15 Medium Access Control Protocols in Mobile Ad Hoc Networks: Problems and Solutions*

Hongqiang Zhai and Yuguang Fang

15.1 Introduction 232
15.2 Problems 233
15.2.1 MAC Layer Related Problems 233
Hidden Terminal Problem • Exposed Terminal Problem • Limitation of NAV Setup Procedure • Receiver Blocking Problem • The Desired Protocol Behaviors to Achieve Maximum Spatial Reuse • Limitation of IEEE 802.11 MAC Using a Single Channel
15.2.2 Flow Level Related Problems 236
Intra-Flow Contention • Inter-Flow Contention
15.2.3 Physical Layer Related Issues 237
Time-Varying Channel Condition and Rate Adaptation • Link Diversity
15.3 DUCHA: A New Dual-Channel MAC Protocol 239
15.3.1 Protocol Overview 239
15.3.2 Basic Message Exchange 239
RTS • CTS/NCTS • Data • Busy Tone • NACK
15.3.3 Solutions to the Aforementioned Problems 240
Solution to the Hidden Terminal Problem • Solution to the Exposed Terminal Problem • Solution to the Receiver Blocking Problem • Maximum Spatial Reuse • Inherent Mechanisms to Solve the Intra-flow Contention Problem
15.3.4 Bandwidth Allocation 241

*This work was supported in part by the Office of Naval Research under Young Investigator Award N000140210464 and the National Science Foundation under Faculty Early Career Development Award ANI-0093241.

15.4 Distributed Flow Control and Medium Access Control . 242
15.4.1 Motivation 242
15.4.2 Solution 242
Rule 1: Assigning High Priority of the Channel Access to the Receiver • Rule 2: Backward-Pressure Scheduling
15.5 Rate Adaptation with Dynamic Fragmentation 245
15.5.1 Fragmentation Scheme 245
15.5.2 Rate-Adaptive MAC Protocol for Fragment Bursts 246
15.6 Opportunistic Media Access Control and Auto Rate Protocol (OMAR) 247
15.6.1 Hybrid Opportunistic Media Access 247
Sender-Initiated Opportunistic Media Access • Receiver-Initiated Opportunistic Media Access
15.6.2 Rate Adaptation 248
15.6.3 Scheduling 249
15.7 Conclusion 249
References 249

15.1 Introduction

Recent advancements in wireless technologies and mankind's long-time dream of free communication are the driving forces behind the proliferation of wireless local area networks (WLANs) and the "hot" research activities in mobile ad hoc networks (MANETs). One of the most active topics is medium access control (MAC) protocols, which coordinate the efficient use of the limited shared wireless resource. However, in these wireless networks, the limited wireless spectrum, time-varying propagation characteristics, distributed multiple access control, low complexity, and energy constraints together impose significant challenges for MAC protocol design to provide reliable wireless communications with high data rates.

Among all MAC protocols, random medium access control (MAC) protocols have been widely studied for wireless networks due to their low cost and easy implementation. IEEE 802.11 MAC[10] is such a protocol that has been successfully deployed in wireless LANs and has also been incorporated in many wireless testbeds and simulation packages for wireless multihop mobile ad hoc networks. It uses four-way handshake procedures (i.e., RTS/CTS/DATA/ACK). The RTS and CTS procedures are used to avoid collisions with long data packets. The value of the NAV (network allocation vector) carried by RTS or CTS is used to reserve the medium to avoid potential collisions (i.e., virtual carrier sensing) and thus mitigate the hidden terminal problem. The ACK is used to confirm successful transmission without errors.

However, there are still many problems that IEEE 802.11 MAC has not adequately addressed. How to design a more effective transmission scheme based on the channel condition is still open and challenging. How to make full use of multiuser diversity in terms of multiple transmitters with the same receiver or the same transmitter with multiple receivers to maximize the throughput is also an interesting issue. In addition, in multihop ad hoc networks, the MAC layer contention or collision becomes much more severe than in the wireless LANs. Due to the MAC layer contention, the interaction or coupling among different traffic flows also deserves serious attention, which may limit the stability and scalability of multihop ad hoc networks.

At the MAC layer, the open shared channel imposes a lot of challenges for medium access control design. The hidden terminals may introduce collisions and the exposed terminals may lead to low throughput efficiency. In addition to these two notorious problems, the receiver blocking problem (i.e., the intended receiver does not respond to the sender with CTS or ACK due to the interference or virtual carrier sensing operational requirements due to the other ongoing transmissions) hence deserves serious consideration. In fact, this problem becomes more severe in multihop ad hoc networking environments and may result in throughput inefficiency, starvation of some traffic flows or nodes or re-routing. Many proposed solutions actually aggravate this problem by not allowing the hidden terminal to transmit. Furthermore, how to maximize spatial reuse by allowing the hidden terminals to receive and the exposed terminals to transmit is a very interesting issue.

Higher layer network protocols may be affected by wireless MAC protocols. It has been shown in many articles that multihop ad hoc networks perform poorly with TCP traffic and heavy UDP traffic.[3,7,15,16,25] This is because all wireless links in the neighborhood share the same wireless resource. All traffic flows passing through these links need to contend for the channel before transmission. Hence, severe MAC layer contention and collision can result in the contention among traffic flows. On the other hand, MAC contention can introduce network congestion with backlogged packets, which implies that network congestion is closely coupled with MAC contention. Some researchers have already noticed this kind of coupling. Fang and McDonald[6] demonstrated that the throughput and delay can be affected by the path coupling, that is, the MAC layer contention among the nodes distributed along the node-disjoint paths. Thus, cross-layer design and optimization is necessary for MANETs.

Moreover, at the physical layer, the time-varying channel condition makes rate adaptation necessary to improve network throughput. The diversity in the link quality due to the various channel conditions could be exploited to design opportunistic packet scheduling. The MAC protocol should be designed accordingly to adapt to the varying channel conditions.

In this chapter we first discuss the identified problems and challenges at different protocol layers to the design of MAC protocol. Then, we present several recently proposed novel schemes to address MAC layer problems, traffic flow-related issues, rate adaptation, and link quality diversity, respectively, in Sections 15.3 through 15.6. Finally, Section 15.7 concludes this chapter.

15.2 Problems

This section first discusses the inherent problems of the IEEE 802.11 MAC protocol in shared wireless channel environments in MANETs, and then illustrates the impact of traffic flows and physical layer channel conditions on the performance of this MAC protocol.

15.2.1 MAC Layer Related Problems

A packet collision over the air is much more severe in multihop environments than that in wireless LANs. Packet losses due to MAC layer contention will definitely affect the performance of the high layer networking schemes such as the TCP congestion control and routing maintenance because a node does not know whether an error is due to the collision or the unreachable address. It has been shown in many articles that multihop ad hoc networks performs poorly with TCP traffic as well as heavy UDP traffic.[3,16,25]

The source of the above problems comes mainly from the MAC layer. The hidden terminals may introduce collisions and the exposed terminals may lead to low throughput efficiency. In addition to these two notorious problems, the receiver blocking problem (i.e., the intended receiver does not respond to the sender with CTS or ACK due to the interference or virtual carrier sensing operational requirements for the other ongoing transmissions) also deserves serious consideration. In fact, this problem becomes more severe in multihop environments and results in throughput inefficiency and starvation of some traffic flows or nodes. The next few subsections describe a few problems in multihop mobile ad hoc networks when the IEEE 802.11 MAC protocol is deployed.

15.2.1.1 Hidden Terminal Problem

A hidden terminal is the one within the sensing range of the receiver, but not in the sensing range of the transmitter. The hidden terminal does not know that the transmitter is transmitting, and hence can initiate a transmission, resulting in a collision at the receiving node of the ongoing transmission.

One simple example is shown in Figure 15.1, where the small circles indicate the edges of the transmission range and the large circles represent the edges of the sensing range. D is the hidden terminal to A when

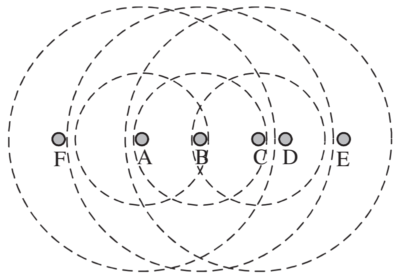

FIGURE 15.1 A simple scenario to illustrate the problems.

