Mobile Networks and Applications 10, 593–606, 2005

A 802.11 Based Slotted Dual-Channel Reservation MAC Protocol for In-Building Multi-Hop Networks

JING ZHU *
Communications Technology Lab, Intel Corporation, 2111 NE 25th Ave., JF3-206, Hillsboro, OR 97124, USA

SUMIT ROY
Department of Electrical Engineering, University of Washington, Seattle, WA 98195-2500, USA

Abstract. In this paper, we propose a novel *slotted Dual-Channel Reservation* (DCR) MAC protocol that uses 802.11 primitives for providing both Quality of Service (QoS) and *fairness*. RTS/CTS handshaking is transmitted on a separated control channel to prevent successive collisions of RTS and CTS packets with existing data transmission. Furthermore, contention for channel access may be initiated *by sender as well as receiver* depending on the channel status for better fairness. A simple slot reservation algorithm in the data channel provides high efficiency. The main conclusion is that *reservation access* benefits both delay performance, and efficiency as well as fairness—the reason being that not only an *exposed* terminal can regain the channel more easily because of dramatically reduced contention (RTS-CTS) traffic, but also a *hidden* terminal receives less collisions in handshaking since any node winning a slot will quit contending on the control channel. Therefore, it is highly recommended to use *reservation access* even if the prevailing traffic is data, e.g. TCP. To enhance delay performance, we devise a fake packet repeating mechanism that can reserve the slot for a connection even if the user temporarily has no packets to transmit. Simulations based on key metrics—*throughput*, *fairness index* and *mean delay* are performed to validate the new protocol and quantify its advantages. The limitations of the proposed DCR-802.11 protocol due to need for global clock synchronization and dual channels are also discussed.

Keywords: IEEE 802.11 protocol, media access control, fairness, QoS

1. Introduction

Increasingly, wireless ad-hoc networks are used to provide connectivity to devices in environments where engineered network infrastructure do not exist or is expensive or difficult to deploy. The proliferation of wireless-enabled personal computing and communication devices such as desktop PCs, laptops and other handhelds (Tablet PCs, PDAs) as well as digital consumer electronic equipment (cameras and video disc players) implies that in-building ad-hoc networks [9] are expected to increase. A preferred architecture for such networks is multi-hop: packets are routed by intermediate nodes prior to reaching the destination. Nodes in the network thus behave both as hosts (generating and consuming network packets) as well as routers (relaying other's packets).

One of the key challenges for such multi-hop networks is the design of an appropriate medium access control (MAC) scheme for delivering multimedia traffic consisting of both data (e.g. http, ftp, email, etc.) and real-time applications (e.g. voice, video, etc.) while satisfying quality of service (QoS) guarantees such as assured throughput and mean delay for real-time applications. This is mainly because an ad-hoc setting has no fixed entities like base stations that can be used to coordinate upstream client communications and suffers from the hidden/exposed terminal problems in multi-hop environment.

The IEEE802.11 MAC protocol for wireless local area networks (WLAN) [4,7] uses the fundamental mechanism called the Distributed Coordination Function (DCF) based on CSMA/CA (Carrier Sensing Multiple Access with Collision Avoidance) as an asynchronous, contention-based access scheme. Retransmission of collided packets is managed according to binary exponential back-off (BEB) rules. DCF is known to suffer from degradation in throughput efficiency as numbers of users increase; further, BEB is historically known to not provide adequate delay bound guarantees. An optional four way handshaking technique—the request-to-send/clear-to-send (RTS/CTS) mechanism—has been adopted by 802.11 to solve the so-called hidden terminal problems and improve throughput. However, all current 802.11 networks are limited to one-hop scenario where it is assumed that all contending terminals can hear each other. When applied to a multi-hop scenario, the base 802.11 MAC with RTS/CTS encounters newer challenges with regard to aggregate throughput and *fairness* [1,19] (i.e. they may exacerbate user starvation scenarios). Thus, our main goal in this work is to present a novel 802.11-based MAC protocol that addresses *all* the key issues of *efficiency*, *fairness*, and *QoS guarantee* in wireless *multi-hop* networks jointly. We also remark that the term *ad-hoc* networks as currently understood has often emphasized a high degree of node *mobility* that naturally poses many challenges for the network (routing, path discovery etc.) layer. But for the in-building networks, a significant fraction of nodes are *static* (though they may be deployed according to semi-random topology due to various constraints) while others have only intermittent mobility; thus in this work, we will ignore node mobility as a first step in our search for an enhance multi-hop MAC protocol.

* Corresponding author.

The remainder of the paper is organized as follows. Section 2 gives a brief literature review of related works. Section 3 introduces the background issues such as dual-channel system and 802.11 fairness and inefficiency problem. Section 4 presents the details of the proposed slotted dual-channel reservation 802.11-based MAC protocol (DCR-802.11). Section 5 analyzes network performance based the key metrics: system capacity, saturation throughput, and mean delay. Section 6 shows numerical results to demonstrate the advantages of using the new protocol compared to the legacy 802.11. Finally, Section 7 concludes the paper and suggests avenues for future work.

2. Literature review

While RTS/CTS or Virtual Carrier Sensing (VCS) is useful in mitigating the hidden terminal problem in multi-hop scenarios, the *exposed terminal* remains largely unsolved, i.e., all neighbors of *both sender and receiver* are blocked by using NAV (Network Allocation Vector). This leads to un-necessary reduction in aggregate throughput in many cases, as has been shown in [10] for a 1-D chain of nodes whose throughput dramatically degrades with the chain length.

Systems with multiple channels have been proposed in the literature to assist the basic RTS/CTS exchange in reducing the exposed terminal problem and thereby increase spatial reuse. Dual Busy Tone Multiple Access (DBTMA) [6] employs two narrow-band channels (say 10 KHz) for the transmit busy tone BTt and the receive busy tone BTr, respectively to assist RTS/CTS in further reducing the impact of hidden and exposed terminals. Compared to the high bandwidth of the main (data) channel, the bandwidth overhead of such additional narrowband control channels is a tolerable cost. However, DBTMA is a pure contention-based MAC scheme and does not consider any QoS constrains and therefore is not suitable for supporting delay sensitive traffic, such as voice.

Recently, many solutions based on reservation access have been proposed in the literature to support QoS. A "piggyback reservation protocol" is adopted in MACA/PR [11], which allows the contention winner to reserve slots for subsequent packets. However, reservation in MACA/PR is achieved by asking all neighbors to exchange their reservation tables. This not only increases overhead significantly per se but is further impacted by errors that may frequently happen during such table exchange in wireless multi-hop environments. D-PRMA [8] is a slotted reservation-based MAC protocol that does not need to exchange reservation tables among neighboring nodes. However D-PRMA is a sender-initiated access scheme, therefore in the case that senders are out of range of each other, namely *hidden terminals* (such as in figure 3(b) and (c)), it suffers from the fairness problem as indicated in [1]. Our new protocol is similar to the D-PRMA scheme in the way that the reservation access is performed but is different in that it addresses the fairness problem that is not targeted by D-PRMA scheme.

Bharghavan et al. [1] pioneered work in addressing the fairness problem of MAC layer in a multi-hop environment. It was shown in [1] that the key problem is the lack of synchronization information about contention periods due to *hidden terminals*. To propagate such information, two new types of packet are introduced in [1]: DS (Data-Sending) and RRTS (Request-for-RTS). Nonetheless, [1] cannot solve all fairness problems because of inevitable collisions of RTS/CTS handshaking with the ongoing DATA transmission. Qiao et al. proposed a priority-based fair medium access control protocol (P-MAC) in [13], which adopts a uniform back-off scheme with only one parameter: Contention Window size CW_{opt}. CW_{opt} is suitably selected to reflect the relative weights among data traffic flows so as to achieve weighted fairness and the number of terminals contending for the wireless medium so as to maximize aggregate throughput. Wang et al. [17] proposed a scheme based on the estimated bandwidth share of all stations. However, a practical difficulty with this method in a multi-hop environment lies in bandwidth estimation, as sender and receiver usually see different contention environments. In this paper, we achieve fairness improvement with two enhancements that we believe are more pragmatic: (i) transmitting the contention signal RTS/CTS on a separate channel from DATA/ACK and (ii) dynamically switching the access between sender-initiated and receiver-initiated modes.

Thus, we propose a 802.11 based slotted dual-channel reservation MAC protocol (DCR-802.11) with the following features:

- Dynamic switching of channel access between sender-initiated and receiver-initiated modes to obtain good fairness.
- Synchronized data transmission to enable simultaneous transmission or reception among nearby nodes where feasible.
- Use a separate (control) channel for access contention resolution to prevent interference with data transmission.
- A simple yet robust reservation scheme to improve efficiency and provide QoS support requiring only one successful contention per data burst, thus dramatically reducing the amount of RTS/CTS transmissions on the control channel.
- A fake-packet repeating mechanism to enhance the delay performance for real-time traffic.

Although each of the above ideas have been separately employed (e.g. D-PRMA, DBTMA, etc.), we believe this work is the first to integrate and adapt their features into one protocol as a potential solution for M^3 (Multi-Channel, Multi-Hop, Multimedia) network, considering *efficiency*, *fairness*, as well as *QoS*.

