness and per-flow fairness. In the per-router fairness model, all the routers have fair channel access opportunities independent of the number of micro-flows delivered by the routers. Thus, the flows may have different throughput, depending on the traffic load of the associated routers. In the per-flow fairness model, when any two routers (which may relay different numbers of flows) contend with each other, all the flows[5] relayed by the two routers have fair channel access opportunities. Thus, a heavy-load router should have more chances to access the channel than a router with light load.

First, we consider how to achieve per-router fairness. From subsection III-A, it is obvious that the opportunity that one router may transmit in a slot largely depends on its mini-slot index in that slot. The smaller the index, the larger the opportunity. In order to fairly allocate the slots to each router, we have an initial mini-slot assignment (pre-determined at the initialization of the network), and rotate the order of the mini-slots slot by slot (i.e., the first mini-slot in the current slot becomes the last one in the next slot, the second mini-slot in the current slot becomes the first one in the next slot, and so on). It is possible that some routers may have less neighbors than others, i.e., the number of neighbors (within two-hop vicinity) of a router may be less than the number of mini-slots. In this case, just rotating the mini-slots may not ensure fair channel access for each router. Consider an example that a router (denoted by A) has 3 one-hop and two-hop neighbors B, C, and D, while the number of mini-slots is 6. A possible mini-slot assignment is shown in Fig. 2-(b). Accordingly, when we rotate the mini-slots, router A gets more chances to access the channel, benefiting from the two idle mini-slots. To solve this problem, we do not use a fixed mini-slot assignment. After a certain period, the order of the mini-slots is re-arranged (e.g., router D is assigned the first mini-slot and router A is assigned the 4th mini-slot), and each router rotates the mini-slots based on the new mini-slot assignment. All the mini-slot assignments are pre-determined and known by all the routers.

Per-flow fairness is achieved based on per-router fairness. Each router needs to exchange the information (i.e., the number of flows relayed by each router) with its one-hop and two-hop neighbors. According to the information, each router determines the fraction of time that it accesses the channel. Then each router adjusts its channel contending behavior accordingly. Consider an example that three routers (A, B, and C) contend with each other, while router A has 1 flow, router B has 2 flows, and router C has 3 flows. According to per-flow fairness, the fractions of channel time allocated to routers A, B, and C are $1/6$, $2/6$, and $3/6$, respectively. With per-router fairness, all the fractions of channel access time of the three routers are $1/3$. For router A, to reduce its time fraction from $1/3$ to $1/6$, it gives up half of its transmission opportunities. Thus, every two times when router A gets a turn to send a jamming signal at mini-slot 1, it gives up sending the jamming signal one time. On the contrary, to increase the time fraction of router C from $1/3$ to $3/6$, router C takes advantage of the transmission chances given up by router A. Router C can send its jamming signal at its own mini-slot upon hearing an idle channel during all the prior mini-slots. For router B, it neither gives up its own transmission opportunities nor takes the chances from others[6], thus maintains the same time fraction as that in per-router fairness.

As compared with per-router fairness, per-flow fairness incurs extra overhead to exchange messages among neighbors, but can allocate resource more fairly to flows. Therefore, there is a tradeoff between flow fairness and resource utilization. As the information exchange is necessary at the flow level instead of the packet level, the overhead is not expected to be significant.

E. Guaranteed Priority Access for Real-time Traffic

Since real-time traffic usually has a strict delay requirement, guaranteed priority access for real-time traffic is necessary in order to provide QoS satisfaction for real-time traffic. Hence, we add an additional mini-slot prior to all the other mini-slots. This extra mini-slot (referred to as real-time mini-slot) is dedicated for real-time traffic and is not rotated. For a router with real-time packet(s) to transmit, in addition to sending a jamming signal in its own mini-slot, it first sends a jamming signal in this real-time mini-slot. Upon hearing the jamming signal in this mini-slot, the routers which have only data packets will not send their own jamming signals, leaving the chance to the router with real-time traffic to send a jamming signal in its corresponding mini-slot. When two or more real-time routers contend for the same slot, the one with the smallest mini-slot index will first send the jamming signal and get the slot.

In order to provide further priority differentiation to real-time packets with different delay requirements, we can have an additional number of real-time mini-slots. For real-time mini-slot i $(i = 1, 2, \ldots)$, a corresponding urgency level U_i $(U_1 < U_2 < U_3 \ldots)$ is pre-defined. The smaller the U_i, the more urgent the level is. The urgency of a real-time packet is measured by the packet due time[7] and the remaining hops to the destination. The packet is more urgent if the due time is smaller and the number of the remaining hops is larger. Considering a router with a real-time packet j having the remaining time t_j to the due time and the remaining hops n_j to the destination, if $U_{i-1} < \frac{t_j}{n_j} \leq U_i$ (where $U_0 = 0$), then the router sends a jamming signal at real-time mini-slot

[5]Note that the flow here is not referred to as the end-to-end multi-hop flow, but the one-hop sub-flow from the relay router to the next hop.

[6]When the mini-slot of router B is not the last one, after hearing an idle channel during all the prior mini-slots, it does not send its jamming signal and leave the chance to router C. However, if the mini-slot of router B is the last one, it will transmit at the current slot to achieve spatial reuse.

[7]The due time of a real-time packet is the packet generation time plus the packet delay bound. We assume that this information is included in the packet header and known by the traversed routers.

i if all the prior real-time mini-slots are idle. Once a router hears the jamming signal (which means that another one-hop or two-hop-away router has a more urgent real-time packet), it will quit the contention for the current slot.

F. Congestion Avoidance

In the wireless backbone, it is very likely that some routers (referred to as bottleneck routers) located at the center of the network or near the gateway need to relay more traffic than other routers. In the case of per-router fairness, with an absolute fair channel access for each router, the traffic arrival rate will be higher than the traffic departure rate at the bottleneck routers. As a result, the packets will be accumulated, eventually causing buffer overflow at the bottleneck routers. It is possible that multi-hop data flows pass through bottleneck routers. Buffer overflow at the routers results in resource waste and low end-to-end throughput. Transmission control protocol (TCP) is the most popular protocol to deal with network congestion at the transport layer. However, TCP suffers from severe performance degradation in wireless networks, due to the fact that it is difficult for the source nodes at the transport layer to know explicitly whether a packet loss is due to buffer overflow or temporary link failure [16]. In order to avoid congestion effectively in the mesh backbone, we propose a straightforward mechanism at the MAC layer. Each router keeps track of its packet arrivals and departures. For each one-hop neighbor, the router records the number of arrived packets (denoted by C_a) and departed packets (denoted by C_d) which are from that neighbor. If the difference between C_a and C_d is larger than a pre-defined threshold[8], the router sends a message to the neighbor to suspend its transmissions to this router. When the difference between C_a and C_d decreases to a certain value, the router sends a message to resume the transmissions. This approach avoids buffer overflow at intermediate hops of multi-hop flows. The control propagates hop by hop to the source node and regulates the source rate depending on the network congestion status.