A is transmitting to B, and it cannot sense A's transmission but may still interfere with B's reception if D begins a transmission.

15.2.1.2 Exposed Terminal Problem

An exposed terminal is the one within the sensing range of the transmitter but not within that of the receiver. The exposed node senses the medium busy and does not transmit when the transmitter transmits, leading to bandwidth under-utilization. In 15.1, F is the exposed terminal to A when A is transmitting to B. F senses A's transmission and keeps silent, although F can transmit to other nodes outside of A's sensing range without interfering with B's reception.

In fact, in the four-way handshake procedures in IEEE 802.11 MAC, either RTS and CTS or DATA and ACK bidirectional packets are exchanged. Thus, the exposed node of one transmitter-receiver pair is also the hidden node of the other pair. So, in addition to the hidden terminal, the exposed terminal of the transmitter should not initiate any new transmission during the whole transmission process to avoid collision with the short packets ACK or CTS in IEEE 802.11 MAC. Thus, the carrier sensing strategy based on the RTS/CTS handshake will lead to a significant deficiency in spatial reuse.

15.2.1.3 Limitation of NAV Setup Procedure

IEEE 802.11 family protocols adopt short control packets (i.e., RTS/CTS) to resolve the long data packet collision and NAV setup procedures to claim the reservation for the channel for a certain period to avoid collision from the hidden terminals. This implies that any node that hears RTS/CTS correctly must set its NAV carried the received packets and keeps silent during the NAV period.

The NAV setup procedure cannot work properly when there are collisions. All kinds of packets, RTS, CTS, DATA or ACK, can be corrupted due to collisions. For example, in Figure 15.1, A wants to send packets to B. They exchange RTS and CTS. If E is transmitting when B transmits CTS to A, B's CTS and E's transmission will collide at C, and C cannot set its NAV according to the corrupted CTS from B.

NAV setup procedure is redundant if a node is continuously doing carrier sense. For example, in Figure 15.1, we can observe that both A's and B's transmission ranges are covered by the common area of A's and B's sensing ranges. If there is no collision, C will set NAV correctly when receiving B's CTS. However, it can also sense A's transmission, so NAV setup procedure is just redundant to prevent C from transmitting. RTS's NAV is not necessary either because any node that can receive RTS correctly can also sense B's CTS and succeeding A's DATA and B's ACK, and will not initiate new transmission to interrupt the ongoing transmission.

The NAV setup procedure does not help solve the hidden terminal problems even if there are no other collisions to prevent the CTS from setting up the neighbors' NAV. For example, in Figure 15.1, D is the hidden terminal to A. It cannot sense A's transmission and cannot receive B's CTS correctly either, because it is out of the transmission range of B. Thus, when A is transmitting a long data packet to B, D may begin to transmit a packet, which will result in a *collision* at B.

15.2.1.4 Receiver Blocking Problem

The blocked receiver is the one that cannot respond to the RTS intended for this receiver due to the other ongoing transmission in its sensing range. This may result in unnecessary retransmissions of RTS requests and subsequent DATA packet discarding. When the intended receiver is in the range of some ongoing transmission, it cannot respond to the sender's RTS according to the carrier sensing strategy in the IEEE 802.11 standard. The sender may attempt to retransmit several times if the backoff window is shorter than the long data packet. Then, the backoff window size becomes larger and larger when the RTS transmission fails and the window size is doubled, until the sender finally discards the packet. If the ongoing transmission finishes before the new sender reaches its maximum number of retransmissions allowed, the packet in the queue of an old sender will have higher priority than a new one because the old sender resets its backoff window size and is much shorter in size than that of a new one. So the old sender has a high probability of continuing to transmit and the new one continues doubling the backoff window size and discards packets when the maximum number of transmission attempts is reached. This will therefore result in serious unfairness among flows and severe packet discarding.

For example, as shown in Figure 15.1, when D is transmitting to E, A will not receive the intended CTS from B if it sends RTS to B. This is because B cannot correctly receive A's RTS due to collision from D's transmission. Thus, A keeps retransmitting and doubling the contention window until it discards the packet. If D has a burst of traffic, it will continuously occupy the channel, which will starve the flow from A to B.

The hidden terminal problem could make the receiver blocking problem worse. In the above example, even if A has a chance to transmit a packet to B, its hidden terminal D could start transmission and collide with A's transmission at B because D cannot sense A's transmission. Therefore, A almost has no chance to successfully transmit a packet to B when D has packets destined to E.

15.2.1.5 The Desired Protocol Behaviors to Achieve Maximum Spatial Reuse

The desired MAC protocol for multihop and wireless mobile ad hoc networks should at least resolve the hidden terminal problem, the exposed terminal problem, and the receiver blocking problem. Therefore, the ideal protocol should guarantee that there is only one receiver in the range of the transmitter and there is also only one transmitter in the range of the receiver. The exposed nodes may start to transmit despite the ongoing transmission. The hidden nodes cannot initiate any transmissions but may receive packets. Thus, to maximize the spatial reuse or network capacity, it should allow multiple transmitters to transmit in the range of any transmitter and multiple receivers in the range of any receiver to receive. In addition, the transmitter should know whether its intended receiver is blocked or is just outside its transmission range in case it does not receive the returned CTS to avoid packet discarding and the undesirable protocol behaviors at the higher layer, such as unnecessary rerouting requests.

15.2.1.6 Limitation of IEEE 802.11 MAC Using a Single Channel

The collisions between RTS/CTS and DATA/ACK, and that between DATA and ACK, are the culprits preventing us from achieving the aforementioned desired protocol behaviors.

The exposed terminal cannot initiate new transmission because its transmission would have prevented the current transmitter from correctly receiving the CTS or the ACK due to a possible collision.

The hidden terminal, which cannot sense the transmission or correctly receive the CTS, may initiate a new transmission, which will cause collision to the current ongoing transmission. In addition, it should not become a receiver because its CTS/ACK may collide with the current transmission. Moreover, its DATA packet reception can be corrupted by the ACK packet from the current receiver.

If the intended receiver for a new transmission is in the range of the ongoing transmission, it may not be able to correctly receive RTS and/or sense the busy medium, which prevents it from returning the CTS. Thus, the new sender cannot distinguish whether the intended receiver is blocked or out of its transmission range.

To summarize, many aforementioned problems cannot be solved if a single channel is used in the IEEE 802.11 MAC protocol.

15.2.2 Flow Level Related Problems

In wireless multihop ad hoc networks, nodes must cooperate to forward each other's packets through the networks. Due to contention for the shared channel, the throughput of each single node is limited not only by the raw channel capacity, but also by the transmissions in its neighborhood. Thus, each multihop flow encounters contentions not only from other flows that pass by the neighborhood (i.e., the *inter-flow contention*), but also from the transmissions of itself because the transmission at each hop must contend the channel with the upstream and downstream nodes (i.e., the *intra-flow contention*). These two kinds of flow contentions could result in severe collisions and congestion, and seriously limit the performance of ad hoc networks. In the following paragraphs, we discuss in detail their impacts on the performance of MANETs.

15.2.2.1 Intra-Flow Contention

The *intra-flow contention* here means the MAC layer contentions for the shared channel among nodes that are in each other's interference range along the path of the same flow. Li et al.[15] have observed that the IEEE 802.11 protocol fails to achieve optimum chain scheduling. Nodes in a chain experience different amounts of competition, as shown in Figure 15.2, where the small circle denotes a node's valid transmission range and the large circle denotes a node's interference range. Thus, the transmission of node 0 in a seven-node chain experiences interference from three subsequent nodes, while the transmission of node 2 is interfered by five other nodes. This implies that node 0 (i.e., the source) could actually inject more packets into the chain than the subsequent nodes can forward. These packets are eventually dropped at the two subsequent nodes. On the other hand, the redundant transmissions from node 0 grab the transmission opportunities of node 1 and node 2 because they cannot simultaneously transmit, and hence keep the end-to-end throughput far from the maximum value. We call this problem the *intra-flow contention* problem.