3. Background

3.1. Description of slotted dual-channel 802.11 system

Tables 1 and 2 summarize the relevant physical and MAC layer parameters defined in IEEE 802.11 standards. 802.11b

Table 1
Parameter of IEEE802.11 PHY a/b.

	IEEE802.11b	IEEE802.11a
PLCP Preamble	144 bits (144 μs)	288 bits (16 μs)
PLCP Header	48 bits (48 μs)	40 bits (4 μs)
Channel Bit Rate	1-11 Mbit/s	6-54 Mbits/s
SIFS	10 μs	16 μs
Backoff-SlotTime	20 μs	9 μs
$CW_{\min}$	31	15
DIFS	50 μs	34 μs

Table 2
Parameter of IEEE802.11 MAC.

RTS	20 bytes
CTS	14 bytes
ACK	14 bytes
Propagation delay	1 μs

adopts DSSS (Direct Sequence Spread Spectrum) as the physical layer mechanism to provide variable data rates between 1–11 Mbps operating at 2.4 GHz, while 802.11a uses OFDM (Orthogonal Frequency Division Multiplex) modulation to support rates up to 54 Mbps in the 5 GHz band.

The following notations are introduced for subsequent use:

O: The number of data-slots in a frame
T_s: The duration of data-slot
σ: The duration of back-off slot
R_c: Bit rate on the control (RTS/CTS) channel
R_d: Bit rate on the data (DATA/ACK) channel
L_{RTS}: Length of RTS packet
L_{CTS}: Length of CTS packet
L_{DATA}: Length of DATA packet
L_{ACK}: Length of ACK packet
δ: Propagation delay

As illustrated in figure 1, the intended dual-channel 802.11 system has two half-duplex sub-channels at different frequencies for *control* and *data*, respectively. These two sub-channels carry RTS/CTS and DATA/ACK respectively, and are scheduled so as to avoid simultaneous transmissions. A frame is composed of O slots where the slot duration is given by

$$T_s = \frac{L_{\mathrm{DATA}} + L_{\mathrm{ACK}}}{R_d} + 2\delta + 2SIFS. \quad (1)$$

To simplify QoS control and channel reservation, we fix the slot duration and the frame length. A contention winner on the control channel can transmit a data packet in the same slot of the next frame on the data channel. Consequently, the duration of a control slot is equal to a data slot duration. If a data packet exceeds the payload size of a data slot, it will be segmented into smaller pieces before transmission.

While RTS/CTS handshaking is optional in current 802.11 protocol, it indicated in [2] that RTS/CTS should be used in the majority of the practical cases because of its capability to cope with hidden terminals and reduce the collisions. Here, RTS/CTS handshaking is mandatory for the proposed DCR-802.11 protocol.

3.2. *Inefficiency and unfairness problems of 802.11 in multi-hop environment*

3.2.1. *Inefficiency due to exposed terminals*

A 802.11 node signals its intent to access the channel by broadcasting an RTS (Ready-to-Send) packet with the destination address included in MAC header. The destination, on receiving RTS, replies by sending CTS (Clear-to-Send) packet to the sender. RTS and CTS packets include the expected duration of time for which the channel will be in use subsequently. Other hosts that overhear RTS or CTS packet must update their Network Allocation Vector (NAV) accordingly and defer their transmission for the duration specified in the NAV. This process is called *virtual carrier sensing*, which allows the area around the sender and receiver to be reserved for the forthcoming data exchange, thus avoiding the hidden terminal problem.

Figure 2 illustrates the operation of IEEE802.11 with four collinear nodes A, B, C, and D, where each node's range equals the distance to it's neighbors. We define *interference period* for a node as the time during which any signal from the node will cause the collision of the reference transmission. In another word, a node during its interference period should not transmit. In this example, when $B \leftrightarrow C$ are communicating, A and D are forced to defer. As we can see, the NAV blocking period is set much longer than the actual interference period, especially for node A, leading to inefficiency. For instance, it would be desirable to allow node A to transmit when it will not be interfered with; however, this is difficult to implement in an asynchronous system because of lack of global timing information.

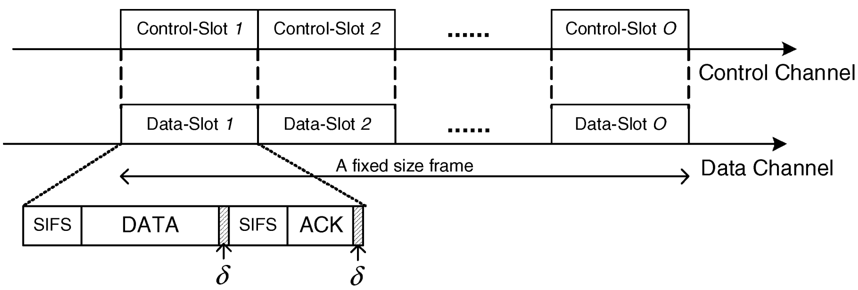


Figure 1. Structure of frame.

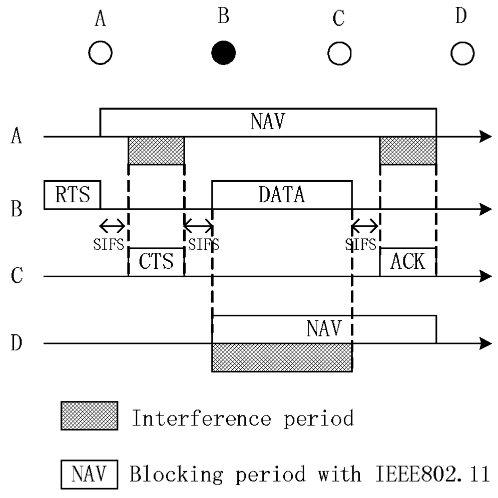


Figure 2. Operations of IEEE802.11 DCF.

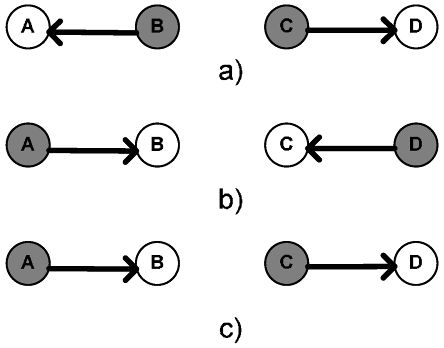


Figure 3. Three typical scenarios for studying fairness problem of Multi-Hop networks.

3.2.2. Unfairness due to hidden terminals

Figure 3 shows three typical scenarios for studying the unfairness problem of 802.11 MAC in multi-hop networks [1]. The four nodes A, B, C and D are separated by inter-node distance equal to the (common) transmission range so that each node can only reach its nearest neighbors. The black arrow indicates the direction of data transmission and colors grey and white denote transmitter and receiver respectively. A 200-second simulation run is performed for each scenario, and the average throughput is estimated over the last 100 seconds for both link $A \sim B$ and $C \sim D$. Here we assume heavy traffic load so that all transmitters always have data to send. All nodes are configured at the data rate 1 Mbps. The simulation results in Table 3 clearly shows that only Scenario (a) has good fairness—the main reason being the lack of synchronization information about contention periods, as identified in [1]. In other words, the contenders must be able to hear each other in order to achieve fairness. However, the legacy 802.11 only uses sender-initiated access leading to unfairness whenever a collision happens at the receiver and the senders cannot hear each other, i.e. *hidden terminals* (such as in Scenarios b and c).

From this example, we can clearly see that the purpose of RTS/CTS handshaking is to protect subsequent *data* collision at the receiver from hidden terminals; however, RTS packet itself is still highly vulnerable to interference from a data transmission and other RTS/CTS packets. Successive failures of handshaking will result in exponential back-off, which is the main source of unfairness in Scenarios (b) and (c). To solve this, we incorporate the new protocol with three key elements:

(i) protect RTS/CTS packets from data packets by transmitting them on a separate (control) channel;
(ii) protect RTS/CTS packets from other RTS/CTS packets by using reservation access to reduce the amount of handshaking needed;
(iii) initiate contention access at a non-hidden terminal by dynamically switching the access between sender and receiver.

Table 3
IEEE802.11b throughput comparison.

	Throughput of link $A \sim B$ (Mbps)	Throughput of link $C \sim D$ (Mbps)	Total throughput (Mbps)
Scenario (a)	0.43	0.43	0.86
Scenario (b)	0.25	0.61	0.86
Scenario (c)	0	0.86	0.86

4. Fundamentals of the slotted dual-channel reservation 802.11-based MAC protocol (DCR-802.11)

The key innovations behind the proposed DCR-802.11 are (i) eliminating the interference between data transmission and RTS/CTS handshaking by using separate sub-channels, (ii) improving the fairness by dynamically switching the access between sender-initiated and receiver-initiated modes, (iii) reducing the amount of handshaking and achieve higher efficiency by using reservation access, and (iv) solving the exposed terminal problem by synchronizing data transmission. Figure 4 illustrates the control flow of the novel DCR-802.11 MAC algorithm with three main stages: *contention access*, *reservation access*, and *fake packet repeating*.