G. Operation Procedure of the Proposed MAC Scheme

To better explain the proposed MAC scheme, we use a simple example to illustrate the operation procedure. A chain topology with six routers is considered, as shown in Fig. 3-(a). At the initialization of the network, the mini-slot assignment algorithm is performed by a router, and the result is broadcast to all the routers. The mini-slot assignment result is shown in Fig. 3-(b), where routers A and D are assigned mini-slot 1, routers B and E are assigned mini-slot 2, routers C and F are assigned mini-slot 3, and mini-slot 0 is reserved for real-time traffic. Suppose at the beginning, all the six routers have data packets to transmit. For slot 1, since no router has real-time packet, mini-slot 0 will remain idle. After hearing an

[8]From the simulations, we observe that our scheme is not sensitive to the value of the threshold. Varying the threshold in a large range results in only slight differences in the performance of relay efficiency. Therefore, the choice of this threshold is not a critical issue for the proposed scheme as long as it does not incur buffer overflow. For a router with a large buffer size, a large threshold can be chosen so that a small control overhead can be achieved. On the contrary, for a router with a small buffer size, a small threshold should be chosen to ensure that buffer overflow does not happen.

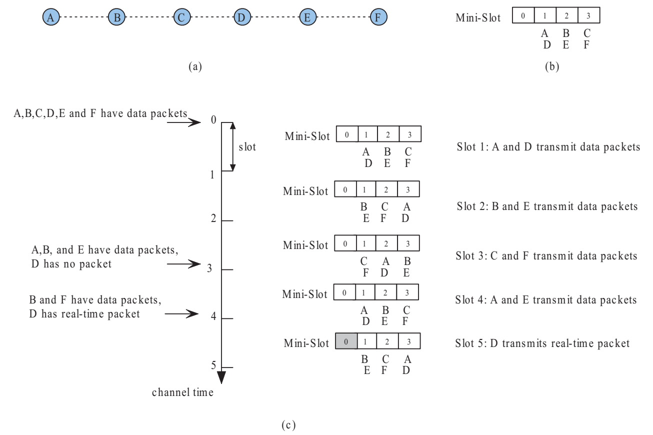

Fig. 3. An example to illustrate the operation procedure of the proposed MAC scheme.

idle channel at mini-slot 0, routers A and D will send jamming signals at mini-slot 1. Consequently, routers B, C, E, and F will hear a busy channel at mini-slot 1, thus defer their own transmissions at slot 1. Therefore, routers A and D transmit their data packets at slot 1 without worrying about collisions. At the end of each slot, all the routers will rotate their mini-slot indices, and the corresponding results are shown in Fig. 3-(c). For slot 2, routers B and E are associated with mini-slot 1, and for slot 3, routers C and F are associated with mini-slot 1. Similarly, routers B and E will transmit their data packets at slot 2, and routers C and F will transmit their data packets at slot 3. At the end of slot 3, suppose routers A, B, and E have data packets and router D has no packet. For slot 4, mini-slot 1 is assigned to routers A and D again. As in the case of slot 1, router A will send jamming signal at mini-slot 1. As router D has no packet to transmit, it keeps silent at mini-slot 1. Consequently, router E will hear an idle channel at both mini-slots 0 and 1. Therefore, it will send its own jamming signal at mini-slot 2. As a result, routers A and E will transmit data packets at slot 4. Suppose at the end of slot 4, routers B and F have data packets and router D has a real-time packet. For slot 5, router D will first send jamming signal at mini-slot 0, and this jamming signal can be heard by both router B and router F. Thus, routers B and F will not send their own jamming signals. After hearing an idle channel at mini-slots 1 and 2, router D will send jamming signal again at mini-slot 3. As a result, router D transmits the real-time packet at slot 5.

IV. Performance Analysis

To make the analysis tractable, we consider a simplified case that 1) there is one real-time mini-slot, and all the real-time packets are treated equally; 2) per-router fairness is considered. We assume that the voice and video call arrivals at each source node are independent and follow a Poisson process, and the call duration has an exponential distribution.

A. Real-time Traffic Access Delay Bound

The access delay is defined as the time period from the instant that a packet becomes the head in the buffer to the

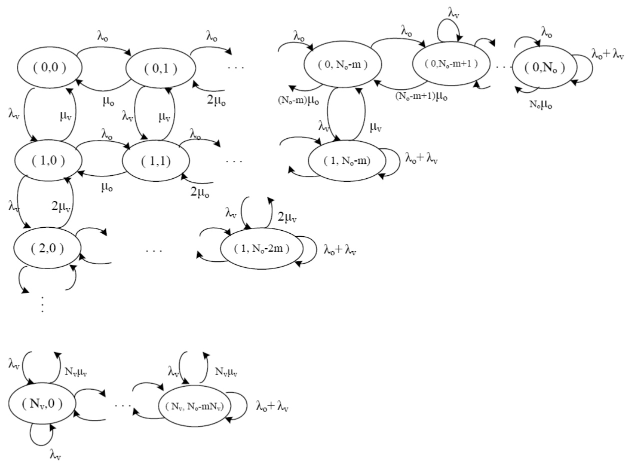

Fig. 4. The state transition diagram.

instant that the packet departs from the router. Let T_s denote the time duration of one slot, and N the number of mini-slots, including the real-time mini-slot. Consider the worst case that the target router has $N-2$ one-hop and two-hop neighbors, and all of them have real-time packets to transmit. After the target router transmits one packet, it takes $T_s(N-1)$ for the target router to transmit the next one. Thus, the access delay bound of real-time traffic at each hop is $T_s(N-1)$, which is independent to the traffic load of the networks.

B. Data Traffic Access Delay

Since guaranteed priority access is provided to real-time traffic, the real-time traffic load will impact the data traffic access delay. The voice call is represented by an `on/off` model [17]. The durations of the `on` and `off` states are independently and exponentially distributed with parameters α and β, respectively. At an `on` state, voice packets are generated periodically with an inter-arrival time I_o, while no voice packet is generated at an `off` state. For a video call, the video frames are generated periodically with an inter-arrival time I_v. The video frame usually has a large and variable size [18]. Suppose that it takes one slot to transmit one voice packet, and M slots (on average) to transmit one video frame. Considering a target router, we refer to its two-hop vicinity as the target area.

To obtain the data traffic access delay, we first need to derive the fraction of channel time occupied by voice and video traffic. We define a two-dimensional state $(n_v,\ n_o)$, where n_v and n_o are the numbers of video calls and voice calls, respectively, being served by the routers within the target area. Denote the average arrival rates of voice and video calls that traverse the target area as λ_o and λ_v, respectively, and the average call duration as μ_o^{-1} and μ_v^{-1}, respectively. We assume that call admission control is in place to guarantee the QoS of voice and video calls, and the maximal number of acceptable voice and video calls within the target area are denoted by N_o and N_v, respectively. The state transition diagram is shown in Fig. 4. Since a video call requires more resources than a voice call, when there is 1 video call being served, the maximal number of supported voice calls is $N_o - M\frac{I_o}{I_v}$, denoted by $N_o - m$. Define $p_{i,j}$ as the joint probability that i video calls and j voice calls being served. The balance equations for the two-dimensional state space of Fig. 4 are given in (1) at the top of next page.