15.2.2.2 Inter-Flow Contention

In addition to the above contentions inside a multihop flow, the contentions between flows could also seriously decrease network throughput. If two or more flows pass through the same region, the forwarding nodes of each flow encounter contentions not only from its own flow, but also from other flows. Thus, the previous hops of these flows could actually inject more packets into the region than the nodes in the region can forward. These packets are eventually dropped by the congested nodes. On the other hand, the transmissions of these packets grab the transmission opportunities of the congested nodes, and hence impact the end-to-end throughput of the flows passing through the region. As shown in Figure 15.3, where there are two flows, one is from 0 to 6 and the other is from 7 to 12. Obviously, node 3 encounters the most frequent contentions and has little chance to successfully transmit packets to its downstream nodes. The packets will accumulate at and be dropped by nodes 3, 9, 2, 8, and 1. We call this problem as the *inter-flow contention* problem.

In the shared channel environments in multihop ad hoc networks, these two kinds of contentions are widespread and result in congestion at some nodes, where packets continuously accumulate, which then

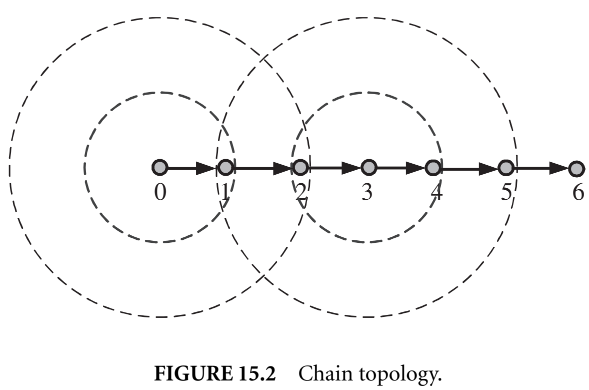

FIGURE 15.2 Chain topology.

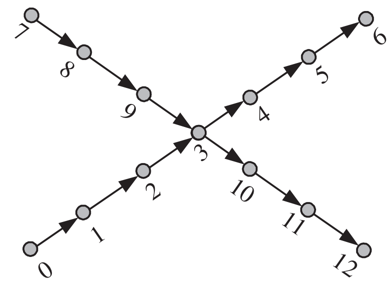

FIGURE 15.3 Cross traffic.

aggravates the contentions and finally results in packet dropping. This not only greatly decreases the end-to-end throughput, but also increases the end-to-end delay due to the long queueing delay.

15.2.3 Physical Layer Related Issues

15.2.3.1 Time-Varying Channel Condition and Rate Adaptation

A typical wireless communication link in wireless local area networks (LANs) is time-varying, and how to more effectively design transmission schemes based on the channel condition is challenging. Many adaptive transmission schemes have been proposed in the literature to enhance throughput performance. Many of these schemes are designed by varying the data rate, the transmission power, or the packet length. One of the popular schemes is based on rate adaptation, the adaptive transmission method that employs different modulation and coding schemes to adjust the data rate based on the channel condition (in terms of the signal-to-noise ratio [SNR]). The basic idea is to employ a high-level modulation scheme when a higher SNR is detected, as long as the target error rate is satisfied. The target error rate can be characterized by the bit error rate (BER), the symbol error rate (SER), or the packet error rate (PER), as specified by the designer. For receiver-based rate-adaptation schemes, the receiver usually carries out the channel estimation and rate selection, and the selected rate is then fed back to the transmitter.

Most of these protocols are receiver-based and employ the RTS/CTS collision avoidance handshake specified in the IEEE 802.11 standard. However, these protocols have not considered the possibility of bursty transmission of fragments in the corresponding rate adaptation schemes. The fixed preamble at the physical layer and the fixed inter-frame spacing (IFS) at the MAC layer have relatively large overheads when a high data rate is used and the transmission time for the payload is relatively short. Thus, reducing the overhead at a high data rate is essential for improving protocol efficiency.

15.2.3.2 Link Diversity

One of most interesting approaches to combating scarce spectrum resources and channel variations in wireless environments is opportunistic multiuser communication. Following the philosophy of cross-layer design, opportunistic multiuser communication utilizes the physical layer information feedback from multiple users, that is, multiuser diversity, to optimize media access control, scheduling, and rate adaptation. By allowing users with good link quality to transmit data in appropriately chosen modulation schemes, system performance in terms of goodput and energy efficiency can be greatly improved.

As the counterpart of multi-downlink diversity and multi-uplink diversity in cellular networks, multiuser diversity in ad hoc networks can be characterized as *Multi-Output Link Diversity* and *Multi-Input Link Diversity*.

Multi-Output Link Diversity is the diversity of instantaneous channel quality and congestion status of output links. Multiple flows may originate from or pass through a given node and take different neighbors as the next hop forwarding nodes or destinations. After this node acquires a transmission opportunity, it can choose a link with good instantaneous quality to transmit data in the given cycle. For example, as shown in Figure 15.4, node 1 is interfered by ongoing transmission of node 5 and the link of $0 \rightarrow 2$

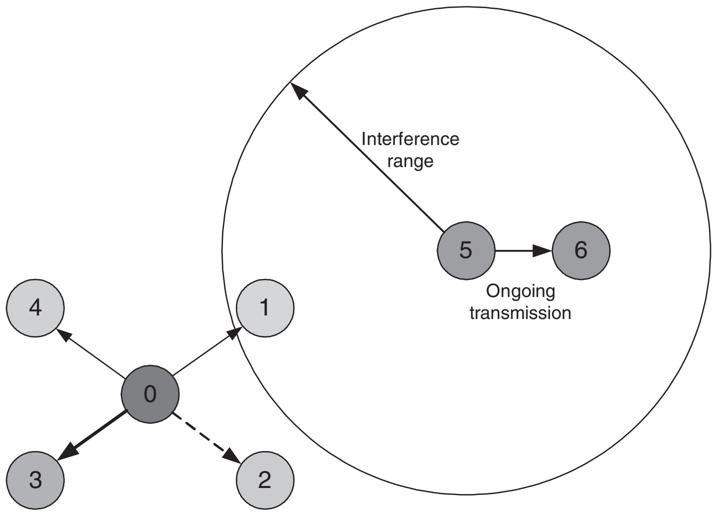

FIGURE 15.4 Multi-output link diversity.

suffers deep fading or shadowing. The link of 0 → 4 has instantaneous quality to support basic data rate transmission. The link quality of 0 → 3 happens to be "on-peak." At this time, it is better for node 0 to transmit data to node 3 or 4 rather than to node 1 or 2. Thus, the Head-of-Line blocking problem[1] can be alleviated and higher throughput can be achieved.

Multi-Input Link Diversity is the diversity of the channel quality and queue status of input links. Multiple flows originating from or passing through different neighbors take a given node as the next hop forwarding node or destination. Differences in instantaneous channel qualities of those input links form the multi-input link diversity. For example, as shown in Figure 15.5, node 1 is in the carrier sensing range of ongoing transmission of node 5. Similar to the previous example, node 3 or 4 instead of node 1 or 2 should take the opportunity to transmit packets to node 0 in this scenario.

Although diversity techniques have been widely studied and shown feasible and efficient in cellular networks, previous schemes may not apply to MANETs because they are based on an infrastructure where the base station acts as the central controller and dedicated control channels are normally available to feed back channel state periodically. To the best of our knowledge, multiuser diversity is still under investigation.

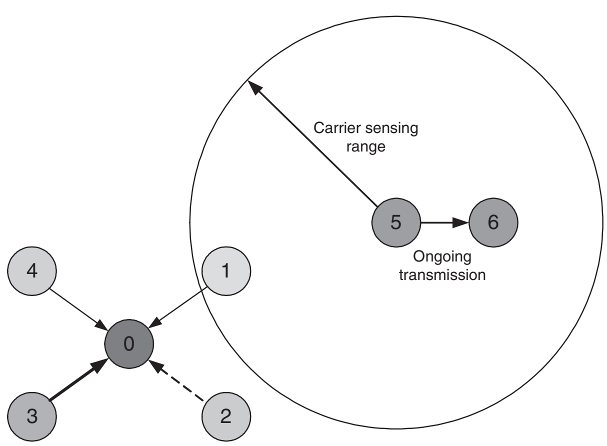

FIGURE 15.5 Multi-input link diversity.

However, there is little work that provides a comprehensive and realistic study of multiuser diversity with desired goals in protocol design of ad hoc networks.

Thus far, we have discussed a set of problems that the IEEE MAC protocol may present when we deploy it in multihop wireless ad hoc networks. The next few sections present some possible solutions that we have investigated recently to overcome or mitigate these problems.