4.1. Contention access

4.1.1. Sender and receiver initiated access for better fairness

One of the main enhancements in the proposed DCR-802.11 vis-a-vis the legacy 802.11 protocol is dynamically switching the access between sender-initiated and receiver-initiated states to achieve good fairness performance; the main idea is that an access attempt should be initiated by the one who can hear other contenders. For example, in Scenario (C) of figure 3, A is out of the transmission range of the contender C leading to a collision at B; hence B should initiate the access attempt rather than A. To fulfill this task, a new one-bit indicator (denoted as BLK) carried by CTS packets is used to inform sender the current status of the receiver: "blocked" or "available". Moreover, both CTS and

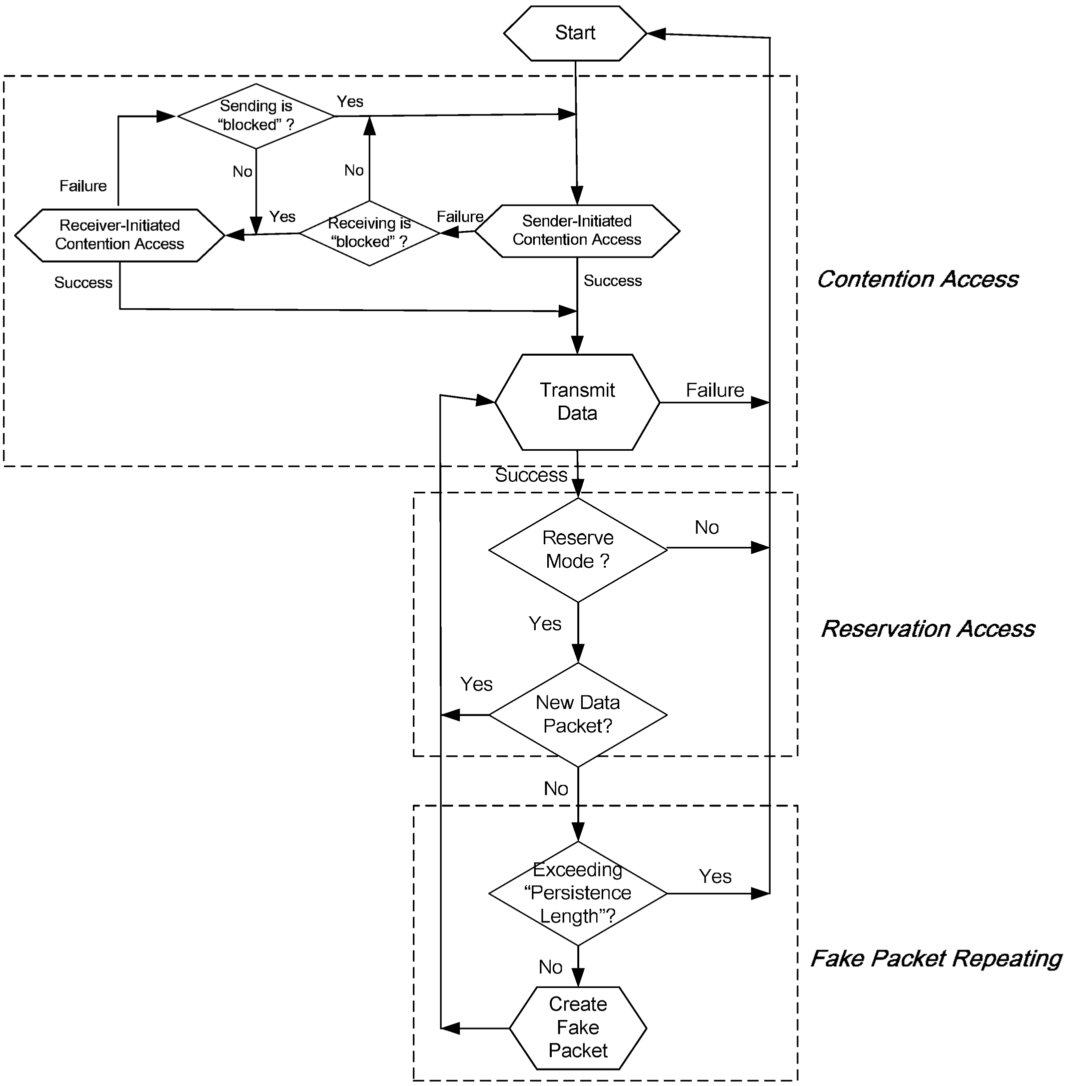


Figure 4. Flow chart of DCR-802.11 MAC algorithm.

RTS packets need another one-bit flag (denoted as RCV) [1] to indicate whether the access is initiated by "sender" or "receiver". An enhanced RTS/CTS handshaking rule is defined as follows:

- After successfully receiving a RTS, a station (sender or receiver) will send back CTS. If blocked, the station will mark the CTS packet with "blocked", then schedule an access attempt.
- After successfully receiving a CTS with "blocked" flag on, a station (sender or receiver) will freeze the current access attempt and wait for handshaking signals from the other node (sender or receiver).
- RTS from receiver and CTS from sender are marked as receiver-initiated; while RTS from sender and CTS from receiver are marked as sender-initiated.
- For a sender-initiated access, the sender will schedule a data transmission in the same data-slot of the next frame after successfully receiving the requested CTS packet marked with "available".
- For a receiver-initiated access, the sender will schedule a data transmission in the same data-slot of the next frame after successfully receiving an expected RTS packet as being "available" for sending data and feeding back a CTS packet with "available" on.

In addition, DCR-802.11 adopts 802.11-similar collision resolution and avoidance mechanisms:

(i) A station successfully receiving a RTS or CTS packet intended for another station will be blocked for the remaining duration of the slot.

(ii) The failure to receive a requested CTS or ACK implies collision, and the binary exponential back-off (BEB) algorithm is used for collision resolution.

[1] Totally, three new bits—RSV, BLK, and RCV in MAC header are used by the DCR-802.11

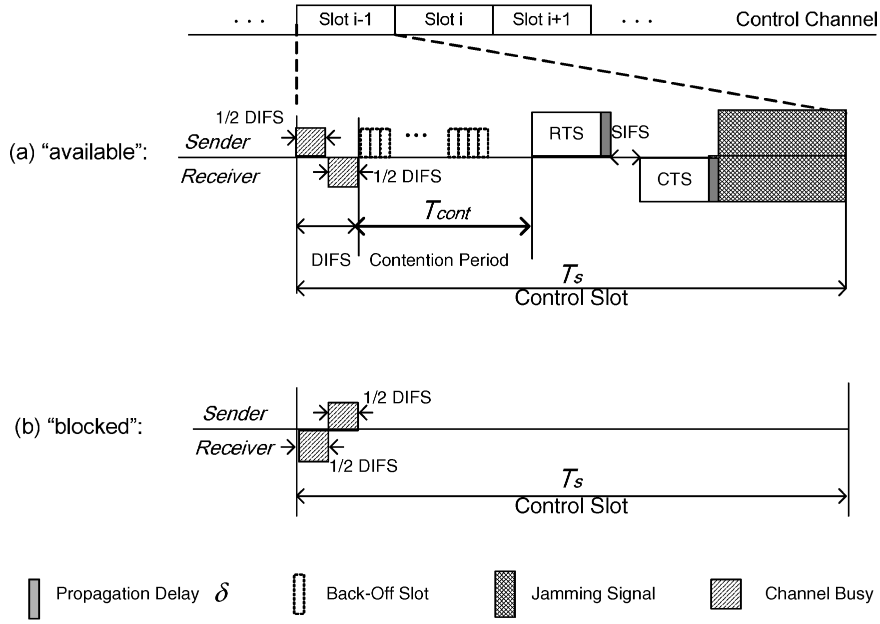


Figure 5. Signaling of basic contention access on control slot: (a) "available" and (b) "blocked".

4.1.2. Contention signaling

The basic contention access is conducted on control channel through the RTS-CTS signaling. The winner of control slot i in *this* frame can transmit a data packet in slot i of the *next* frame.

Figure 5 shows the signaling of basic contention access on a control slot. At the beginning of the slot, the station listens to the control channel in the first DIFS duration subject to the following:

- The station is prohibited from *receiving* data if it detects a signal in the first $\frac{1}{2}$DIFS, because at least one of its neighbors will *send* data in the data-slot next frame,
- The station is prohibited from *sending* data if it detects a signal in the second $\frac{1}{2}$DIFS, because at least one of its neighbors will *receive* data in the data-slot next frame.

A station is deemed to be in "available" or "blocked" state for sending or receiving data in every slot according to the above rules. If a station, for example, detects signal in the first half DIFS, but not in the second half, it will be "blocked" for sending data but "available" for receiving data (see the receiver in figure 5(a)). In figure 5(b), both sender and receiver are "blocked".

Only an "available" station can initiate an access attempt, for which a count-down process with a start value randomly chosen from $[0,\ CW]$ (CW: Contention Window) is triggered for every idle back-off slot till the end of the contention period in this data-slot, and will be continued in the next data-slot if not finished. Like 802.11, the back-off process is frozen if any signal is detected on the control channel, and restarted after an idle time of DIFS. A RTS/CTS handshaking is initiated as the counter reaches zero. The winner (a sender-receiver pair) of handshaking acquires the right of sending data in the same data-slot in the next frame.