For the case of $i = N_v$, we need to consider three possibilities: $N_o - mN_v > 1$, $N_o - mN_v = 1$, and $N_o - mN_v = 0$. When $N_o - mN_v > 1$, the corresponding balance equations are

$$\begin{array}{ll}
(\lambda_o + i\mu_v)p_{i,j} = \lambda_v p_{i-1,j} + \mu_o p_{i,j+1} & i = N_v, j = 0; \\
(\lambda_o + i\mu_v + j\mu_o)p_{i,j} = \lambda_v p_{i-1,j} & i = N_v, \\
\quad +\lambda_o p_{i,j-1} + (j+1)\mu_o p_{i,j+1} & 1 \le j \le N_o - mN_v - 1; \\
(i\mu_v + j\mu_o)p_{i,j} = \lambda_v p_{i-1,j} + \lambda_o p_{i,j-1} & i = N_v, j = N_o - mN_v.
\end{array}$$

When $N_o - mN_v = 1$, the corresponding balance equations are

$$\begin{array}{ll}
(\lambda_o + i\mu_v)p_{i,j} = \lambda_v p_{i-1,j} + \mu_o p_{i,j+1} & i = N_v, j = 0; \\
(i\mu_v + \mu_o)p_{i,j} = \lambda_v p_{i-1,j} + \lambda_o p_{i,j-1} & i = N_v, j = 1.
\end{array}$$

When $N_o - mN_v = 0$, the corresponding balance equations is

$$i\mu_v p_{i,0} = \lambda_v p_{i-1,0} \qquad i = N_v, j = 0.$$

Based on the above balance equations, the probability distribution of state $(n_v,\ n_o)$ can be derived. A voice/video call may traverse several hops within the target area. Let h_n^o and h_n^v denote the average number of hops that voice and video calls traverse the target area, respectively. As voice traffic only generates packets during an `on` period, at any time instant, each voice call is at the `on` state with probability $\beta/(\alpha+\beta)$. During I_o (i.e., the voice packet inter-arrival duration), each voice call which is at the `on` state generates one voice packet. Thus, given n_o voice calls being served in the target area, the average channel time occupied by these n_o voice calls during I_o is given by

$$\overline{T_o}(n_o) = \sum_{i=1}^{n_o} \binom{n_o}{i} (\frac{\beta}{\alpha+\beta})^i (\frac{\alpha}{\alpha+\beta})^{n_o - i} i T_s h_n^o, \quad 0 \le i \le n_o. \tag{2}$$

For a video call with the frame inter-arrival duration I_v, the average number of video frames that a video call generates during I_o is I_o/I_v. Thus, given n_v video calls, the average channel time occupied by these n_v video call during I_o is given by

$$\overline{T_v}(n_v) = \frac{I_o}{I_v} n_v M T_s h_n^v. \tag{3}$$

Thus, the fraction of channel time occupied by real-time traffic is given by

$$f = \sum_{\text{all state } (n_v, n_o)} \frac{(\overline{T_o}(n_o) + \overline{T_v}(n_v))p_{n_v,n_o}}{I_o}. \tag{4}$$

In our scheme, the residual channel time left by real-time traffic is fairly shared by all the routers with data traffic. For data traffic access delay, we consider two cases: saturated case and unsaturated case. In the saturated case, all the routers with data traffic always have data packets to transmit. In the unsaturated case, the average date packet arrival rate at each router is denoted as λ_d. First consider the saturated case. Given an arbitrary time slot, the probability that the target router can transmit its data packet in that slot is given by

$$\begin{array}{ll}
(\lambda_o+\lambda_v)p_{0,0}=\mu_o p_{0,1}+\mu_v p_{1,0} & i=0, j=0; \\
(\lambda_o+\lambda_v+j\mu_o)p_{0,j}=\lambda_o p_{0,j-1}+(j+1)\mu_o p_{0,j+1}+\mu_v p_{1,j} & i=0, 1\le j\le N_o-m; \\
(\lambda_o+j\mu_o)p_{0,j}=\lambda_o p_{0,j-1}+(j+1)\mu_o p_{0,j+1} & i=0, N_o-m+1\le j\le N_o-1; \\
N_o\mu_o p_{0,N_o}=\lambda_o p_{0,N_o-1} & i=0, j=N_o; \\
(\lambda_o+\lambda_v+i\mu_v)p_{i,0}=\lambda_v p_{i-1,0}+(i+1)\mu_v p_{i+1,0}+\mu_o p_{i,1} & 1\le i\le N_v-1, j=0; \\
(\lambda_o+\lambda_v+i\mu_v+j\mu_o)p_{i,j}=\lambda_v p_{i-1,j}+\lambda_o p_{i,j-1}+(j+1)\mu_o p_{i,j+1}+(i+1)\mu_v p_{i+1,j} & 1\le i\le N_v-1, 1\le j\le N_o-m(i+1); \\
(\lambda_o+i\mu_v+j\mu_o)p_{i,j}=\lambda_v p_{i-1,j}+\lambda_o p_{i,j-1}+(j+1)\mu_o p_{i,j+1} & 1\le i\le N_v-1, \\
 & N_o-m(i+1)+1\le j\le N_o-im-1; \\
(j\mu_o+i\mu_v)p_{i,j}=\lambda_v p_{i-1,j}+\lambda_o p_{i,j-1} & 1\le i\le N_v-1, j=N_o-im.
\end{array} \tag{1}$$

$$p=(1-f)\frac{1}{K} \tag{5}$$

where f is given by (4), and K is the number of data routers within the target area that fairly share the residual channel time left by the real-time traffic. Thus, the data traffic access delay of the target router is $\frac{T_s}{p}$. For the unsaturated case, denote the data traffic access delay as d_a. The data packet arrivals and departures at each router can be considered as a queue, and the queue utilization is $\rho=\lambda_d d_a$ $(\rho\le 1)$. For each router with data traffic, at an arbitrary time, the router has data packet(s) to transmit with probability ρ and has no data packet to transmit with probability $1-\rho$. Thus, (5) can be re-written as

$$p=(1-f)\sum_{i=0}^{K-1}\binom{K-1}{i}\rho^{K-1-i}(1-\rho)^i\frac{1}{K-i}, \quad \rho\le 1. \tag{6}$$

Substituting $p=\frac{T_s}{d_a}$ and $\rho=\lambda_d d_a$ in (6), we can obtain d_a. Note that when $\rho=1$ (i.e., the saturated case), (6) is equivalent to (5).

C. Numerical Results

Simulations are carried out in order to verify the accuracy of the analysis. Since the analysis of real-time access delay bound is straightforward, here we validate the analysis of data traffic access delay. Without loss of generality, we choose the parameters $T_s=0.2$ ms, $K=10$, $N_o=40$, $N_v=5$, $h_n^o=3$, and $h_n^v=3$. The voice packet and video frame inter-arrival durations are 20 ms and 100 ms, respectively. One video frame takes 40 slots (on average) to transmit. The average voice and video call durations are $150s$ and $600s$, respectively. For a voice call, the average `on` and `off` durations are 352 ms and 650 ms [17], respectively.

First, consider the saturated case. We fix the voice call arrival rate λ_o as 0.1 call/s, and vary the video call arrival rate λ_v from 0.01 call/s to 0.1 call/s. Table I compares the simulation and analytical results of the data traffic access delay. They agree with each other well.

Second, consider the unsaturated case. We fix λ_o and λ_v as 0.1 and 0.05 call/s, respectively, and vary the average date packet arrival rate λ_d from 30 to 60 packet/s. Fig. 5 shows the data traffic access delay. Note that when $\lambda_d=60$ packet/s, ρ equals to 1 (i.e., $\lambda_d d_a=1$), the data routers become saturated. The data access delay increases sharply when the system approaches the saturated case. It is clear that the simulation results match well with the analytical results.

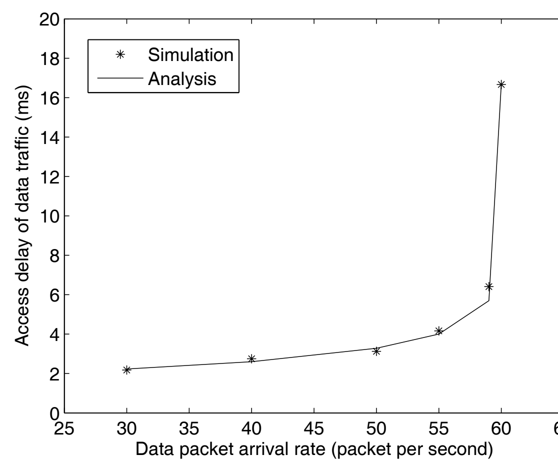

Fig. 5. The data traffic access delay with different data packet arrival rates.