15.3 DUCHA: A New Dual-Channel MAC Protocol

This section presents a new dual-channel MAC protocol (DUCHA) for multihop mobile ad hoc networks to mainly address the MAC layer related problems discussed above. More details can be found in Ref. 28.

15.3.1 Protocol Overview

To achieve the desired protocol behavior, we utilize two channels (dual-channel) for control packets and data packets, separately. RTS and CTS are transmitted over the control channel. Negative CTS (NCTS) is used to solve the receiver blocking problem and is also transmitted in the control channel. Data is transmitted over the data channel. An outband receiver-based busy tone[9,19] is used to solve the hidden terminal problem. The ACK is not necessary here because our protocol can guarantee that there is no collision to data packets. To deal with wireless channel error, we introduce the NACK signal, which is the continuing busy tone signal when the receiver determines that the received data packet is corrupted. The sender will not misinterpret this NACK signal because there are no other receivers in its sensing range and hence no interfering NACK signals and will conclude that the transmission is successful if no NACK signal is sensed.

Our protocol DUCHA adopts the same transmission power and capture threshold CP_{Thresh} in the control channel and the data channel. And the transmission power level for correct receiving RX_{Thresh} is also the same for the two channels.

15.3.2 Basic Message Exchange

15.3.2.1 RTS

Any node must sense the control channel idle at least for DIFS long and sense no busy tone signal before initiating a new transmission of an RTS. If it senses the noisy (busy) control channel longer than or equal to the RTS period, it should defer long enough (at least for SIFS + CTS + 2 × max-propagation-delay) to avoid possible collision to the CTS's reception at some sender. For example, in Figure 15.1, when A finishes transmitting its RTS to B, F should wait at least long enough for A to finish receiving the possible returning CTS/NCTS from B.

15.3.2.2 CTS/NCTS

Any node correctly receiving the RTS should return CTS after SIFS spacing, regardless of the control channel status if the DATA channel is idle.

If both the control and data channels are busy, it ignores the RTS to avoid possible interference to the CTS's reception at other RTS's transmitter. If the control channel is idle for at least one CTS packet long and the data channel is busy, it returns NCTS. The NCTS estimates the remaining data transmission time in its duration field according to the difference between the transmission time of maximum data packet and the length it has sensed a busy medium in the data channel.

15.3.2.3 Data

RTS's transmitter should start data transmission after correctly receiving the CTS if no busy tone signal is sensed. If the sender receives an NCTS, it defers its transmission according to the duration field of NCTS. Otherwise, it assumes that a collision occurred and will then double its backoff window and defer its transmission.

15.3.2.4 Busy Tone

The intended receiver begins to sense the data channel after it transmits CTS. If the receiver does not receive a signal with enough power in the data channel in the due time that the first few bits of the data packet reaches it, it will assume that the sender does not transmit data and finish the receiving procedure. Otherwise, it transmits a busy tone signal to prevent possible transmissions from hidden terminals.

15.3.2.5 NACK

The intended receiver has a timer to indicate when it should finish the reception of the data packet according to the duration field in the previously received RTS. If the timer expires and has not received the correct data packet, it assumes that the data transmission fails and sends NACK by continuing the busy tone signal for an appropriate time period. If it correctly receives the data packet, it stops the busy tone signal and finishes the receiving procedure.

The sender assumes that its data transmission is successful if there is no NACK signal sensed over the busy tone channel during the NACK period. Otherwise, it assumes that its transmission fails because of wireless channel error and then starts the retransmission procedure.

In addition, during the NACK period, in addition to the data transmission period, any other nodes in the sensing range of the sender are not allowed to become the receiver of data packets, and any other nodes in the sensing range of the receiver are not allowed to become the sender of data packets. This is to avoid confusion between NACK signals and the normal busy tone signals.

In the above message exchange, our protocol transmits or receives packets in only one channel at any time. We only use receive busy tone signals and not transmit busy tone signals. Thus it is necessary to sense the data channel before transmitting CTS/NCTS packets to avoid becoming a receiver in the sensing range of the transmitters of some ongoing data packet transmissions.

15.3.3 Solutions to the Aforementioned Problems

In the following discussions, we use examples to illustrate how our DUCHA solves those well-know problems.

15.3.3.1 Solution to the Hidden Terminal Problem

As shown in Figure 15.1, B broadcasts a busy tone signal when it receives a data packet from A. The hidden terminal of A (i.e., D) could hear B's busy tone signal and thus will not transmit in the data channel to avoid interference with B's reception. Thus, the busy tone signal from the data's receiver prevents any hidden terminals of the intended sender from interfering with the reception. Moreover, no DATA packets are dropped due to the hidden terminal problem.

15.3.3.2 Solution to the Exposed Terminal Problem

In Figure 15.1, B is the exposed terminal of D when D is transmitting a data packet to E. B could initiate RTS/CTS exchange with A although it can sense D's transmission in the data channel. After the RTS/CTS exchange is successful between B and A, B begins to transmit the data packet to A. Because A is out of the sensing range of D and E is out of sensing range of B, both A and E could correctly receive the data packet destined to them. Thus, the exposed terminal could transmit data packets in DUCHA, which could improve the spatial reuse ratio.

15.3.3.3 Solution to the Receiver Blocking Problem

In Figure 15.1, B is the blocked receiver in the IEEE 802.11 MAC protocol when D is transmitting data packets to E. In our protocol DUCHA, B can correctly receive A's RTS in the control channel while D sends data packets in the data channel. Then B returns the NCTS to A because it senses a busy medium in the DATA channel. The duration field of NCTS estimates the remaining busy period in the data channel required to finish D's transmission. When A receives the NCTS, it defers its transmission and stops the unnecessary retransmissions. It retries the transmission after the period indicated in the duration field of NCTS. Once the RTS/CTS exchange is successful between A and B, A begins to transmit the data packet to B. B will correctly receive the data packet because there is no hidden terminal problem for receiving data packets.

15.3.3.4 Maximum Spatial Reuse

As discussed above, the exposed terminals could transmit data packets. Furthermore, in our protocol, the hidden terminal could receive data packets although it cannot transmit. In Figure 15.1, D is the hidden terminal of A when A is transmitting the data packet to B. After the RTS/CTS exchange between E and D is successful in the control channel, E can transmit data packets to D. Because D is out of A's sensing range and B is out of E's sensing range, both D and E can correctly receive the intended data packets. Thus, our protocol DUCHA can achieve maximum spatial reuse by allowing multiple transmitters or multiple receivers in the sensing range of each other to communicate. At the same time, there are no collisions for data packets or for the NACK signals because there is only one transmitter in its intended receiver's sensing range and only one receiver in its intended transmitter's sensing range.

15.3.3.5 Inherent Mechanisms to Solve the Intra-flow Contention Problem

In our DUCHA protocol, the receiver of data packets have the highest priority to access the channel for the next data transmission. When one node correctly receives a data packet, it could immediately start the backoff procedure for the new transmission, while the upstream and downstream nodes in its sensing range are prevented from transmitting data packets during the NACK period. In fact, this could achieve optimum packet scheduling for chain topology and is similar for any single flow scenario.

For example, in Figure 15.2, node 1 has the highest priority to access the channel when it receives one packet from node 0 and hence immediately forwards the packet to node 2. For the same reason, node 2 immediately forwards the received packet to node 3. Then node 3 forwards the received packet to node 4. Because node 0 can sense node 1 and 2's transmissions, it will not interfere with these two nodes. Node 0 cannot send packets to node 1 when node 3 forwards packet to 4 because node 1 is in the interference range of node 3. When node 4 forwards packet to node 5, node 0 may have the chance to send a packet to node 1. In general, nodes that are four hops away from each other along the path could simultaneously send packets to their next hops. Thus, the procedure could utilize 1/4 of the channel bandwidth, the maximum throughput that can be approached by the chain topology.[15]

15.3.4 Bandwidth Allocation

We split the whole bandwidth into control and data channels. While the nodes are negotiating the transmission by RTS and CTS in the control channel, there is no transmission in the data channel for these nodes. On the other hand, when the nodes are transmitting data packets in data channel, the bandwidth in the control channel is not fully utilized. There exists an optimal bandwidth allocation for the two channels to reach the maximum throughput.

