WIRELESS COMMUNICATIONS AND MOBILE COMPUTING
Wirel. Commun. Mob. Comput. 2003; **3**:935–958 (DOI: 10.1002/wcm.118)

Medium access control protocols for wireless mobile ad hoc networks: issues and approaches[‡]

Teerawat Issariyakul,[1] Ekram Hossain[2,*,†] and Dong In Kim[3]

[1]*TRLabs Winnipeg, 100-135 Innovation Drive, Winnipeg, MB, R3Y 1P6, Canada (e-mail: teerawat@win.trlabs.ca)*
[2]*Department of Electrical and Computer Engineering at the University of Manitoba, Winnipeg, MB, R3T 5V6, Canada (e-mail: ekram@ee.umanitoba.ca)*
[3]*School of Engineering Science, Simon Fraser University, Burnaby, BC, V5A 1S6, Canada (e-mail: dikim@sfu.ca)*

Summary

In this article, a comprehensive survey of the medium access control (MAC) approaches for wireless mobile ad hoc networks is presented. The complexity in MAC design for wireless ad hoc networks arises due to node mobility, radio link vulnerability and the lack of central coordination. A series of studies on MAC design has been conducted in the literature to improve medium access performance in different aspects as identified by the different performance metrics. Tradeoffs among the different performance metrics (such as between throughput and fairness) dictate the design of a suitable MAC protocol. We compare the different proposed MAC approaches, identify their problems and discuss the possible remedies. The interactions among the MAC and the higher layer protocols such as routing and transport layer protocols are discussed and some interesting research issues are also identified. Copyright © 2003 John Wiley & Sons, Ltd.

KEY WORDS: mobile ad hoc networks; medium access control (MAC); differentiated services MAC; collision avoidance; fairness; energy efficiency

1. Introduction

The simplicity in deployment makes mobile ad hoc networks (which do not require any infrastructure) suitable for a variety of applications such as collaborative computing, disaster recovery, battle field communication. With the proliferation of communications and computing devices such as mobile phones, laptops, or PDAs, personal area networking (PAN), which is an ad hoc networking-based technology, has recently gained much interest.

Medium access control (MAC) protocols play an important role in the performance of the *mobile ad hoc networks* (MANETs). A MAC protocol defines how each mobile unit can share the limited wireless bandwidth resource in an efficient manner. Recent

---

*Correspondence to: Ekram Hossain, Department of Electrical and Computer Engineering, University of Manitoba, Winnipeg, MB, R3T 5V6, Canada.

[†] E-mail: ekram@ee.umanitoba.ca

[‡] This work was supported in part by the Telecommunications Research Labs (TR*Labs*), Winnipeg, Canada and in part by a Grant from the Natural Sciences and Engineering Research Council (NSERC) of Canada.

Contract grant sponsors: Telecommunication Research Labs (TRLabs); NSERC.

researches in this area have focused on designing MAC protocols with optimized performance metrics including throughput and delay, fairness, stability, support for multimedia and energy efficiency.

The organization for the rest of the article is as follows. Section 2 describes the general concepts and the performance metrics of MAC protocols for MANETs. The problems and the issues related to designing MAC protocols for MANETs are described in Section 3. Section 4 briefly describes the MAC protocol for IEEE 802.11 WLAN in the ad hoc mode of operation. The different approaches to MAC design are presented in Section 5 in a comprehensive fashion. In Section 6, observations and recommendations about optimizing MAC performance and a qualitative comparison among the different approaches are presented. In Section 7, the issue of inter-layer protocol interaction is discussed and some other research issues are identified.

2. MANET Topology and MAC Performance Metrics

A mobile ad hoc network is a collection of heterogeneous communications and computing devices which can communicate with one another (within their transmission range) without any central coordination (Figure 1). It is essentially infrastructureless and there is no need for any fixed radio-base station or router. Such a network is *self-organizing* and *adaptive*.

In general, communication between a *Base Station* (BS) and a mobile node in a centralized architecture is performed by using a deterministic MAC protocol such as *frequency division multiple access* (FDMA), *time division multiple access* (TDMA), or *code division multiple access* (CDMA). Due to the lack of centralized control in a mobile ad hoc network, a distributed MAC protocol, such as ALOHA or *carrier-sense multiple access* (CSMA), which utilizes only information local to each node, is used to control access of mobile nodes to a shared wireless channel. The protocol should try to reduce the number of possible collisions among transmissions from neighboring nodes and thereby increase the channel utilization. In addition, efficient use of the limited battery power, service fairness and provision of Quality of Service (QoS) are some of the other issues which are needed to be taken into consideration. In fact, finding an optimal MAC mechanism in an ad hoc network often results in *NP-complete*[§] problems such as a graph-coloring problem.