As shown in figure 5(a), a station after sending a CTS with "blocked" flag off will jam the control channel till the end of this slot to prevent multiple successful RTS/CTS exchanges within the same data slot as this will lead to certain collisions on the subsequent data transmissions. Figure 6 shows an example of two winners in a slot when jamming is not used for the usual four collinear nodes scenario. Two active links – $A \to B$ and $C \to D$ are trying to transmit packets, and both initiate an access attempt in data-slot i. Although A failed in the first attempt due to the collision with the RTS packet from C, its second attempt was successful. As a result, A and C both are winners for data-slot i in the coming frame.

4.1.3. Contention duration T_{cont}

The RTS-CTS exchange must be completed in a slot duration to avoid ambiguity, as illustrated in figure 5(a). Clearly, the duration of the contention period, denoted as T_{cont}, is bounded by

$$T_{\text{cont}} \leq T_s - \left(DIFS + \frac{L_{\text{RTS}} + L_{\text{CTS}}}{R_c} + \delta + SIFS \right). \quad (2)$$

Inserting T_s from Eq. (1):

$$T_{\text{cont}} \leq \frac{L_{\text{DATA}} + L_{\text{ACK}}}{R_d} - \frac{L_{\text{RTS}} + L_{\text{CTS}}}{R_c} + \delta + SIFS - DIFS. \quad (3)$$

T_{cont} cannot be too short since the smaller the contention period, the lower the success probability of contention resolution, i.e., determining a winner. Let's consider an extreme case with only two stations (a sender and a receiver). With no contention from other nodes, all access attempts are successful, leading to a constant CW value equal to $CW_{\min}$. In order to make full use of data-slots, the sender must finish its access

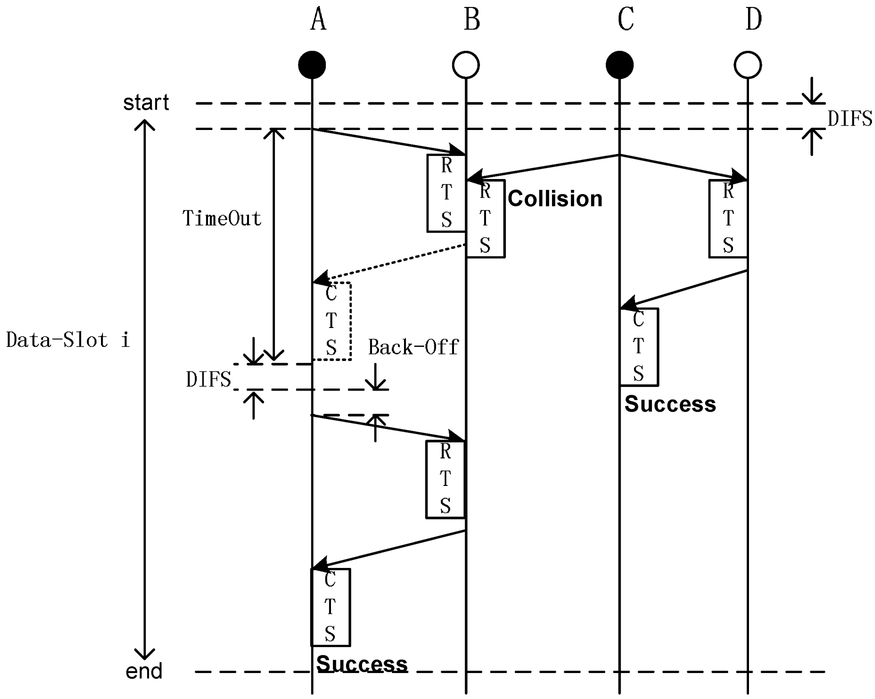


Figure 6. Two winners in a slot.

attempt always in one contention period, leading to

$$T_{\text{cont}} \geq CW_{\min} \times \sigma, \tag{4}$$

where σ is the duration of back-off slot. Combining Eqs. (2) and (3) leads to

$$R_c \geq \frac{L_{\text{RTS}} + L_{\text{CTS}}}{\frac{L_{\text{DATA}}+L_{\text{ACK}}}{R_d} - CW_{\min} \times \sigma + \delta + SIFS - DIFS}, \tag{5}$$

which gives a lower bound of the control channel bandwidth.

4.2. Reservation access for higher efficiency

DCR-802.11 has two modes: RSV (ReSerVe) and Non-RSV, as shown in figure 4. All users must undergo contention on packet-by-packet basis in the Non-RSV mode; in the RSV mode, contention is necessary only for the first packet of a burst arrival, and subsequent packets in the burst are transmitted by reservation. A reserved data-slot is released if there is no data to be transmitted.

During data transmission, no signal is transmitted on the control channel if using the Non-RSV mode; therefore after the first idle DIFS time, all other stations can start contending for the same data-slot of the next frame. In the RSV mode, a jamming signal with duration $\frac{1}{2}$DIFS is sent on the control channel during the first(second) $\frac{1}{2}$DIFS time by sender(receiver) respectively to prevent their neighbors from accessing control channel, so the same data-slot in the next frame is automatically reserved for the pair sending data in the current frame. A one-bit indicator (called "RSV") carried by DATA and RTS is needed to notify the receiver whether to send such jamming signal in the next frame. A receiver will schedule the jamming signal only after receiving a correct DATA packet with "RSV" on. To prevent unnecessary reservation, the sender must make sure that it does have a data packet for the reserved data-slot before marking "RSV".

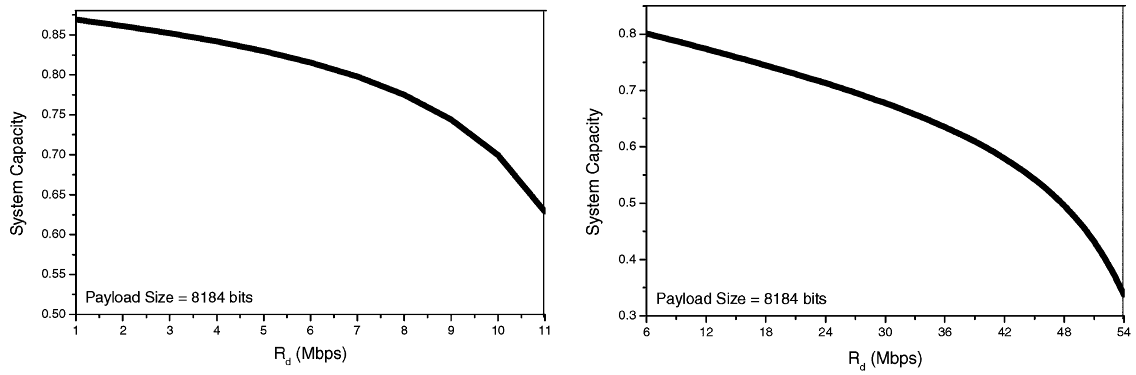
4.3. Fake-packet repeating mechanism for enhanced delay performance

In the RSV mode, a data-slot is reserved as long as data transmission is performed. Here, we propose a *fake-packet repeating* mechanism, which allows a connection to continue occupying a data-slot by having the sender transmit fake packets even if it has no data to transmit temporarily. Figure 4 illustrates the fake-packet repeating mechanism. A new parameter —*persistence length* is used to define the maximum number of successive fake packets. With a sufficiently long *persistence length*, the RSV mode of the new DCR-802.11 protocol can provide CBR (constant bit rate) service, and therefore support delay sensitive multimedia traffic, such as voice and video, but at the cost of wasting more bandwidth in sending fake packets.

5. Performance analysis

5.1. System capacity

Let η denote system capacity (i.e. the ratio of maximum data throughput to the total link data rate) and γ be the payload



a) 802.11 b b) 802.11 a

Figure 7. System capacity with Payload size 8184 bits with (a) IEEE802.11b and (b) IEEE802.11a.

ratio of a MAC data packet:

$$\eta = \gamma \frac{L_{\mathrm{DATA}}/T_s}{(R_c + R_d)} = \gamma \frac{L_{\mathrm{DATA}}}{(R_c + R_d)\left(\frac{L_{\mathrm{DATA}}+L_{\mathrm{ACK}}}{R_d} + 2(\delta + SIFS)\right)}, \tag{6}$$

where R_c is bounded by equation (5).

Figure 7 shows the system capacity with the payload size 8184 bits and the lower-bound of R_c for IEEE 802.11 a and b. Clearly, the system capacity decreases as data rate R_d increases because of the MAC inefficiency as reported in [18], i.e. the constant MAC overhead, such as DIFS, SIFS, back-off slot duration is more significant relative to reduced payload duration as the data rate is increased.

5.2. Saturation throughput

Previously introduced in [2], saturation throughput is an important metric that defines the normalized throughput achieved by a one-hop network where all terminals can hear each other under the assumption of heavily offered load, i.e., terminals always have packets to send.

The DCR-802.11 protocol has two modes: RSV and Non-RSV. Obviously, saturation throughput in the RSV mode is equal to system capacity, since all packets except the first one are transmitted without contention. In the Non-RSV mode where every packet must contend for data-slot before transmission, heavy loads will lead to multiple RTS collisions, and therefore the probability of event of no winner emerging in a slot increases.