V. PERFORMANCE EVALUATION

We evaluate the performance of the proposed scheme by extensive simulations. As no representative distributed MAC scheme for wireless mesh networks is available for comparison so far, we compare the proposed scheme with IEEE 802.11 and distributed packet reservation multiple access (DPRMA) [19]. IEEE 802.11 is the most popular distributed MAC scheme for WLANs. DPRMA is proposed for multi-hop wireless ad hoc networks, and supports both voice and data traffic. Similar to the proposed MAC scheme, DPRMA is also a slotted-channel based distributed MAC scheme. For voice traffic, we choose the GSM 6.10 codec as an example. The voice packet size is 109 bytes with 33-byte payload and 76-byte RTP/UDP/IP and MAC headers. The voice packet inter-arrival period is 20 ms. For video traffic, we choose the H.264 codec [20], which is the most efficient video compression technology and is widely implemented. The H.264 defines a set of profiles with different video bit rates for various classes of applications. Here, we use H.264 with video bit rate of 384 kbps. The frame rate is 30 frame/s. For data traffic, the data packet arrivals follow a Poisson process with various arrival rates. Other simulation parameters are chosen according to the IEEE 802.11g/e[9] standards [21], [22], given in Table II, where the channel rate is to transmit voice/video/data packets, and the basic rate is to transmit RTS and CTS (in IEEE 802.11). For

[9]IEEE 802.11g includes two slot times. One is short slot time (i.e., 9 μs) for a 802.11g-only network, and the other is long slot time (i.e., 20 μs) for a mixed-mode 802.11b/g network. Here we use short slot time in the simulation.

TABLE I
THE AVERAGE DATA TRAFFIC ACCESS DELAY (MS) WITH DIFFERENT λ_v (CALL/S) WHILE $\lambda_o = 0.1$ CALL/S

video call arrival rate λ_v		0.01	0.025	0.05	0.075	0.1
data access delay	simulation	7.99	12.47	16.76	20.56	26.55
	analysis	7.09	11.42	16.67	21.62	26.91

TABLE II
SYSTEM PARAMETERS USED IN SIMULATION.

Parameter	Value
AIFS[AC_data]	37 μs
AIFS[AC_voice]	28 μs
AIFS[AC_video]	28 μs
SIFS	10 μs
$CW_{\min}$[voice]	7
$CW_{\max}$[voice]	15
$CW_{\min}$[video]	15
$CW_{\max}$[video]	31
$CW_{\min}$[data]	31
$CW_{\min}$[data]	1023
PHY preamble	20 μs
RTS frame size	20 bytes
CTS frame size	14 bytes
data packet size	1000 bytes
channel rate	54 Mbps
basic rate	24 Mbps
Mini-Slot time (Proposed)/Slot time (IEEE 802.11g)	9 μs
Frame time (DPRMA)	20 ms
Contention Probability (DPRMA)	0.6

fair comparison, we choose the time duration of one mini-slot in our scheme the same as the smallest time unit in IEEE 802.11. Our scheme and DPRMA have the same preamble overhead to transmit one voice/video/data packet as that in IEEE 802.11.

A. The Delay Performance for Real-time Traffic

We consider the case that there is one real-time mini-slot in our scheme. A chain topology as shown in Fig. 6-(a) is considered. We first compare the video packet delay performance of our scheme with IEEE 802.11. There are two video flows, flow 1 having 4 hops from router 1 to the gateway, and flow 2 having one hop from router 4 to the gateway. Each flow is an aggregated video flow, including 10 video calls. To demonstrate the performance of priority access for real-time traffic, we let the two video flow experience various contention degrees with data traffic. First, consider the case that there is no other data flow contending with these two video flows. Then we increase the number of data contenders N_{dc} near each router gradually from 1 to 5, each contender having a data flow with source rate 3 Mbps to the gateway.

Table III compares the packet delay of the two video flows. It can be seen that as the number of data contenders increases, the video packet delays increase in IEEE 802.11. Especially for flow 1 (with a relatively long path), the video packet delay increases significantly from 2.12 ms to 128.24 ms when the number of data contenders varies from 0 to 5. On the contrary, in our scheme, the video packet delays remain stable for all the numbers of data contenders. These results demonstrate that in IEEE 802.11, the delay performance of real-time traffic is degraded when the data traffic load increases. Especially for real-time flows with a long path, such performance degradation is significant. The reason is that IEEE 802.11 provides statistical priority access, where the prioritized access for real-time traffic is only guaranteed in a long term, but not for every contention[10]. Such a statistical priority access is difficult to satisfy the delay requirement of real-time traffic since the real-time traffic may suffer from performance degradation due to a high data traffic load [9]. Video flows with a long

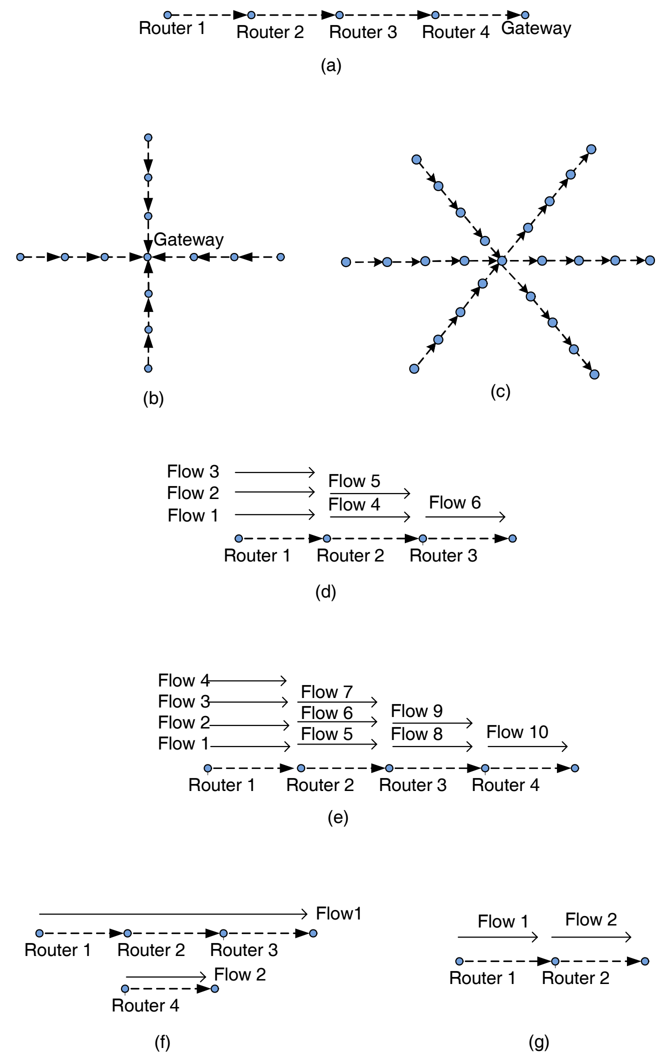

Fig. 6. The simulation topologies.

[10]In IEEE 802.11, since each node continues to count down its backoff timer once the channel becomes idle for a certain period, a data packet with a probably large initial backoff timer will eventually count down its backoff timer to a small value, most likely smaller than the backoff timer of a new backlogged real-time packet. Then the data packet gains the channel, resulting in the real-time packet waiting for a long time for the next competition [23].