For simplicity of analysis, we assume that there are no collisions to all the packets, and the spacings between RTS, CTS, and data are fixed and can be neglected when compared to the control and data frames. The maximum throughput is determined by the packet length and the data rate of each channel. Let L_R, L_C, and L_D be lengths of RTS, CTS, and DATA, respectively; R_c and R_d be data rates of control and data channels, respectively; and *BW* be the total data rate (bandwdith). We observe that maximizing throughput is equivalent to minimizing the total time for a successful transmission of a packet, say, T_p. Thus, the problem is to minimize T_p under the condition $R_c + R_d = BW$. We can easily obtain

$$\begin{aligned} T_p &= \frac{L_R + L_C}{R_c} + \frac{L_D}{R_d} \geq \frac{(\sqrt{L_R + L_C} + \sqrt{L_D})^2}{R_c + R_d} \\ T_p &= \frac{(\sqrt{L_R + L_C} + \sqrt{L_D})^2}{BW}, \quad \text{when } \frac{R_c}{R_d} = \frac{\sqrt{L_R + L_C}}{\sqrt{L_D}} \end{aligned} \tag{15.1}$$

In IEEE 802.11, the total time for successful packet transmission when there is no transmission error (due to collisions or channel condition) is

$$T'_p = \frac{L_R + L_C + L_D}{BW}. \tag{15.2}$$

We can observe that the bandwidth splitting sacrifices bandwidth utilization ($T_p > T'_p$). However, our protocol can eliminate the collisions to data packets and greatly improve spatial reuse, leading to performance improvement for multihop ad hoc networks.

15.4 Distributed Flow Control and Medium Access Control

This section proposes a scheme to address flow level related issues by optimizing the cross-layer interaction between the MAC layer and the higher layer. More details can be found in Refs. 26 and 27.

15.4.1 Motivation

Considering the fact that contentions are from the transmission attempts of packets at different nodes, which are generated by various traffic flows, it is natural to exploit flow control to schedule the packet transmissions to reduce the collisions and congestion at the MAC layer. The intuitive solution is to allow the downstream nodes and the congested ones to transmit packets while keeping others silent, and hence smoothly forward each packet to the destination without encountering severe collisions or excessive delay at the forwarding nodes. This motivates us to develop our scheme presented in the next subsection.

15.4.2 Solution

We present a framework of flow control over the MAC layer and queue management to address the collisions and congestion problem due to the *intra-flow contention* and *inter-flow contention*. Based on the framework, a multihop packet scheduling algorithm is incorporated into the IEEE 802.11 MAC protocol. The salient feature here is to generalize the optimum packet scheduling of chain topology, which allows nodes four hops away to transmit simultaneously, to any traffic flows in general topology.

The framework includes multiple mechanisms. The *fast relay* assigns high priority of channel access to the downstream nodes when they receives packets, which reduces a lot of intra-flow contentions. The *backward-pressure congestion control* gives transmission opportunity to the congested node and keeps its upstream nodes silent. This could not only reduce excessive contentions in the congested area, but also quickly eliminate the congestion. It is also a quick method to notify the source to slow the sending rate down by exploiting the RTS/CTS of the IEEE 802.11 MAC protocol. The *receiver-initiated transmission scheme* uses a three-way handshake to resume the blocked flow at the upstream nodes when the congestion is cleared. It is a timely and economical approach with even less control overhead than the normal four-way handshake transmission in the IEEE 802.11 protocol. We discuss each of these mechanisms in detail in the next subsections.

15.4.2.1 Rule 1: Assigning High Priority of the Channel Access to the Receiver

In each multihop flow, the intermediate node on the path needs to contend for the shared channel with the upstream nodes when forwarding the received packet to the next hop. One way to avoid the first few nodes on the path to inject more packets than the succeeding nodes can forward is to assign high priority of channel access to each node when it receives a packet. This can achieve better scheduling for the chain topology.

For example, in Figure 15.2, node 1 has the highest priority when it receives one packet from node 0 and then forwards the packet to node 2. Node 2 immediately forwards the received packet from node 1 and forwards it to node 3. It is the same for node 3, which immediately forwards the received packet to node 4. Because node 0 can sense the transmissions of nodes 1 and 2, it will not interfere with these two nodes. Node 0 cannot send packets to node 1 when node 3 forwards packet to node 4 because node 1 is in the interference range of node 3. When node 4 forwards packet to node 5, node 0 may have chance

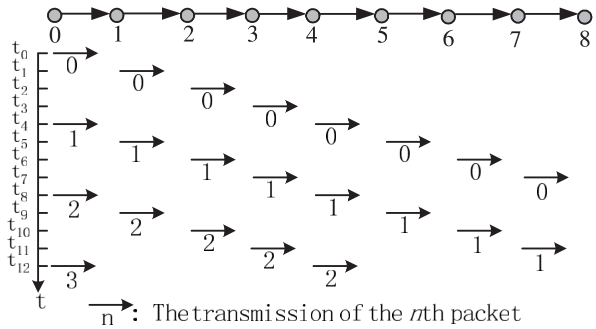

FIGURE 15.6 Optimum packet scheduling for chain topology. To simplify the illustration of how our scheme works, we use chain topology in this figure and the following ones, which is conceptually the same with any random multihop path in mobile ad hoc networks.

Au: Text in figure is not "readable"

to send a packet to node 1. Similar procedures are adopted by succeeding nodes along the path. Node 0 and node 4 can simultaneously send packets to their next hops, and a similar event happens to nodes that are four hops away from each other along the path. Thus, the procedure can utilize 1/4 of the channel bandwidth, the maximum throughput that can be approached by the chain topology.[15]

To incorporate this procedure into the IEEE 802.11 MAC protocol, our scheme OPET sets the initial value of the backoff window size of each receiver to 8. When it finishes the transmission, the scheme resets its contention window size to the normal value of 32.[10] The example in Figure 15.6 shows the optimum packet scheduling for the chain topology implemented by our scheme.

Rule 1 only considers the interference in a single flow. If the next hop of the current receiver is busy or interfered by other transmission, the receiver cannot seize the channel even with the highest priority. So we introduce backward-pressure scheduling to deal with inter-flow contention.

15.4.2.2 Rule 2: Backward-Pressure Scheduling

If one flow encounters congestion, it should decrease its sending rate to alleviate contention for the shared channel. Therefore, other flows in the neighborhood can obtain more channel bandwidth to transmit their packets to achieve higher utilization efficiency of the limited channel resource.

In addition to lowering the sending rate of the source, it is necessary to prevent the node, referred to as the *restricted node* in the following discussions, from transmitting packets to its next hop if the latter has already had many packets from the same flow. This can yield the transmission opportunity to the next hop as well as alleviate the congestion status.

The backward-pressure scheduling procedure takes advantage of RTS/CTS exchange in the IEEE 802.11 MAC protocol to restrict the transmission from the upstream nodes. A negative CTS (NCTS) should respond to the RTS when the intended receiver has reached the *backward-pressure threshold* for this flow. To uniquely identify each flow, RTS for the multi-hop flows (RTSM) should include two more fields than RTS, that is, the source address and the flow ID.

Our scheme OPET sets the *backward-pressure threshold* to 1, which indicates the upper limit of the number of packets for each flow at each intermediate node. As discussed, the optimum chain throughput in the IEEE 802.11 MAC protocol is 1/4 of the chain bandwidth and therefore the optimum threshold for the backward-pressure objective is 1/4, which is similar in operations for any single path. Because 1/4 is difficult to implement in the actual protocol, we select the nearest integer 1 as the value of this threshold.