The DCR-802.11 protocol employs the same back-off algorithm as the legacy 802.11, for which an accurate analytical model has been proposed to calculate the saturation throughput by Bianchi [2]. Let τ be the probability that a station starts a RTS/CTS handshaking in a back-off slot and p be the probability that a transmitted packet collides, note that p is also the probability that, in a time slot, at least one of the $\hat{n} - 1$ remaining stations transmits, Bianchi's model gives

$$\begin{cases} \begin{cases} \tau = \dfrac{2(1-2p)}{(1-2p)(W+1) + pW(1-(2p)^m)} & \hat{n} > 1 \\ p = 1 - (1-\tau)^{\hat{n}-1} & \end{cases} \\ \tau = 1 & \hat{n} = 1 \\ \left(W = CW_{\min},\ m = \log_2\left(\dfrac{CW_{\max}}{CW_{\min}}\right)\right) \end{cases}, \tag{7}$$

where $\hat{n}$ is the total number of active stations contending for a same data slot. Using numerical methods, we can derive the value of τ from equation (7). Each station can only access one data slot per frame, therefore

$$n \geq \hat{n} \geq n - O, \quad (O < n) \tag{8}$$

where n is the total number of active stations sharing a common channel and O is the number of data slots per frame. Let P_s be the probability of successful contention for a data slot as given by

$$P_s = \hat{n}\tau(1-\tau)^{\hat{n}-1}\left[1 + (1-\tau)^{\hat{n}} + (1-\tau)^{2\hat{n}} + \cdots + (1-\tau)^{(K-1)\hat{n}}\right] = \hat{n}\tau(1-\tau)^{\hat{n}-1}\frac{1-(1-\tau)^{\hat{n}(K-1)+1}}{1-(1-\tau)^{\hat{n}}}, \tag{9}$$

where K is the total number of back-off slots per data slot. Obviously,

$$K \geq CW_{\min}. \tag{10}$$

Let μ denote saturation throughput, given by

$$\mu = (R_c + R_d)\eta P_s, \quad (O < n) \tag{11}$$

If $O = n - 1$, we have $P_s = 1$, leading to $\mu = \eta(R_c + R_d)$.

Now we study the case of $O \geq n$. Figure 8 illustrates an example with $O = n = 4$, where we assume that each station acquired a data slot in the ith frame. Because one station can only access one data slot, and contention can not be performed in the same slot as data transmission, no access attempt occurs in the first data slot of the ith frame, leading to an empty data slot in the $i+1$st frame. Similarly, the second data slot of the ith

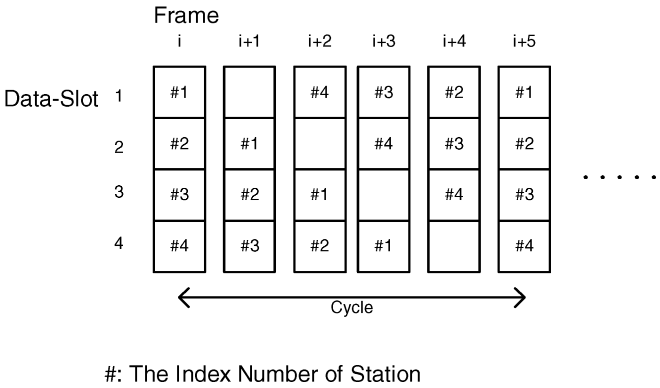


Figure 8. Saturation throughput of DCR802.11 Non-RSV mode with $O = n$.

frame only has one contender: station #1, which thus acquires that slot in the next frame; and similarly by induction for subsequent data slots in a frame. As we can see from figure 8, slot occupation in the $i + 5$th frame is the same as in the ith frame. Therefore, in steady state, the saturation throughput for the Non-RSV mode with $O = n$ can be simply calculated by

$$\mu = \frac{n}{n+1}\eta(R_c + R_d). \tag{12}$$

Similarly, it is not difficult to get the result for the case of $O > n$ as

$$\mu = \frac{n}{n+1+(O-n)}\eta(R_c + R_d). \tag{13}$$

In summary, we have

$$\mu = \begin{cases} (R_c + R_d)\eta P_s & (O \leq n-1) \\ (R_c + R_d)\eta & (O = n-1) \\ (R_c + R_d)\eta \dfrac{n}{n+1+(O-n)} & (O \geq n) \end{cases}. \tag{14}$$

The above equation gives the saturation throughput of the Non-RSV mode. Obviously, $O = n - 1$ is the optimal configuration. If the number of data slots in a frame exceeds $n-1$, the saturation throughput will be impacted due to the constraint that any station can only access one slot per frame. The problem could be solved by allowing a station to access multiple slots or making the number of data slots in a frame as few as possible (say 1). Since the approach of allowing multiple slots leads to many other issues such as slot management, which are out of scope of the paper, we use "O = 1" in our simulations and following discussions for simplicity.

A failed handshaking results in a wasted data-slot; therefore the Non-RSV mode with "O = 1" is equivalent to 802.11 DCF basic scheme (without RTS/CTS handshaking). However, the efficiency can be significantly improved by using the RSV mode. Let e be the average number of data transmissions following each successful contention; then equation (11) is modified to

$$\mu = (R_c + R_d)\eta \frac{P_s(1+e)}{P_s(1+e) + (1-P_s)}, \quad (O < n, e \geq 0). \tag{15}$$

Obviously, $e \gg 1$ leads to $\mu \approx (R_c + R_d)\eta$.

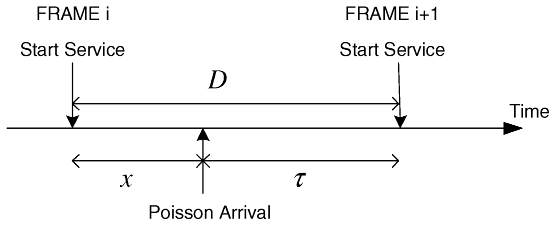


Figure 9. Synchronization delay τ.

5.3. Delay performance of slotted M/D/1 system

We define *delay* as the duration from the arrival of a packet to receiving the acknowledgement for the packet, denoted as d. Traffic arrival is modelled as a Poisson process. We only consider the RSV-mode with fake-packet repeating mechanism, which is aimed at CBR service. Assume that the *persistence length* is large enough such that all packets excluding the first are transmitted by reservation. In such case, a DCR-802.11 system can be modelled as a slotted M/D/1 system.

Denote service time as D and traffic load as ρ. The mean delay of an un-slotted M/D/1 system [15] is $\frac{D}{1-\rho}(1 - \frac{\rho}{2})$. In a slotted M/D/1 system, we must consider the extra delay due to synchronization (denoted as τ) (see figure 9, $x = D - \tau$). Under the assumption of Poisson arrival, x is an exponentially distributed random variable with mean $\frac{1}{\lambda}$ where λ is the arrival rate. Hence,

$$E(\tau) = E(D - x) = D - E(x) = D - \frac{\int_0^D \lambda e^{-\lambda x} x dx}{\int_0^D \lambda e^{-\lambda x} dx}$$
$$= D\left(\frac{1}{1-e^{-\rho}} - \frac{1}{\rho}\right). \tag{16}$$

Because each connection only uses one data slot in a frame, while service time is determined by frame length, i.e., $D = OT_s$. The ACK is received $(O - 1)$ data slots earlier than the start of next service. Thus, we have

$$E(d) = \frac{OT_s}{1-\rho}\left(1 - \frac{\rho}{2}\right) + OT_s\left(\frac{1}{1-e^{-\rho}} - \frac{1}{\rho}\right)$$
$$-(O-1)T_s, \tag{17}$$

With $O = 1$, the above equation simplifies to

$$E(d) = \frac{T_s}{1-\rho}\left(1 - \frac{\rho}{2}\right) + T_s\left(\frac{1}{1-e^{-\rho}} - \frac{1}{\rho}\right). \tag{18}$$

6. Numerical results and discussions

The simulation was performed using a MATLAB-based simulator [5] presented in [14], which accurately implements the base 802.11 protocol. A new module was developed for the proposed DCR-802.11b protocol with $R_d = 1$ Mbps and payload size of 8184 bits.

Multimedia traffic is usually better approximated by a correlated arrivals model such as MMPP (Markovian Modulated

Poison Process), than a pure Poisson process [16]. Nevertheless, for ease of simulation we use the Poisson arrival model; thus the evaluated throughput etc. are not intended to be taken as accurate estimates for a realistic scenario but is only used to demonstrate the *relative improvements* that are achievable with our enhanced DCR-802.11 MAC vis-a-vis the baseline 802.11 protocol.

The number of data-slots per frame on performance is set to 1, i.e., $O = 1$, so we do not consider the effect of slot management and allocation. R_c is set to the lower-bound to acquire the highest system capacity, i.e. $R_c = 0.082\ Mps$ and $\eta = 0.87$. The buffer size of MAC layer is fixed with 20 MAC data packets. Two key metrics—*throughput* and *fairness index* are defined as

$$\text{Throughput} = \text{Total Achieved Bit Rate}; \tag{19}$$

$$\text{Fairness Index} = \frac{\left(\sum_{i=1}^{n} \lambda_i\right)^2}{n \times \left(\sum_{i=1}^{n} \lambda_i^2\right)}, \tag{20}$$

where λ_i is throughput of the ith link.