TABLE III

THE AVERAGE REAL-TIME PACKET DELAY (MS) WITH DIFFERENT NUMBER OF DATA CONTENDERS NEAR EACH ROUTER

N_{dc}		0	1	2	3	4	5
802.11 (video)	flow 1	2.12	4.01	26.02	63.96	93.55	128.24
	flow 2	0.44	0.75	1.88	1.96	2.73	4.40
Proposed (video)	flow 1	0.67	0.67	0.67	0.68	0.67	0.68
	flow 2	0.21	0.21	0.21	0.21	0.21	0.21

N_{dc}		0	1	2	3	4	5
DPRMA (voice)	flow 1	1.89	2.37	2.61	24.97	38.23	54.21
	flow 2	0.47	0.71	0.95	1.18	1.66	4.97
Proposed (voice)	flow 1	0.63	0.63	0.63	0.63	0.63	0.63
	flow 2	0.21	0.21	0.21	0.21	0.21	0.21

path also suffer from the priority reversal problem[11] caused by hidden terminals, resulting in worse delay performance. Contrary to IEEE 802.11, our scheme can achieve guaranteed priority access for real-time traffic. In addition, as the routers within the two-hop neighborhood are not allowed to transmit simultaneously, hidden terminals do not exist, neither does the priority reversal problem. As a result, our scheme achieves a small delay for both long-path and short-path real-time flows regardless of the data traffic load.

As DPRMA does not support video traffic, we compare the voice packet delay performance of our scheme with DPRMA. The simulation scenario is similar to the preceding one except that the video flows are replaced with voice flows. Table III also compares the packet delay of the two voice flows. As expected, the voice packet delay remains stable in our scheme in all the cases. When the number of data contenders increases, DPRMA experiences a longer and longer delay to reserve a slot for voice traffic, resulting in a longer voice packet delay.

B. *Fairness and End-to-End Throughput of Data Flows*

For data flows, the QoS metrics of interest are fairness and end-to-end throughput. Here, we consider two scenarios: the chain topology in Fig. 6-(a) with 4 data flows, where flow i $(i = 1, \ldots, 4)$ is from router i $(i = 1, \ldots, 4)$ to the gateway; and the cross topology with 12 data flows in Fig. 6-(b). In the cross topology, the center node is the gateway, and each router has a data flow to the gateway.

For the chain topology, we vary the data packet arrival rate of each flow, and obtain the end-to-end throughput of the 4 flows, given in Table IV. The end-to-end throughput is measured by the data rate received at the gateway. For IEEE 802.11, all the flows have the same throughput when the traffic load is low. However, when the traffic load becomes high, the resources are not fairly allocated to each flow. The throughput of the flow with the shortest path (i.e., flow 4) is much larger than that of the flow with the longest path (i.e., flow 1). The unfairness is due to the hidden terminal problem [13]. On the contrary, the throughputs of the 4 flows in our scheme and DPRMA are almost the same under the varying traffic load, indicating that our scheme and DPRMA have improved fairness performance over IEEE 802.11. In addition, the aggregated throughput in our scheme is always larger than those in IEEE 802.11 and DPRMA. By avoiding collisions and reducing the control overhead, our scheme achieves higher resource utilization than IEEE 802.11 and DPRMA. Similar results are observed in the cross topology, as shown in Table IV. In contrast to IEEE 802.11 that the one-hop flows achieve much higher throughput than the two-hop and three-hop flows when the traffic load increases, in our scheme and DPRMA, all the flows have almost the same throughput in all the cases.

[11]Priority reversal problem is that a real-time node may lose its priority and seldom have a chance to transmit its packets when it contends with a data node, which is a hidden terminal of the real-time node [13].

TABLE IV

THE END-TO-END THROUGHPUTS (MBPS) OF DATA FLOWS IN THE CHAIN TOPOLOGY AND CROSS TOPOLOGY

The chain topology

Source rate of each flow (Mbps)		1	3	5	7	9
802.11	Flow 1	0.97	0.82	0.53	0.23	0.19
	Flow 2	0.97	0.98	0.59	0.54	0.26
	Flow 3	0.97	2.83	2.59	2.48	2.14
	Flow 4	0.97	2.86	4.42	5.91	6.78
	Aggregate	3.88	7.49	7.83	9.16	9.37
Proposed	Flow 1	0.98	2.98	4.40	4.42	4.42
	Flow 2	0.99	2.99	4.41	4.42	4.43
	Flow 3	0.99	2.99	4.42	4.43	4.42
	Flow 4	0.98	2.98	4.44	4.44	4.44
	Aggregate	3.94	11.94	17.67	17.71	17.71
DPRMA	Flow 1	0.99	2.96	3.24	3.34	3.35
	Flow 2	0.98	2.96	3.25	3.36	3.37
	Flow 3	0.99	2.97	3.25	3.38	3.40
	Flow 4	0.99	2.98	3.28	3.41	3.98
	Aggregate	3.95	11.87	13.02	13.49	14.10

The cross topology

Source rate of each flow (Mbps)		0.5	1	2	3	4
802.11	One-hop flow	0.49	0.95	1.63	1.92	1.49
	Two-hop flow	0.49	0.80	0.54	0.39	0.30
	Three-hop flow	0.48	0.79	0.53	0.38	0.29
	Aggregate	5.84	10.16	10.80	10.76	8.32
Proposed	One-hop flow	0.49	0.99	1.99	2.98	3.06
	Two-hop flow	0.49	0.99	1.99	2.98	3.02
	Three-hop flow	0.49	0.99	1.99	2.98	3.03
	Aggregate	5.88	11.88	23.88	35.76	36.44
DPRMA	One-hop flow	0.49	0.97	1.05	1.15	1.20
	Two-hop flow	0.49	0.97	1.05	1.15	1.20
	Three-hop flow	0.49	0.96	1.04	1.14	1.20
	Aggregate	5.88	11.60	12.56	13.76	14.40

C. *Relay Efficiency*

As mentioned earlier, if a MAC scheme is designed without considering congestion avoidance, it is very likely that the

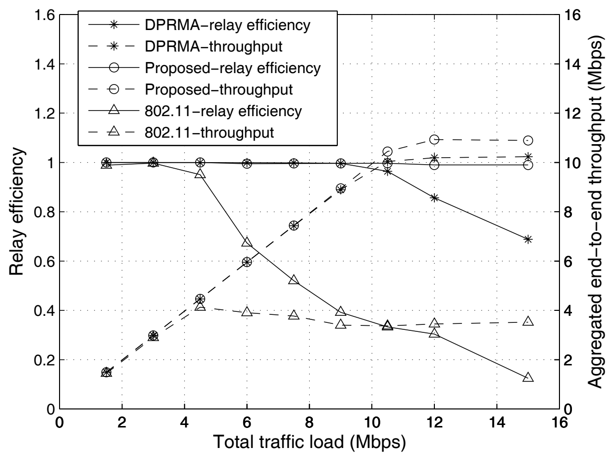

Fig. 7. The relay efficiency and aggregated end-to-end throughput in the cross topology.