Our scheme OPET adopts the receiver-initiated transmission mechanism to resume the *restricted* transmission. It uses the three-way handshake CTS/DATA/ACK instead of the normal four-way handshake RTS/CTS/DATA/ACK, because the downstream node already knows that the restricted node has packets destined to it. The CTS to resume the transmission (CTSR) should include two more fields than CTS, the source address and the flow ID, to uniquely specify the flow. CTSR, as well as CTS, has no information

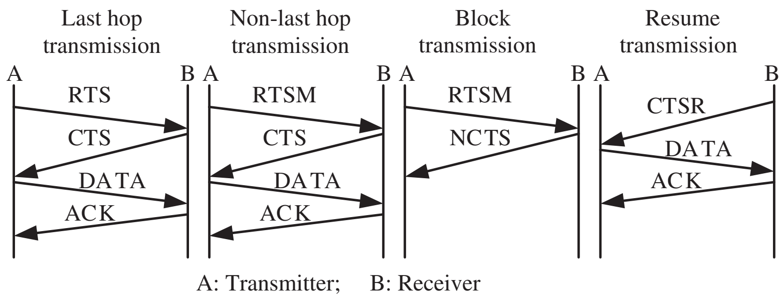

FIGURE 15.7 Message sequence for packet transmission.

about its transmitter as that in RTS. The two fields (i.e., the source address and the flow ID) are used to uniquely specify the next hop that the flow should pass through; hence, we assign different flow IDs to the flows from the same application but with different paths if multipath routing is used. The procedure of transmitting CTSR is similar to that of RTS and allows multiple retransmissions before dropping it. Different message sequences at different situations are shown in Figure 15.7.

To use the receiver-initiated transmission mechanism, we must consider that the mobility in ad hoc networks could result in link breakage followed by the transmission failure of CTSR. And CTSR may also collide several times and be dropped. The blocked node should drop CTSR after multiple retransmissions, as in the mechanism for RTS transmission. The restricted node should start a timer and begin retransmission if its intended receiver has not sent CTSR back in a long time, which we set 1 second in our study of the proposed scheme.

One simple example to illustrate how our scheme works is shown in Figure 15.8 and Figure 15.9. To simplify the illustration, we use chain topology, which is conceptually the same as any random multihop path in mobile ad hoc networks. When node 4 has congestion and cannot forward packet 0 to its downstream node 5, as shown in Figure 15.8, the flow along the chain will accumulate one packet at each node from node 1 to node 4 and then prevent nodes 0, 1, 2, and 3 from contending for the channel in order to reduce contention at congested node 4. After eliminating the congestion at node 4, the transmission will be resumed by the congested node, as shown in Figure 15.9.

It is important to note that the control overhead of backward-pressure scheduling is relatively low. The information of backward-pressure is carried by the original message sequence RTS/CTS in IEEE 802.11. And the blocked flows are resumed by a three-way handshake procedure with less overhead than the original four-way handshake. Moreover, our scheme only maintains several short entries for each active flow with at least one packet queueing up at the considered node to indicate the *blocked* status. We observe that in a mobile ad hoc network, the number of active flows per node is restricted by the limited bandwidth and processing capability, and hence is much smaller than that in wired networks; thus, the scalability problem should not be of major concern in our scheme.

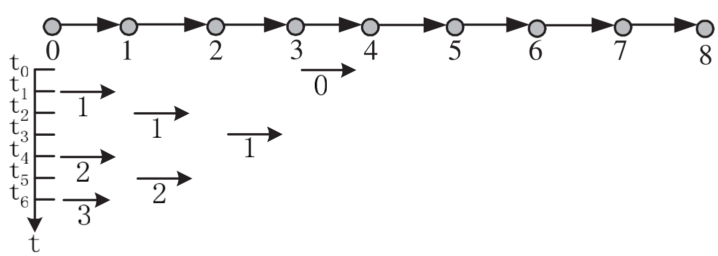

FIGURE 15.8 Packet scheduling when congestion occurs at node 4. The congestion can result from interference or contention from any crossing flow such that node 4 cannot grab the channel in time.

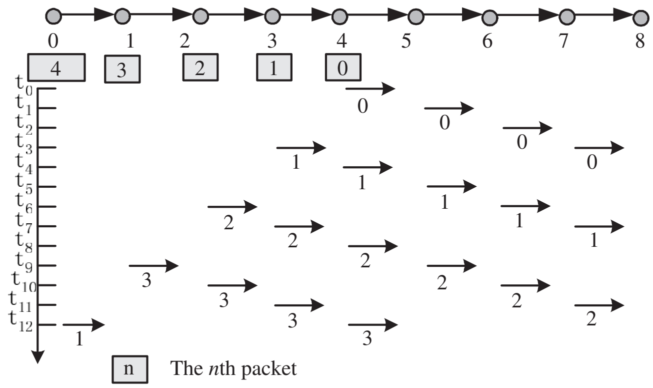

FIGURE 15.9 Packet scheduling after eliminating the congestion at node 4. After backward-pressure scheduling takes effect, the upstream nodes of this flow and all other crossing flows yield the transmission opportunity to the congested node. Thus, node 4 can quickly forward the backlogged packets and hence the congestion is eliminated.

Extensive simulation experiments are carried out to validate their performance. It turns out that our scheme can maintain stable performance with high throughput independent of traffic status, and improve the aggregated throughput by up to more than 12 times, especially for multihop flows under heavy traffic load. At the same time, it also improves the fairness among flows, and has much smaller delay and much less control overhead compared to the IEEE 802.11 MAC protocol. Moreover, it is scalable for large networks where there are more multihop flows with longer paths without incurring explosion of control packets under heavy load as the original 802.11 MAC protocol does.

15.5 Rate Adaptation with Dynamic Fragmentation

A rate-adaptive protocol with dynamic fragmentation is proposed to enhance throughput based on fragment transmission bursts and channel information. Instead of using one fragmentation threshold as in the IEEE 802.11 standard, we propose the use of multiple thresholds for different data rates so that more data can be transmitted at higher data rates when the channel is good. In our proposed scheme, whenever the rate for the next transmission is chosen based on the channel information from the previous fragment transmission, a new fragment is then generated using the fragment threshold for the new rate. This way, the channel condition can be more effectively utilized to squeeze more bits into the medium. Further details can be found in Refs. 12 through 14.

15.5.1 Fragmentation Scheme

The proposed dynamic fragmentation scheme contains the following key changes, as compared to IEEE 802.11 MAC, to enhance throughput in the time-varying wireless environment:

- The transmission durations of all fragments, except the last fragment, are set the same in the physical layer, regardless of the data rate.
- Different *aFragmentationThreshold*s for different rates are used, based on the channel condition; namely, a Rate-based Fragmentation Thresholding (RFT) scheme is employed.
- A new fragment is generated from the fragmentation process only when the rate is decided for the next fragment transmission, namely, Dynamic Fragmentation (DF).

In IEEE 802.11, with a single *aFragmentationThreshold*, the sizes of fragments are equal regardless of the channel condition. Therefore, the channel access time for a fragment varies with respect to the selected rate. For example, the channel access time for a fragment at the base rate is longer than that for a fragment at a higher rate. It is generally assumed that the channel remains unchanged during the transmission of a fragment at the base rate. Thus, more data frames can, in fact, be transmitted when a higher rate is used in the same duration provided that the SNR is high enough to support the higher rate. Due to this observation, the OAR protocol[17] proposes a multi-packet transmission scheme. However, multi-packet transmission has a higher overhead because of additional MAC headers, PHY headers, preambles in data and ACK, and SIFS idle times.

To overcome the shortcoming of multi-packet transmission, we fix the time duration of all data transmissions except for the last fragment. To generate fragments with the same time duration in a physical layer, the number of bits in a fragment should be varied based on the selected rate. Thus, it is necessary to have different *aFragmentationThreshold*s for different data rates selected by the receiver.

In the fragmentation process in IEEE 802.11 MAC, an MSDU is fragmented into equal-sized fragments that remain unchanged until all fragments in the burst are transmitted. If the channel quality is constant during the transmission of the fragment burst, the target PER (packet error rate) can be met. However, this is not guaranteed in a wireless LAN for two reasons. The first reason is that different fragments of a burst experience different channel qualities because of the time-varying nature of a wireless channel. The second reason is that after the transmission of a fragment fails, the sender contends for the channel again to transmit remaining fragments; thus, the channel quality is not guaranteed to be the same as that at the time that the first fragment is transmitted. To achieve the target PER, both the data rate and the fragment size should vary according to the changing channel condition. Moreover, to better match the varying channel, instead of generating all fragments before transmitting the first fragment, each fragment should be generated at each time when the rate is chosen for the next transmission. As a result, the fragments in a burst may not be of the same size. Figure 15.10 illustrates the process of the proposed dynamic fragmentation scheme. Notice that when the transmission of a fragment fails, the size of the retransmitted fragment may not be the same as that of the originally transmitted fragment because the channel condition may have changed.