6.1. DCR-802.11 one-hop performance

We study the effect of number of competing stations in a one-hop scenario on aggregate throughput. Total network traffic load is fixed and is assigned equally to each link. Normalized throughput for two values of load: 0.65 Mbps and 2 Mbps are investigated in figure 10. The theoretical value of saturation throughput for the Non-RSV mode is shown as well.

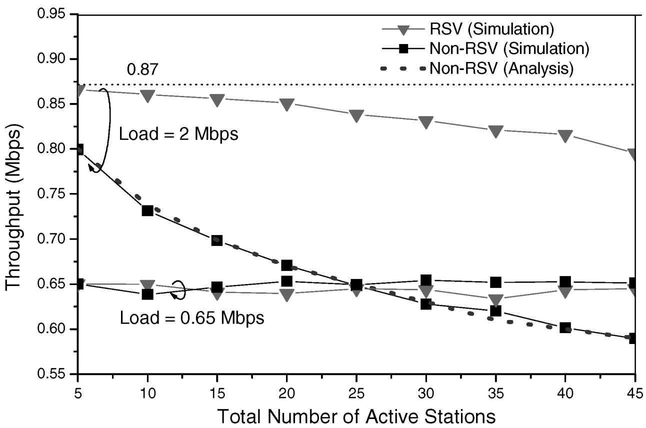


Figure 10. Effect of total number of active stations.

In general, throughput degrades due to the increased collision probability as the total number of active stations increases. Furthermore, the RSV mode achieves much higher throughput than the Non-RSV mode at heavy load (=2 Mbps). At light load of 0.65 Mbps (i.e., traffic arrival for each node is sparse), the RSV mode is equivalent to the Non-RSV mode since they achieve almost the same throughput in figure 10. Finally, our analysis results of saturation throughput for the Non-RSV mode match the simulation results very well, where the analysis results were obtained by using equation (14) with

$$\begin{cases} O = 1 \\ \hat{n} \approx n - 1 \\ K = CW_{\min} \end{cases}. \tag{21}$$

6.2. Multi-hop performance: DCR-802.11 vs. 802.11

First we study symmetric 8-node scenario, as shown in figure 11, which can be seen as an extended version of Scenario (a) and (b) in figure 3, with four links $1 \to 2$, $3 \to 4$, $5 \to 6$, and $7 \to 8$. Arrows indicate the direction of data flow and dotted lines indicates the neighbors of a node. We assume the same traffic arrival rate for all links, and define

$$\text{Load} = \text{Total Number of Arriving Packets per Second} \times 8184. \tag{22}$$

Figure 12 illustrates the fairness as the function of the load, showing that the proposed DCR-802.11 (RSV-mode and Non-RSV mode) can achieve perfect fairness ($\approx$1) in both symmetric scenarios, while the legacy 802.11 does not work well at heavy load ($>$1.7 Mbps) in Scenario b) due to its unfairness problem as discussed in Section II B). We further compare the throughput performance in figure 13. Clearly, the DCR-802.11 RSV mode significantly improves the efficiency in a multi-hop environment. The gain is about 100% at saturation in this case mainly due to higher spatial reuse by allowing the neighboring nodes to send or receive in a same data-slot. However in the Non-RSV mode, this gain is offset by the overhead of the contention procedure—single access attempt per data-slot.

Next we focus on asymmetric scenario (see figure 14), an extended version of Scenario (C) in figure 3. Obviously, only one link can be active at one time in this scenario; therefore both DCR-802.11 and the legacy 802.11 have almost the same throughput performance. With respect to the fairness, the

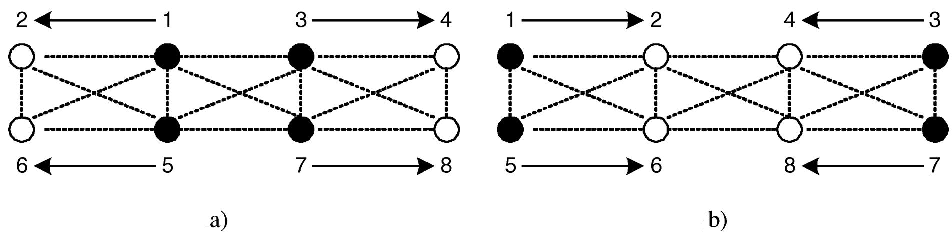
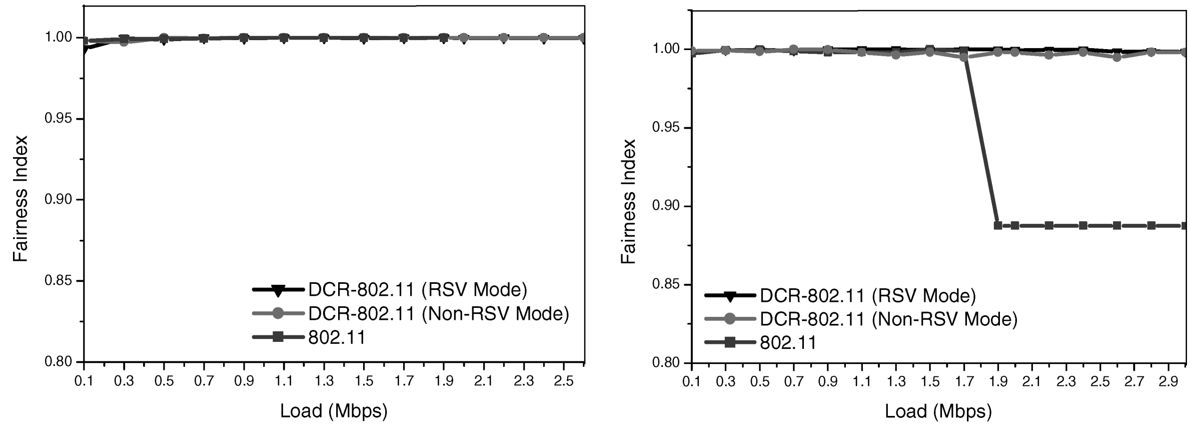


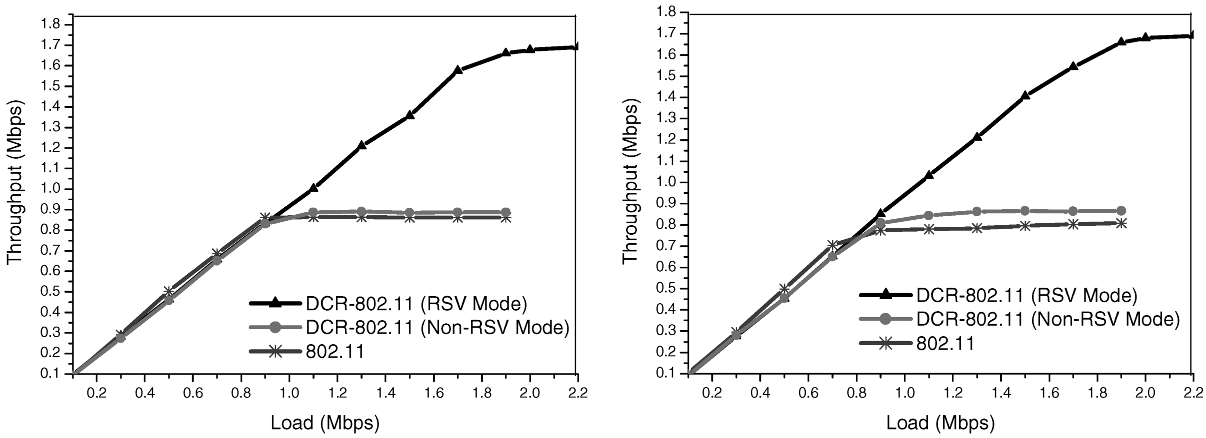
Figure 11. Symmetric 8-nodes scenario.



Scenario a) in Fig. 11

Scenario b) in Fig. 11

Figure 12. Fairness performance of DCR-802.11 for symmetric 8-node scenario.



Scenario a) in Fig. 11

Scenario b) in Fig. 11

Figure 13. Throughput comparison of DCR-802.11 with 802.11 for symmetric 8-node scenario.