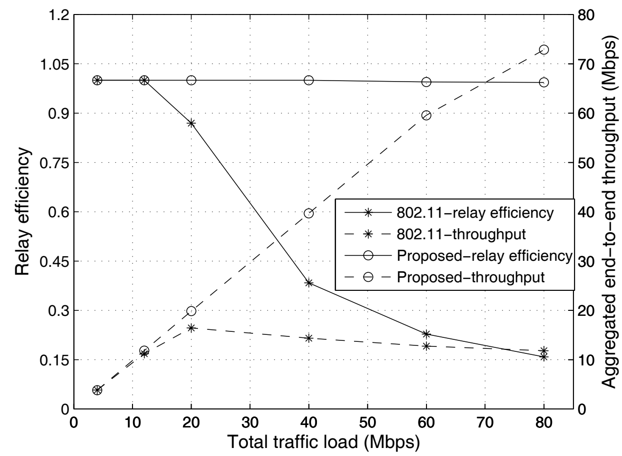

Fig. 8. The relay efficiency and aggregated end-to-end throughput of the proposed scheme and IEEE 802.11 in the random topology.

source nodes may inject more packets than what the bottleneck routers can forward. As a result, some packets sent by the source nodes are dropped by the bottleneck routers due to buffer overflow, leading to a waste of wireless channel and power resources. Relay efficiency is defined as the ratio of the sum of the packets received at all the destinations to the sum of the packets sent by all the sources. This metric reflects how much the resources are wasted in relaying. The smaller the relay efficiency, the more the resources are wasted at the bottleneck routers. The scenario of three data flows as shown in Fig. 6-(c) is considered, where the router at the center is a bottleneck router, relaying all the data flows.

Fig. 7 compares the relay efficiency of the three schemes. It is clear that our scheme achieves close to 100% relay efficiency in all the cases, while IEEE 802.11 and DPRMA have a decreased efficiency when the traffic load increases. The result implies that, without congestion control, IEEE 802.11 and DPRMA drop more and more packets at the bottleneck router, as the traffic load increases. For comparison, the aggregated end-to-end throughputs of the three schemes are also shown in Fig. 7. By avoding packet dropping at the bottleneck router, our scheme utilizes the resources more efficiently and achieves a higher end-to-end throughput. Notice that, in our scheme, the aggregated packet sending rate at all the source nodes is bounded at about 11 Mbps[12] although the total traffic load (i.e., the aggregated packet arrival rate at all the source nodes) may be higher than 11 Mbps. This result indicates that our scheme effectively controls the rate that a source injects packets to the network, so that network congestion can be avoided.

D. Performance in Random Topology

To evaluate the performance of the propose scheme in a more general case, we consider a random topology. In the simulations, 100 routers are uniformly placed deterministically in a 1000 m $\times$ 1000 m area. Two routers are one-hop neighbors if the distance between them is less than or equal to 100 m. For comparison with IEEE 802.11, fifty flows are considered, with 5 voice flows, 5 video flows, and 40 data flows. The source and destination of each flow are randomly selected, and a shortest path from the source to the destination is pre-determined, so that each intermediate router knows its up-stream and down-stream routers. We vary the traffic load of each data flow from 0.1 Mbps to 2 Mbps and observe that the end-to-end delays of one randomly picked voice and video flow remain unchanged at 0.63 ms and 1.46 ms, respectively, over the data traffic load range. This result confirms again that our scheme provides guaranteed priority access to real-time traffic regardless of data traffic load.

Fig. 8 compares the relay efficiency and aggregated end-to-end throughput of our scheme and IEEE 802.11. We can see that the relay efficiency of our scheme is almost 100% under all the cases of data traffic load. However, in IEEE 802.11, the relay efficiency drops rapidly when the data traffic load increases. Fig. 8 also compares the aggregated end-to-end throughput of the 40 data flows in IEEE 802.11 and in our scheme. When the traffic load increases, our scheme achieves a much higher end-to-end throughput than IEEE 802.11.

To investigate the fairness performance, here we use the commonly used Jain's Fairness Index given by $\frac{(\sum_{i=1}^{n} x_i)^2}{n\sum_{i=1}^{n} x_i^2}$ [24], where x_i is the throughput of the ith data flow, and n is the number of data flows. Jain's Fairness Index has the value from 0 to 1. The higher the value, the better the fairness performance. Fig. 9 compares the Fairness Index values of our scheme and IEEE 802.11. When the total traffic load is low (i.e., less than 12 Mbps), the Fairness Index values of both schemes are 1, while with the increase of traffic load, our scheme achieves much better fairness performance than IEEE 802.11. Note that when the traffic load is at 80 Mbps, the Fairnesses Index value in our scheme is less than 1. Since flows are randomly chosen, some flows may experience more contentions than others. Due to the capacity limit, the flows with more contentions cannot further increase their throughput, while other flows with less contentions may still

[12]Since the relay efficiency is close to 100% in our scheme, the packet sending rate at the source nodes almost equals to the packet receiving rate at the destinations (i.e., the end-to-end throughput).

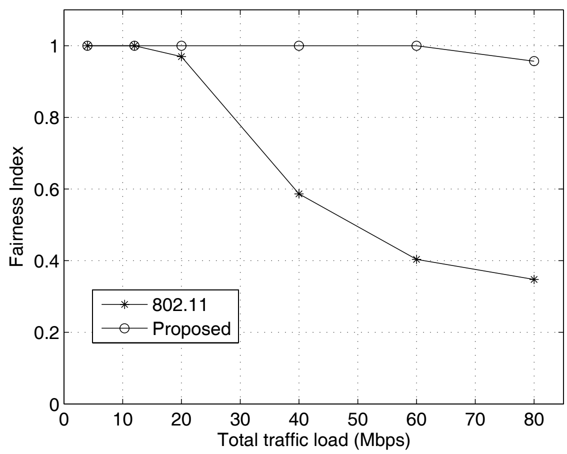

Fig. 9. The fairness index in the random topology.

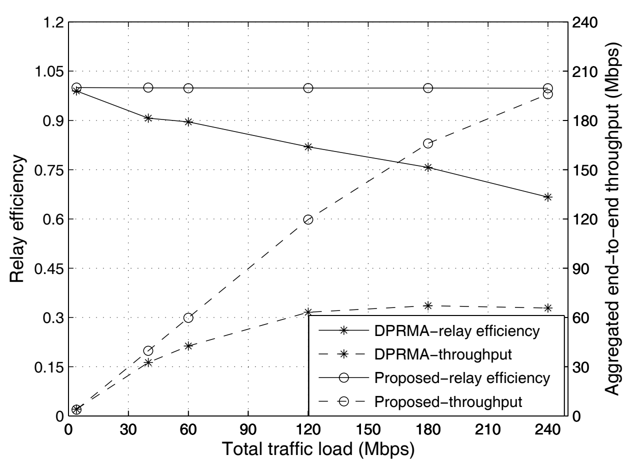

Fig. 10. The relay efficiency and aggregated end-to-end throughput of the proposed scheme and DPRMA in the random topology.

increase their throughput.

For comparison with DPRMA, we consider 50 randomly chosen flows with 10 voice flows and 40 data flows (as DPRMA does not support video traffic). Again, for a randomly picked voice flow, the end-to-end delay remains unchanged at 0.72 ms in our scheme, when the data traffic load increases. However, for DPRMA, the delay increases from 1.66 ms to 21.65 ms when the traffic load of each data flows varies from 0.1 Mbps to 6 Mbps. Fig. 10 compares the relay efficiency and aggregated end-to-end throughput of our scheme and DPRMA. Similar results are observed as in Fig. 8. As DPRMA and the proposed scheme have a similar fairness performance, the results are not listed in the paper.

From the simulations, we also observe that our scheme is not sensitive to the value of the threshold for congestion control as long as it is not very large. Varying the threshold from 5 packets to 50 packets results in only slight differences in the performance of relay efficiency.