15.5.2 Rate-Adaptive MAC Protocol for Fragment Bursts

With fragment burst transmission and rate adaptation for each fragment, data and ACK frames also participate in the rate adaptation process in the same way as RTS/CTS frames do. To support the rate adaptation process of a fragment burst, the physical layer header is modified as shown in Figure 15.11. The *signal* field in the PLCP header is divided into two 4-bit subfields, namely the current rate and the next rate subfields. The *current rate* subfield indicates the data rate of the current frame, while the *next rate* subfield indicates the selected data rate for the next incoming data frame. The values of the two subfields

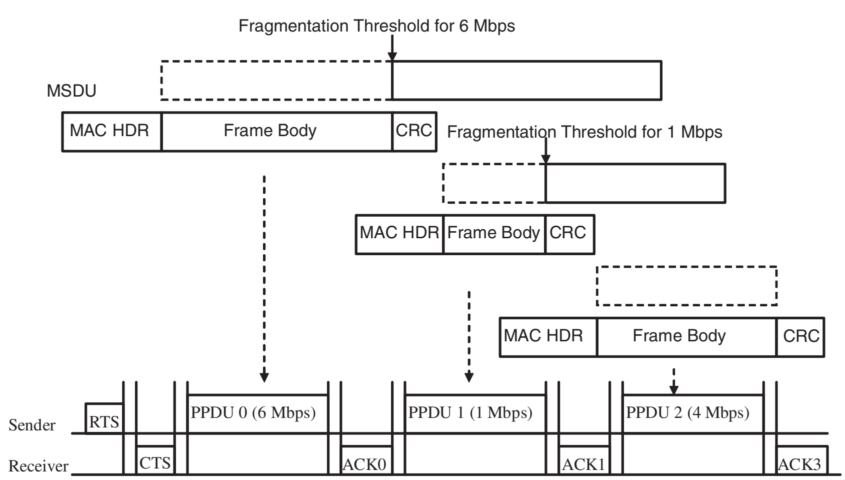

FIGURE 15.10 Dynamic fragmentation process and the timeline of data transmission.

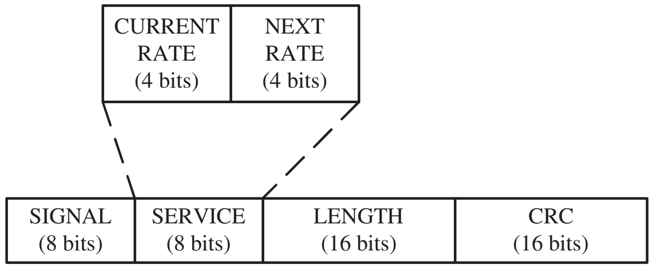

FIGURE 15.11 Physical layer header format in the proposed protocol.

in PLCP headers for RTS and data frames are the same because the *next rate* subfields in these headers indicate rates of frames transmitted from the receiver. After a sender sends an RTS frame at the base rate, a receiver estimates the channel and sends back a CTS frame to the sender with the selected rate stored in the *next rate* subfield. The sender modulates the fragment with the rate and sends a data frame to the receiver. After receiving the frame, the receiver predicts the channel condition for the next data frame and sends an ACK frame to the sender with the selected rate.

15.6 Opportunistic Media Access Control and Auto Rate Protocol (OMAR)

Due to the physical locations of various nodes in ad hoc networks, multiple nodes with various link qualities can transmit to (receive from) a common node; how to schedule the transmissions to utilize this diversity (user diversity) is challenging. The fundamental idea of OMAR is to exploit this diversity discussed in Section 15.2.3.2 through a collision avoidance process, which is necessary for CSMA/CA-based MAC. Based on sender-initiated and receiver-initiated collision avoidance, we introduce Multicast RTS (Request to Send) and Multicast RTR (Ready to Receive) to exploit multi-output link diversity and multi-input link diversity, respectively. Before the transmission of RTS (RTR) from a sender, each node selects a list of candidate receivers (senders), each with a different priority level, according to specific scheduling policy. The intended sender (receiver) multicasts a channel probing message (i.e., RTS or RTR) to the selected group of candidate receivers (senders). Each candidate receiver (sender) evaluates the instantaneous link quality based on the received channel-probing message. The candidate receiver (sender) with the highest priority among those with channel quality better than a certain level (threshold) is granted to access the medium. The details appear in Refs. 20 through 22.

The major components adopted in our scheme are hybrid opportunistic media access, rate adaptation, and packet scheduling. We detail these mechanisms in the following sections.

15.6.1 Hybrid Opportunistic Media Access

15.6.1.1 Sender-Initiated Opportunistic Media Access

Recognizing that RTS/CTS is a common mechanism to avoid collision in sender-initiated CSMA/CA, we extend the RTS/CTS handshake procedure to probe channels and utilize multi-output link diversity. Proposed Multicast RTS and Prioritized CTS are given as follows.

15.6.1.1.1 *Multicast-RTS*

The RTS used by IEEE 802.11 is a unicast short packet in that only one receiver is targeted. In our protocol, we use multiple candidate receiver addresses in RTS and request those receivers in the receiver list to receive the RTS and measure the channel quality simultaneously. Of course, each node must use an omni-directional antenna. The targeted data rate is added to the RTS for the declaration of the data rate that the sender wants to achieve at a given directed link. We dynamically set the targeted data rate according to recently measured channel conditions among those candidate receivers in the list. Each node monitors the transmissions of its neighbors and records the received power. In addition, considering that both MPDU size and data rate are variable, we use the packet size rather than the duration into RTS for each candidate receiver so that the corresponding receiver can derive duration according to the selected data rate based on the channel condition.

Anyone except the candidate receiver that receives the MRTS should tentatively keep silent to avoid possible collision before the sender receives the CTS. After the selected qualified receiver determines the transmission duration and send back CTS, the sender sets the duration field accordingly in the MAC header of the data frame[10] for the final NAV setting. The MAC header is sent at the basic rate so that all overhearing nodes can decode.

15.6.1.1.2 *Prioritized CTS*

The candidate receivers evaluate the channel condition based on physical-layer analysis of the received RTS message. If the channel quality is better than a certain level and its NAV is zero, the given receiver is a good candidate. To avoid collision when there are two or more good candidates, different inter-frame spacings (IFSs) are employed according to the listing order of intended receivers in the RTS. For example, the IFS of the nth receiver equals to $SIFS + (n-1)^{*}Time\_slot$. The receiver with the shortest IFS among those having the capability to receive data packet would reply CTS first. Because all candidate receivers are within one-hop transmission range of the sender and the physical carrier sensing range is normally larger than two hops of transmission,[15] the CTS may be powerful enough for all other qualified candidate receivers to hear or sense. Suppose that busy tone[9,19] is available; it would further enhance the physical carrier sensing capability. The intended receiver turns on the busy tone upon receiving the CTS; thus, those receivers that detect CTS or busy tone from other sources would yield the opportunity to the one transmitting CTS in the first place; that is, the one with the good channel condition and the highest priority. The duration to be advertised in the CTS is set to 2*SIFS plus transmission time for DATA and ACK.

15.6.1.2 Receiver-Initiated Opportunistic Media Access

RTR, as discussed in the literature,[2,8,18,23] is also a unicast packet. To reduce the control packets (i.e., RTR to probe channel and queue status, especially when the link condition changes significantly and/or each candidate sender has no packet with high probability when the receiver polls it), we propose to use the Multicast RTR to poll the candidate senders. A candidate sender list is included in the frame of Multicast RTR. The noise power level (dB) indicates the interference and power level at the receiver. Here we assume that link gain is symmetric and RTR is sent at the default power level. The candidate receivers can derive link gains according to the receiving power. By the link gain and nomic transmit power of DATA, the sender can calculate the expected receiving power at the receiver. Finally, the candidate senders can determine the average SINR with known interference power level.

Au: define SINR

Upon hearing the RTR, the candidate senders estimate the channel gain. The idle candidate sender with link gain better than certain level is allowed to access media. Similar to the sender-initiated strategy, IFS is employed to differentiate the media access priority of candidate senders. The IFS of the nth sender is equal to $SIFS + (n-1)^{*}Time\_slot$. The sender with the highest priority among those having good link conditions transmits data first. Similar to sender-initiated opportunistic media access, we propose to incorporate busy tone to enhance carrier sensing, even if it is not necessary. Thus, other intended senders would yield the opportunity to the one transmitting DATA first, that is, the one with the good channel condition and the highest priority.