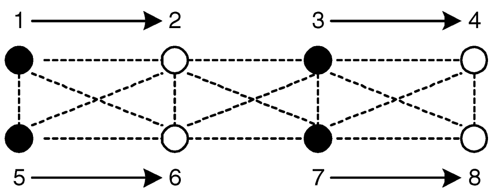


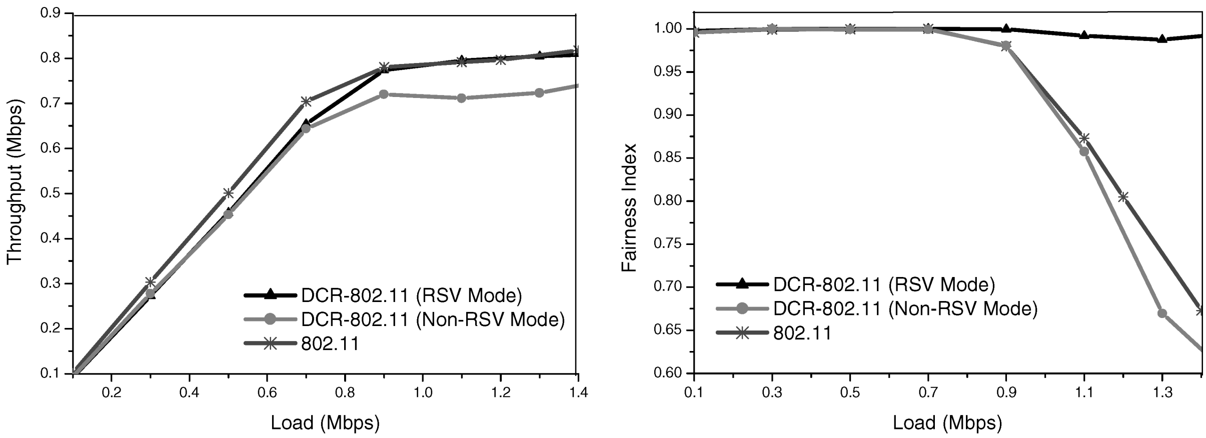
Figure 14. Asymmetric 8-nodes scenario.

DCR-802.11 RSV mode performs best, while the Non-RSV mode performs as bad as the legacy 802.11.

In the DCR-802.11, although the interference between data transmission and RTS/CTS handshaking is eliminated by using separate channels, a RTS may still collide with other RTS packets and the collision probability increases with the number of handshakes, which is proportional to the traffic load in the Non-RSV mode. If RTS packets collide, it is impossible for an access to change from sender-initiated to receiver-initiated; as a result, the DCR-802.11 is not very useful in such case. It is clear in figure 14 that links 3-4 and 7-8 prevail over links 1-2 and 5-6 in RTS contentions. For example, if node 1 and 3 send RTS in a same contention period, node 2 will not receive the RTS due to collision, while node 4 can successfully receive the RTS from node 1. Therefore, a deteriorating fairness performance is inevitable as the traffic load becomes heavier and heavier. One solution to the above problem is to reduce the number of handshakes, which is exactly one of the benefits of using the RSV mode. That also explains why the RSV mode can maintain good fairness performance as the traffic load increases.

In summary, the RSV mode should be used in most cases, especially when traffic load is high. However, if the load at a node in RSV mode is too high such that it always has packet to send, a data slot might be occupied by the node for a very long time. One solution is to limit the traffic from a node by employing a flow control mechanism. For example, we can enforce the node to release its occupied data-slot by setting a timer; we leave this issue for future investigation.

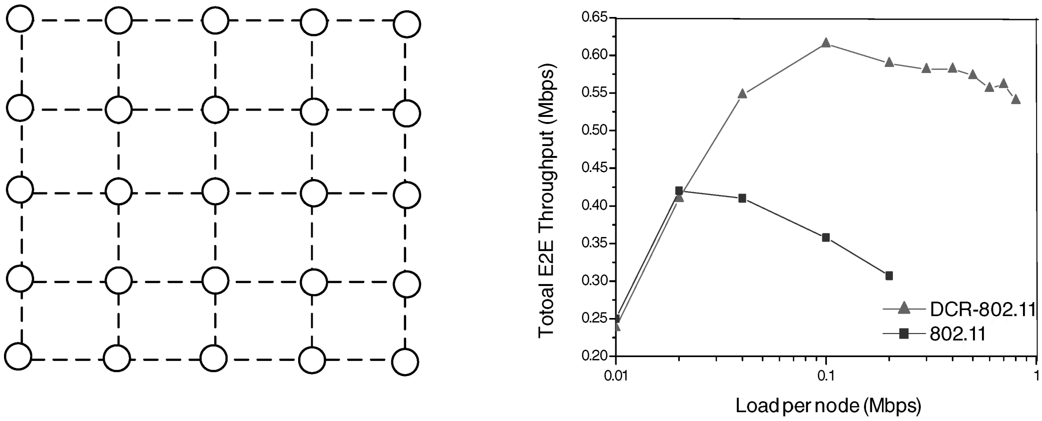
Finally, we compare the end-to-end(E2E) throughput of the DCR 802.11 with the legacy 802.11. The most common traffic pattern used in simulations of ad hoc networks has been random traffic: each source node initiates packets



a) Throughput

b) Fairness Index

Figure 15. Performance comparison of DCR-802.11 with 802.11 for asymmetric 8-node scenario.



a) 5x5 Grid Topology

b) Total E2E Throughput vs. Load per node

Figure 16. E2E throughput comparison of DCR-802.11 with 802.11.

to randomly chosen destinations in the network. We study E2E throughput performance of the DCR-802.11 when all nodes are in the RSV mode. Moreover, to make better use of a reserved slot, all packets with same next-hop address in a queue are transmitted successively. The shortest path routing policy is adopted to find a path from source to destination. A regular network topology -5×5 grid is investigated, shown in figure 16(a) (the dot line indicates wireless link, and each node can only reach its closest neighbors). Traffic arrival is still modeled as Poisson process, and all nodes have same traffic load. figure 16(b) shows the simulation results. It is clearly observed that DCR-802.11 achieves much higher E2E throughput than 802.11, especially at a higher load, implying that the proposed DCR-802.11 is more efficient than 802.11 in supporting multi-hop wireless networks.

6.3. Quality of service

We study the QoS support aspects of the DCR-802.11 using the 8-node scenario in figure 14 but with the signal transmission range set large enough to cover all nodes to emulate a one-hop network.

First, let link 1-2 be in the RSV mode, and other three links in the Non-RSV mode. We assume *greedy* traffic for

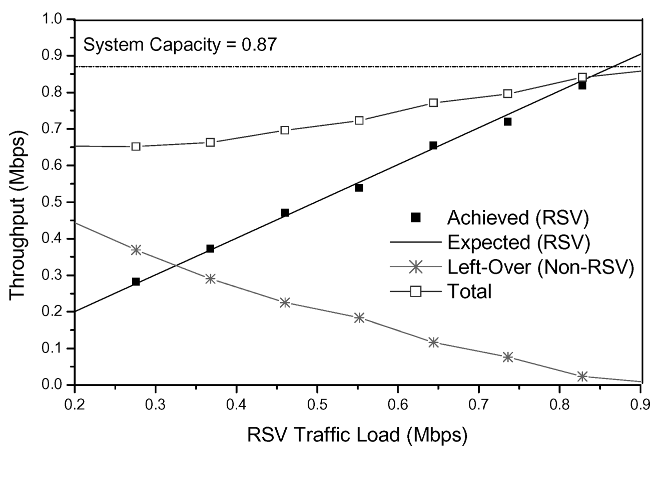
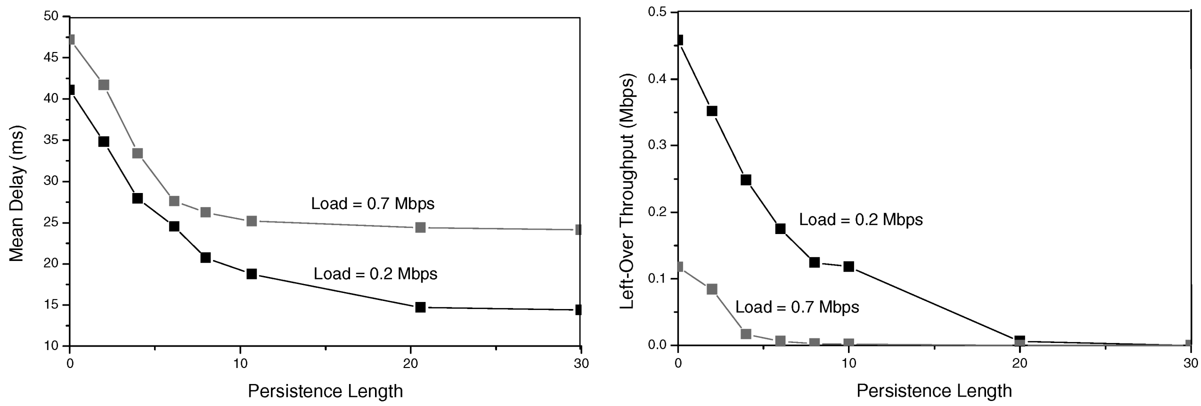


Figure 17. Assured throughput by using Non-RSV mode.



a) Mean Delay

b) Left-Over Throughput

Figure 18. Enhancing delay performance of RSV-mode with fake-packet repeating.

Non-RSV links such that they always have data to send, and a Poisson arrival for link 1-2. Figure 17 illustrates the performance with the load on the RSV link increasing with the following definitions:

- *Expected:* The expected throughput with zero packet drop rate for the RSV link 1-2;
- *Achieved:* The achieved throughput measured in simulation for the RSV link 1-2;
- *Left-Over:* The total throughput achieved by the three Non-RSV links;
- *Total:* The total throughput achieved by all four links.

It is clearly observed that the curve "Expected" overlaps with "Achieved" very well, implying that the RSV mode has higher priority than the Non-RSV mode, and delivers an *assured throughput*.