TABLE V

THE THROUGHPUTS (MBPS) OF THE DATA FLOWS IN SCENARIOS SHOWN IN FIG 6-(D) AND (E)

Scenario (d)

Source rate of each flow (Mbps)		4	6	8	10	15
per-router fairness	Flow 1-3	3.99	5.98	5.29	5.00	4.43
	Flow 4-5	3.99	5.98	7.98	7.44	6.65
	Flow 6	3.99	5.98	7.99	9.99	13.29
per-flow fairness	Flow 1-3	3.99	5.98	6.64	6.64	6.64
	Flow 4-5	3.99	5.98	6.65	6.65	6.65
	Flow 6	3.99	5.98	6.65	6.65	6.65

Scenario (e)

Source rate of each flow (Mbps)		2	4	6	8	10
per-router fairness	Flow 1-4	1.98	3.98	3.34	3.32	3.32
	Flow 5-7	1.98	3.98	4.82	4.42	4.42
	Flow 8-9	1.98	3.98	5.98	6.64	6.64
	Flow 10	1.99	3.99	5.98	7.99	9.99
	Aggregate	19.81	39.81	45.76	47.81	49.81
per-flow fairness	Flow 1-4	1.98	3.96	3.97	3.98	3.98
	Flow 5-7	1.98	3.96	3.98	3.98	3.99
	Flow 8-9	1.98	3.97	3.99	3.99	3.99
	Flow 10	1.99	3.99	5.99	6.64	6.65
	Aggregate	19.81	35.69	41.79	42.48	42.52

E. The Comparison of Per-flow Fairness and Per-router Fairness

In the preceding simulations, one router generates one data flow and per-router fairness is considered. In the following, we consider the cases that different routers generate different numbers of data flows and compare the performance of per-router fairness and per-flow fairness. First, consider the scenario shown in Fig. 6-(d), where there are 6 data flows contending with each other. We vary the data packet arrival rate of each flow, and obtain the throughput of each flow under per-flow fairness and per-router fairness, respectively, given in Table V. It can be seen that, in the case of per-router fairness, all the flows have the same throughput when the traffic load is low. When the traffic load becomes high, flow 6 has the highest throughput, and flows 1 to 3 have the lowest throughput. In the case of per-flow fairness, all the flows have the same throughput in all the cases. Note that a perfect fairness can be achieved in our scheme without considering real-time traffic. However, when real-time traffic is considered, some data flows may have to give their transmission chances to nearby real-time traffic, and a perfect fairness may not be able to achieve.

Second, consider the scenario shown in Fig. 6-(e), which has 10 data flows. Similar observation can be made from Table V. Note that in both per-flow and per-router fairness cases, flow 10 has a much higher throughput than other flows. It is because router 4 can transmit simultaneously with router 1 for spatial reuse. Also note that the aggregate throughput of all the flows with per-flow fairness is lower than that with per-router fairness. In the case of per-flow fairness, each router exchanges the flow information only with its one-hop and two-hop neighbors, and adjusts its channel access time accordingly. Due to the lack of the flow information of the whole network, the resources may not be fully utilized. For example, after exchanging the flow information with routers 2 and 3, router 4 considers the fractions of channel access time of routers 2, 3, and itself are $1/2$, $1/3$, and $1/6$, respectively. However, router 2 can only get $1/3$ of channel time because of the contending flows from router 1. Without knowing this information, router 4 cannot fully utilize the channel time which is not used by router 2.

F. Priority Differentiation of Real-time Packets

In the preceding simulations, the network has sufficient resources to transmit all the real-time traffic with no packet dropping, so all the real-time packets are treated equally. In the following, we consider two scenarios where the real-time traffic load exceeds the network capacity, leading to packet dropping. In these scenarios, further priority differentiation is needed. The first scenario is shown in Fig. 6-(f), where flows 1 and 2 both consist of 11 video calls, flow 1 has 3 hops and flow 2 has 1 hop. The number n_r of real-time mini-slots is chosen to be 10, and the video packet delay bound D_{max} is set as 100 ms. We consider two methods to differentiate the priorities. First, D_{max} is uniformly divided (we refer to this method as uniform priority differentiation), and the urgency level of real-time mini-slot i is given by $U_i = i\frac{D_{max}}{n_r},\ (1 \le i \le n_r)$. Second, D_{max} is non-uniformly divided (we refer to it as non-uniform priority differentiation). We differentiate packets approaching the due time with a small scale, and non-urgent packets with a large scale. Specifically, the urgency level of real-time mini-slot i is given by

$$\begin{cases} U_i = \frac{1}{a} i \frac{D_{max}}{n_r}, & 1 \le i \le \left\lfloor \frac{n_r}{b} \right\rfloor \\ U_i = \frac{D_{max} - \frac{1}{a}\left\lfloor \frac{n_r}{b} \right\rfloor \frac{D_{max}}{n_r}}{n_r - \left\lfloor \frac{n_r}{b} \right\rfloor}\left(i - \left\lfloor \frac{n_r}{b} \right\rfloor\right) & \\ \quad + \frac{1}{a}\left\lfloor \frac{n_r}{b} \right\rfloor \frac{D_{max}}{n_r}, & \left\lfloor \frac{n_r}{b} \right\rfloor + 1 \le i \le n_r \end{cases}$$

where $\lfloor\rfloor$ is the floor function, a and b (both larger than 1) are the adjustable parameters. For the considered scenarios, when $a = 10, b = 2$, the desired priority differentiation performance is achieved. Without considering further priority differentiation, the packet dropping rates of flows 1 and 2 are 3.82% and 0, respectively. With the uniform priority differentiation, they are 2.90% and 0.92%, respectively. With the non-uniform method, they are 1.81% and 1.61%, respectively. It is clear that without further priority differentiation, the packets are not fairly dropped, and all the packet dropping are from the flow with a relatively long path. The uniform priority differentiation is not effective to improve the fairness. With the non-uniform priority differentiation, the packets are dropped more or less fairly between the flows with different hops.

The second scenario is shown in Fig. 6-(g), where flow 1 and flow 2 consist of 40 and 10 video calls, respectively. Without considering further priority differentiation, the packet dropping rates of flows 1 and 2 are 3.05% and 0, respectively. The packets from the heavy-load flows are more likely to be dropped. With non-uniform priority differentiation, the packet dropping rates of flows 1 and 2 are 2.13% and 2.10%, respectively. The packets are dropped fairly between the flows regardless of the traffic load of each flow.

VI. Conclusion

In this paper we have proposed a novel collision-free MAC scheme supporting multimedia traffic for the wireless mesh backbone. The proposed scheme is distributed, simple, and scalable. Taking the unique characteristic of the wireless mesh backbone (i.e., all the routers are located at fixed sites) into consideration, the proposed MAC greatly reduces the control overhead in comparison with conventional contention based MAC schemes (e.g., IEEE 802.11). By eliminating collisions, reducing control overhead, and achieving maximal spatial reuse, the proposed MAC achieves much higher resource utilization than contention based MAC. In addition, the proposed scheme provides guaranteed priority access to real-time traffic and, at the same time, ensures fair channel access to data traffic. Simulation results demonstrate that it significantly improves the delay performance of real-time traffic and the end-to-end data throughput, as compared with IEEE 802.11 and DPRMA. The performance of the proposed scheme is analyzed and verified by computer simulations. This research should provide helpful insights to the development and deployment of future broadband wireless mesh networks.

References

[1] P. Wang and W. Zhuang, "A collision-free MAC scheme for multimedia wireless mesh backbone," in *Proc. IEEE ICC'08*, May 2008.