15.6.2 Rate Adaptation

According to the channel condition evaluated by the physical layer analysis of Multicast RTS or Multicast RTR, SINR is determined and the appropriate modulation scheme can be selected to efficiently use the channel, as discussed in Section 15.5.

15.6.3 Scheduling

Here we discuss how to determine the candidate receiver (sender) list, the MPDU size, and the targeted data rate of each candidate directed link, which is closely related to QoS and energy efficiency. Considering that there are many constraints such as CPU and energy consumption for portable wireless device, one of the crucial requirements for the scheduling algorithm is simplicity. Many scheduling algorithms in the literature,[4] such as Round Robin (RR) and Earliest Timestamp First (ETF), can be tailored to our framework to achieve the desired goals. The scheduling policy in our simulation is based on the window-based weighted Round Robin. The targeted data rate is dynamically set using an algorithm similar to ARF.[11] A station is allowed to transmit multiple packets successively without contending for the media again after accessing the channel, as long as the total access time does not exceed a certain limit. We follow the thought of OAR[17] to grant channel access for multiple packets in proportion to the ratio of the achievable data rate over the basic rate so that the time-share fairness as in IEEE 802.11 can be assured. We show from our simulation study that both throughput and fairness can be significantly enhanced even by this simple scheduling method. We believe that fairness, QoS, and energy consumption can be further enhanced by more advanced scheduling algorithms.

15.7 Conclusion

This chapter discussed several important issues in MAC protocol design in IEEE 802.11 wireless LANs and mobile ad hoc networks, including severe MAC layer contention and collision in multihop environments, traffic flow contention, rate adaptation with dynamic fragmentation, multiple-input link diversity, and multiple-output link diversity. Then we proposed several novel schemes to address these issues, schemes that can greatly improve the performance of wireless networks in terms of throughput, end-to-end delay, fairness, stability, and scalability.

Acknowldegment

The authors would like to extend their appreciation to Byungseo Kim and Jianfeng Wang for their help during the preparation of this chapter.

References

1. P. Bhagwat, P. Bhattacharya, A. Krishna, and S. K. Tripathi, Enhancing Throughput over Wireless LANs Using Channel State Dependent Packet Scheduling, in *Proc. of INFOCOM'96*, 1998. **Au:** 96 versus 1998?
2. V. Bharghavan, A. Demers, S. Shenker, and L. Zhang, MACAW: A Media Access Protocol forWireless LANs, in *ACM SIGCOMM '94*.
3. J. Broch, D.A. Maltz, D.B. Johnson, Y. Hu, and J. Jetcheva, A Performance Comparison of Multihop Wireless Ad Hoc Network Routing Protocols, in *Proc. of the Fourth ACM International Conference on Mobile Computing and Networking (MobiCom'98)*, pp. 85–97, Oct. 1998.
4. Y. Cao and V.O.K. Li, Scheduling Algorithms in Broad-Band Wireless Networks, in *Proc. of the IEEE*, 89(1), January 2001.
5. K. Chen, Y. Xue, and K. Nahrstedt, On Setting TCP's Congestion Window Limit in Mobile Ad Hoc Networks, in *Proc. of IEEE International Conference on Communications (ICC 2003)*, May 2003.
6. Y. Fang and A.B. McDonald, Cross-Layer Performance Effects of Path Coupling in Wireless Ad Hoc Networks: Power and Throughput Implications of IEEE 802.11 MAC, in *Proc. of 21st IEEE International Performance, Computing, and Communications Conference*, April 2002.
7. Z. Fu, P. Zerfos, H. Luo, S. Lu, L. Zhang, and M. Gerla, The Impact of Multihop Wireless Channel on TCP Throughput and Loss, in *Proc. IEEE INFOCOMM'2003*, March 2003.
8. J.J. Garcia-Luna-Aceves and A. Tzamaloukas, Reversing the Collision-Avoidance Handshake in Wireless Networks, *Mobicom'99*.

9. Z.J. Haas and J. Deng, Dual Busy Tone Multiple Access (DBTMA) — A Multiple Access Control for Ad Hoc Networks, *IEEE Transactions on Communications*, 50(6), 975–985, June 2002.
10. *IEEE standard for Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications*, ISO/IEC 8802-11: 1999(E), Aug. 1999.
11. A. Kamerman and L. Monteban, WaveLAN II: A High-Performance Wireless LAN for the Unlicensed Band, Bell Labs Technical Journal, 1997.
12. B. Kim, Y. Fang, T. Wong, and Y. Kwon, Throughput Enhancement through Dynamic Fragmentation in Wireless LANs, submitted for publication.
13. B. Kim, Y. Fang, and T.F. Wong, Rate-Adaptive MAC Protocol in High-Rate Personal Area Networks, in *Proc. of IEEE Wireless Communications and Networking Conference 2004 (WCNC'04)*, Atlanta, GA, March 2004.
14. B. Kim, Y. Fang, T.F. Wong, and Y. Kwon, Dynamic Fragmentation Scheme for Rate-Adaptive Wireless LANs, *IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC'2003)*, Beijing, China, September 7–10, 2003.
15. J. Li, C. Blake, D.S.J. De Couto, H.I. Lee, and R. Morris, Capacity of Ad Hoc Wireless Network, in *Proc. of ACM MobiCom 2001*, July 2001.
16. C. Perkins, E.M. Royer, S.R. Das, and M.K. Marina, Performance Comparison of Two On-Demand Routing Protocols for Ad Hoc Networks, *IEEE Personal Communications*, pp. 16–28, Feb. 2001.
17. B. Sadeghi, V. Kanodia, A. Sabharwal, and E. Knightly, Opportunistic Media Access for Multirate Ad Hoc Networks, in *Proc. ACM MOBICOM' 02*, September 2002.
18. F. Talucci, M. Gerla, and L. Fratta, MACA-BI(MACA by invitation — A Receiver Oriented Access Protocol for Wireless Multihop Networks, *PIMRC'97.*
19. F.A. Tobagi and L. Kleinrock, Packet Switching in Radio Channels. Part II- The Hidden Terminal Problem in Carrier Sense Multiple-Access and the Busy-Tone Solution, *IEEE Trans. Commun.*, Vol. COM-23, pp. 1417–1433, Dec. 1975.
20. J. Wang, H. Zhai, and Y. Fang, OMAR: Utilizing Diversity in Wireless Ad Hoc Networks, submitted for publication.
21. J. Wang, H. Zhai, Y. Fang, and M.C. Yuang, Opportunistic Media Access Control and Rate Adaptation for Wireless Ad Hoc Networks, *IEEE International Conference on Communications (ICC'04)*, Paris, France, June 2004.
22. J. Wang, H. Zhai, and Y. Fang, Opportunistic Packet Scheduling and Media Access Control for Wireless LANs and Multi-hop Ad Hoc Networks, in *Proc. of IEEE Wireless Communications and Networking Conference (WCNC'04)*, Atlanta, March 2004.
23. Y. Wang and J.J. Garcia-Luna-Aceves, A New Hybrid Channel Access Scheme for Ad Hoc Networks, *ACM WINET*, Vol. 10, No. 4, July 2004.
24. S. Wu, Y. Tseng, and J. Sheu, Intelligent Medium Access for Mobile Ad Hoc Networks with Busy Tones and Power Control, *IEEE Journal on Selected Areas in Communications*, 18, 1647–1657, Sept. 2000.
25. S. Xu and T. Safadawi, Does the IEEE 802.11 MAC Protocol Work Well in Multihop Wireless Ad Hoc Networks?, *IEEE Communications Magazine*, pp. 130–137, June 2001.
26. H. Zhai and Y. Fang, Distributed Flow Control and Medium Access in Multihop Ad Hoc Networks, submitted for publication.
27. H. Zhai, J. Wang, and Y. Fang, Distributed Packet Scheduling for Multihop Flows in Ad Hoc Networks, in *Proc. of IEEE Wireless Communications and Networking Conference (WCNC'04)*, Atlanta, March 2004.
28. H. Zhai, J. Wang, and Y. Fang, DUCHA: A New Dual-Channel MAC Protocol for Multihop Mobile Ad Hoc Networks, submitted for publication.