Then, we enhance the RSV link 1-2 using fake-packet repeating mechanism, focusing on *delay* performance. Two values of load on link 1-2 are investigated: 0.2 Mbps and 0.7 Mbps. Figure 18 shows that mean delay (measured in *ms*) is significantly reduced by increasing the persistence length, but at the cost of losing left-over throughput. Figure 19 shows the minimum mean delay obtained by setting the persistence length large enough (say 30). The analytical result from equation (18) is also given in figure 19 for comparison.

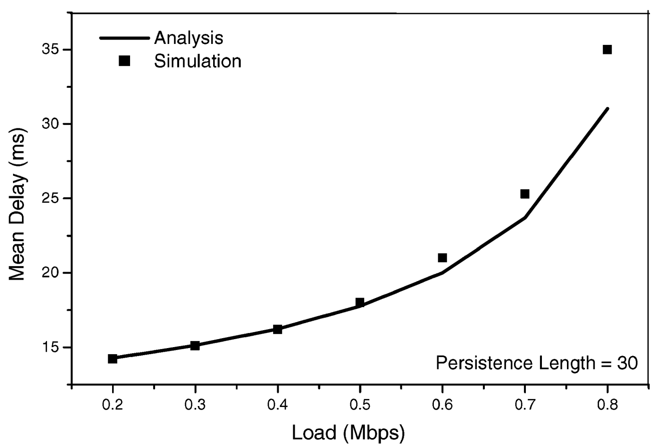


Figure 19. The minimum mean delay of RSV-Mode with fake-packet repeating.

7. Conclusions

In this paper, we studied the key issues of MAC design for wireless Multi-channel Multi-hop Multimedia (M^3) networks, focusing on efficiency, fairness and QoS. A new protocol, namely slotted dual-channel reservation 802.11-based scheme (DCR-802.11), is presented with preliminary simulation results. System metrics are analyzed, including *system capacity*, *saturation throughput*, and *mean delay*. The proposed DCR-802.11 is characterized by its *simplicity*, *efficiency*, and *robustness*, requiring no complex estimation and computation, complicated signaling, or heavy control traffic. Thus, we have shown that DCR-802.11 is a valid solution of M^3.

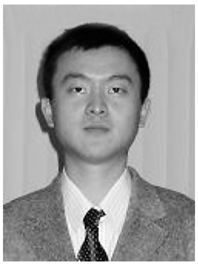
We identify the key limitations of the DCR-802.11 as the requirement for *multiple channels* and global clock *synchronization*. In [6], the authors discussed in detail the issues of implementing two radios in one device that can lead to potential interference. Our approach assumes frequency division to solve the interference problem between control and data traffic (though it does not mitigate the collision between control RTS/CTS packets at high loads, for which reservation is needed). Even if a two-radio device is not available, our scheme can be easily modified and implemented with one radio using time division to separate control and data slots. Synchronization needed for channel slotting in ad hoc environments with node mobility is clearly the greatest challenge for the DCR-802.11. However, with technology evolution and cost reduction of location systems such as Global Position System (GPS), a solution based on integrating such transceivers in each node (such as laptops) may be feasible in the future. This is akin to code division multiple access (CDMA)-based third-generation cellular system, where GPS may be used for synchronization between a base station and cellular phones [3].

Further, we note the following issues that remain unexplored and are left for future work:

(1) We ignored the effect of the total number of data-slots per frame (i.e. frame length) on the performance. However, longer the frame length, the more the RSV links can be supported simultaneously, but the longer the service time for each RSV link as well. It will be interesting to study how to choose an appropriate frame length according to traffic characteristics and network scale.
(2) We allowed each node to access only one data slot per frame no matter which mode is used. However, it might be useful for a node to access multiple data-slots for higher efficiency when the total number of data-slots in a frame exceeds the number of stations, or supporting high-rate multimedia traffic when the output rate provided by one data-slot is not enough.
(3) This paper ignored the impact of the MAC protocol on routing algorithms (such as DSR, AODV). We believe that the proposed DCR-802.11 possesses many properties that will be conducive for an efficient routing protocol.
(4) IEEE802.11e [12] is an on-going effort of IEEE standards for supporting QoS and improving efficiency in wireless LAN. Since the core of the proposed DCR-802.11 is fully compatible with the legacy 802.11 protocols, it is possible to integrate the main functionality of 802.11e into the DCR-802.11 framework to take the advantages of the DCR-802.11, such as reservation, multi-channel and dynamic switching between sender-initiated and receiver-initiated.

References

[1] V. Bharghavan, A. Demers, S. Shenker and L. Zhang, MACAW: A media access protocol for wireless LANs, in: *Proc. of ACM SIGCOMM'94*, 1994.
[2] G. Bianchi, Performance analysis of the IEEE 802.11 distributed coordination function, IEEE JSAC, 28(3) (2000) 535–547.
[3] E.A. Bretz and T.S. Perry, X marks the spot, maybe, IEEE Spectrum, 37 (2000) 26–36.
[4] B.P. Crow and J.G. Kim, IEEE 802.11 wireless local area networks, IEEE Comm. Mag., Sept. 1999.
[5] Sim Eleven, An IEEE 802.11 MATLAB-based Simulator, Available at: http//netlab1.bu.edu/saikat.
[6] Z.J. Hass and J. Deng, Dual busy tone multiple access (DBTMA)—A multiple access control scheme for ad hoc networks, IEEE Trans. Comm. 5(6) (2002) 975–985 .
[7] IEEE Standard for Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) sepcifications, ISO/IEC 8802-11: 1999(E) (Aug. 1999).
[8] S. Jiang, J. Rao, D. He and X. Ling, A simple distributed PRMA for MANETs, IEEE Trans. on Vehi. Tech. , 51(2), 293–305.
[9] L. Krishnamurthy, S. Conner, M. Yarvis, J. Chhabra, C. Ellison, C. Brabenac and E. Tsui, Meeting the demonds of the digital home with high-speed multi-hop wireless networks, Intel Technology Journal, 6(4) (Nov. 15, 2002).
[10] J. Li, C. Blake, D.S.J. De Couto, H.I. Lee and R. Morris, Capacity of ad hoc wireless networks, 7th ACM Mobilecom, 2001.
[11] C.H.R. Lin and M. Gerla, Asychronous multimedia multihop radio network, in: *Proc. IEEE INFOCOM, Kobe*, Japan (Apr. 7–11, 1997) pp. 118–125.
[12] S. Mangold, S. Choi, P. May, O. Klein, G. Hiertz and L. Stibor, *IEEE 802.11e Wireless LAN for Quality of Service*, European Wireless '2002, Florence, Italy, Feb. 2002.
[13] D. Qiao and K.G. Shin, Achieving efficient channel utilization and weighted fairness for data communications in IEEE802.11 WLAN under the DCF, the 10th ieee international workshop on qos, (2002) pp. 227–236.
[14] S. Ray, J.B. Carruthers and D. Starobinski, RTS/CTS-Induced Congestion in Ad Hoc Wireless LANs, *IEEE WCNC2003*, New Orleans, Lousiana, USA, 16–20 March 2003.
[15] M. Schwartz, *Broadband Integrated Networks* (Prentice Hall, 1998).
[16] M. Schwartz, *Broadband Integrated Networks* (Prentice Hall, Inc., 1998).
[17] Y. Wang and B. Bensaou, Achieving fairness in IEEE802.11 DFWMAC with variable packet lengths, IEEE GLOBECOM '01. 6(2001) 3588–3593.
[18] Y. Xiao and J. Rosdahl, Throughput and delay limits of IEEE802.11, IEEE Comm. Letter, 6(8) (Aug. 2002).
[19] S. Xu and T. Saadawi, Does the IEEE 802.11 MAC protocol work well in multihop wireless Ad hoc networks, IEEE Comm. Mag., June, 2001.



Jing Zhu received the B.S. and M.Sc. degrees both in electronic engineering from Tsinghua University, Beijing, P. R. China, in 2000 and 2001, respectively, and Ph.D. in electrical engineering from University of Washington (Seattle) in 2004. Currently, he is a research scientist in the Communications Technology Lab at Intel Corp. He is an IEEE member. His main research interest is high-performance wireless networking and its applications, e.g. Internet-on-Move, Digital Home, and Intelligent Transport System, focusing on cross-layer protocol optimization.
E-mail: jing.z.zhu@intel.com

Sumit Roy received the B. Tech. degree from the Indian Institute of Technology (Kanpur) in 1983, and the M. S. and Ph. D. degrees from the University of California (Santa Barbara), all in Electrical Engineering in 1985 and 1988 respectively, as well as an M. A. in Statistics and Applied Probability in 1988. His previous academic appointments were at the Moore School of Electrical Engineering, University of Pennsylvania, and at the University of Texas, San Antonio. He is presently Prof. of Electrical Engineering, Univ. of Washington where his research interests center around analysis/design of physical and data link layer of communication systems, with a topical emphasis on next generation mobile/wireless networks. He is currently on academic leave at Intel Wireless Technology Lab working on high speed UWB radios and next generation Wireless LANs. His activities for the IEEE Communications Society includes membership of several technical committees and technical program committees for conferences, and he serves as an Editor for the IEEE Transactions on Wireless Communications and Wiley Journal on Wireless Communications and Mobile Computing.
E-mail: roy@ee.washington.edu