[2] I. F. Akyildiz, X. Wang, and W. Wang, "Wireless mesh networks: a survey," *Computer Networks*, vol. 47, no. 4, pp. 445-487, Mar. 2005.

[3] R. Jurdak, C. V. Lopes, and P. Baldi, "A survey, classification and comparative analysis of medium access control protocols for ad hoc networks," *IEEE Commun. Surveys Tutorials*, vol. 6, no. 1, pp. 2-16, First Quarter 2004.

[4] H. T. Cheng, H. Jiang, and W. Zhuang, "Distributed medium access control for wireless mesh networks," *Wireless Commun. Mobile Computing*, vol. 6, no. 6, pp. 845-864, Sept. 2006.

[5] "IEEE 802.11 WG, Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specification," Standard, IEEE, Aug. 1999.

[6] E. Carlson, C. Prehofer, C. Bettstetter, H. Karl, and A. Wolisz, "A distributed end-to-end reservation protocol for IEEE 802.11-based wireless mesh networks," *IEEE J. Select. Areas Commun.*, vol. 24, no. 11, pp. 2018-2027, Nov. 2006.

[7] R. Zhao and B. Walke, "Decentrally controlled wireless multi-hop mesh networks for high quality multi-media communications," in *Proc. ACM MSWiM'05*, Oct. 2005, Montreal, Quebec, Canada.

[8] H. Jiang, P. Wang, W. Zhuang, and X. Shen, "An interference aware distributed resource management scheme for CDMA-based wireless mesh backbone," *IEEE Trans. Wireless Commun.*, vol. 6, no. 12, pp. 4558-4567, Dec. 2007.

[9] J. W. Robinson and T. S. Randhawa, "Saturation throughput analysis of IEEE 802.11e enhanced distributed coordination function," *IEEE J. Select. Areas Commun.*, vol. 22, no. 5, pp. 917-928, June 2004.

[10] K. Tabata, Y. Kishi, S. Konishi, and S. Nomoto, "A study on the autonomous network synchronization scheme for mesh wireless network," in *Proc. IEEE PIMRC'03*, vol. 1, pp. 829-833, Sept. 2003.

[11] Y. Yan, M. Tomisawa, Y. Gong, Y. Guan, G. Wu, and C. Law, "Joint timing and frequency synchronization for IEEE 802.16 OFDM systems," in *Proc. IEEE Mobile WiMAX Symposium'07*, pp. 25-29, Mar. 2007.

[12] "IEEE Standard for Local and metropolitan area networks Part 16: Air Interface for Fixed and Mobile Broadband Wireless Access Systems Amendment 2: Physical and Medium Access Control Layers for Combined Fixed and Mobile Operation in Licensed Bands and Corrigendum 1," 2006

[13] P. Wang, H. Jiang, and W. Zhuang, "A dual busy-tone MAC scheme supporting voice/data traffic in wireless ad hoc networks," in *Proc. IEEE Globecom'06*, Nov.-Dec. 2006.

[14] A. A. Bertossi and M. A. Bonuccelli, "Code assignment for hidden terminal interference avoidance in multihop packet radio networks," *IEEE/ACM Trans. Networking*, vol. 3, no. 4, pp. 441-449, 1995.
[15] K. R. Chowdhury, P. Chanda, D. P. Agrawal, and Q.-A. Zeng, "DCA—a distributed channel allocation scheme for wireless sensor networks," in *Proc. IEEE PIMRC'05*, vol. 2, pp. 1297-1304, Sept. 2005.
[16] Q. He, L. Cai, X. Shen, and P. H. Ho, "Improving TCP performance over wireless ad hoc networks with busy tone assisted scheme," *EURASIP J. Wireless Commun. Networking*, vol. 2006, article ID 92039, pp. 1-11, 2006.
[17] P. Wang, H. Jiang, and W. Zhuang, "Capacity improvement and analysis for voice/data traffic over WLANs," *IEEE Trans. Wireless Commun.*, vol. 6, no. 4, pp. 1530-1541, Apr. 2007.
[18] Y. Wang, J. Ostemann, and Y. Q. Zhang, *Video Processing and Communications*. Prentice Hall, 2002.
[19] S. Jiang, J. Rao, D. He, X. Ling, and C. C. Ko, "A simple distributed PRMA for MANETs," *IEEE Trans. Veh. Technol.*, vol. 51, no. 2, pp. 293-305, 2002.
[20] E. Ong, M. Loke, W. Lin, Z. Lu; S. Yao,"Video quality metrics—an analysis for low bit rate videos," in *Proc. IEEE ICASSP'07*, vol. 1, Apr. 2007, pp. I-889-I-892.
[21] IEEE 802.11 WG, IEEE 802.11g, "IEEE Standard for Information technology-Telecommunications and information exchange between systems-Local and metropolitan area networks-Specific requirements - Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications: Amendment 4: Further Higher Data Rate Extension in the 2.4 GHz Band," 2003.
[22] IEEE 802.11 WG, IEEE 802.11e/D11, "IEEE Standard for Information technology-Telecommunications and information exchange between systems-Local and metropolitan area networks-Specific requirements - Part 11: Wireless Medium Access Control (MAC) and Physical Layer (PHY) specifications: Amendment 7: Medium Access Control(MAC) Quality of Service (QoS) Enhancements," Oct. 2004.
[23] X. Yang and N. H. Vaidya, "Priority scheduling in wireless ad hoc networks," in *Proc. ACM MobiHoc'02*, pp. 71-79, June 2002.
[24] R. Jain, A. Durresi, and G. Babic, "Throughput fairness index: an explanation," ATM Forum Document Number: ATM_Forum/99-0045, Feb. 1999.

Ping Wang received the B.E. in 1994 and M.E. degrees in 1997, from Huazhong University of Science and Technology, China, and the Ph.D. degree in 2008, from the University of Waterloo, Canada, all in electrical engineering. She is currently an assistant professor at School of Computer Engineering, Nanyang Technological University, Singapore. Her current research interests include QoS provisioning and resource allocation in multimedia wireless communications. She is a co-recipient of a Best Paper Award from IEEE ICC 2007. She is an Editor of EURASIP JOURNAL ON WIRELESS COMMUNICATIONS AND NETWORKING and INTERNATIONAL JOURNAL OF ULTRA WIDEBAND COMMUNICATIONS AND SYSTEMS.

Weihua Zhuang received the B.Sc. and M.Sc. degrees from Dalian Maritime University, China, and the Ph.D. degree from the University of New Brunswick, Canada, all in electrical engineering. Since October 1993, she has been with the Department of Electrical and Computer Engineering, University of Waterloo, Canada, where she is a Professor. She is a co-author of the textbook Wireless Communications and Networking (Prentice Hall, 2003). She is a co-recipient of a Best Paper Award from IEEE ICC 2007, a Best Student Paper Award from IEEE WCNC 2007, and the Best Paper Awards from Int. Conf. Heterogeneous Networking for Quality, Reliability, Security and Robustness (QShine 2007 and 2008). Her current research interests include wireless communications and networks, and radio positioning.

Dr. Zhuang received the Outstanding Performance Award in 2005, 2006, and 2008 from the University of Waterloo for outstanding achievements in teaching, research, and service, and the Premier's Research Excellence Award (PREA) in 2001 from the Ontario Government for demonstrated excellence of scientific and academic contributions. She is the Editor-in-Chief of IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY, and an Editor of IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, EURASIP JOURNAL ON WIRELESS COMMUNICATIONS AND NETWORKING, and INTERNATIONAL JOURNAL OF SENSOR NETWORKS. She is a Fellow of the IEEE and an IEEE Communications Society Distinguished Lecturer.