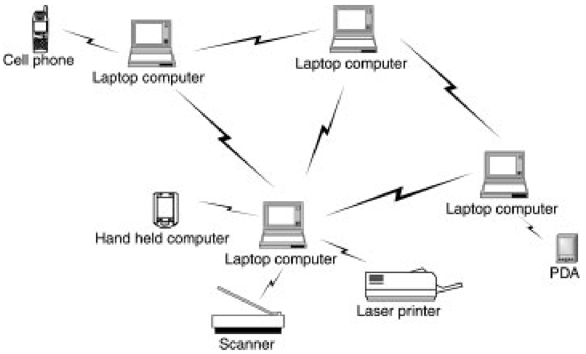


Fig. 1. Mobile ad hoc network architecture.

While evaluating an MAC protocol for a wireless mobile ad hoc network, the following performance measures should be considered:

- *Throughput and delay*: Throughput is generally measured as the percentage of successfully transmitted radio link level frames per unit time. Transmission delay is defined as the interval between the frame arrival time at the MAC layer of a transmitter and the time at which the transmitter realizes that the transmitted frame has been successfully received by the receiver.
- *Fairness*: Generally, fairness measures how fair the channel allocation is among the flows in the different mobile nodes. The node mobility and the unreliability of radio channels are the two main factors that impact fairness.
- *Energy efficiency*: Generally, energy efficiency is measured as the fraction of the useful energy consumption (for successful frame transmission) to the total energy spent.
- *Multimedia support*: It is the ability of an MAC protocol to accommodate traffic with different service requirements such as throughput, delay and frame loss rate.

3. Issues in Designing MAC Protocol for MANETs

In a mobile ad hoc network, node mobility, vulnerability of the radio channel(s) and the lack of any central coordination give rise to the following

---

[§] No polynomial time algorithm exists to solve an NP-complete problem.

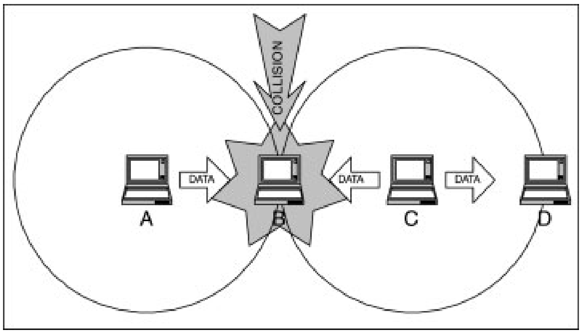


Fig. 2. Illustration of the hidden node problem.

problems which need to be taken into consideration while designing an MAC protocol.

- *Hidden node problem*: A hidden node is a node which is out of range of a transmitter node (node A in Figure 2), but in the range of a receiver node (node B in Figure 2) [1]. A hidden node does not hear the data sent from a transmitter to a receiver (node C is hidden from node A). When node C transmits to node D, the transmission collides with that from node A to node B. Obviously, the hidden nodes lead to higher collision probability. Generally, the probability of successful frame transmission decreases as the distance between source and destination increases and/or the traffic load increases [1].
- *Exposed node problem*: An exposed node (node C is exposed to B in Figure 3) is a node which is out of range of a receiver (node A), but in the range of the corresponding transmitter (node B). Node C defers transmission (to node D) upon detecting data from node B, even though a transmission from node C does not interfere with the reception at node A. The link utilization may be significantly impaired due to the exposed node problem. This would impact the higher layer protocol (e.g. TCP) performance considerably.
- *Radio link vulnerability*: Wireless channel capacity is limited due to high bit-error rate. The high bit error rate in a wireless environment can be attributed to the effects such as noise, interference, free-space loss, shadowing and multipath fading. Again, the channel errors are location dependent and bursty in nature. The radio link vulnerability may tremendously impact the utilization of the radio channel(s) and the service fairness among different mobile nodes (and flows).

  To deal with the radio link vulnerability, schemes such as *forward error correction* (FEC) and *automatic repeat request* (ARQ) have been developed. Unfortunately, they result in inefficient bandwidth utilization. Again, increase in transmission power to combat with the above undesirable radio propagation properties can broaden interference region, thereby resulting in the reduction of spatial reuse.
- *Capture problem*: Capture is an ability of a mobile node to perfectly receive a signal (presumably one with the dominating signal level) in the presence of more than one simultaneous transmissions. In fact, capture effect is a favorable feature in that it improves the utilization of the channel [2], but it may cause unfairness among mobile nodes.

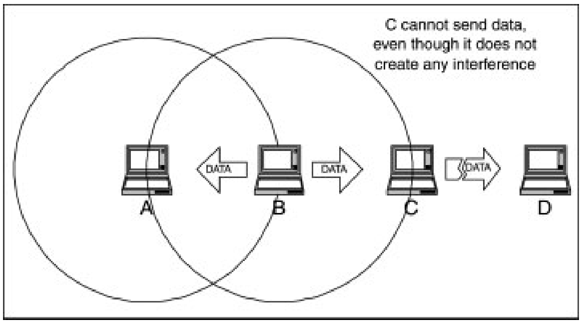


Fig. 3. Illustration of the exposed node problem.

4. IEEE 802.11 WLAN: Ad Hoc Mode of Operation

Known as *distributed foundation wireless medium access control* (DFWMAC), the 802.11 MAC with *distributed coordination function* (DCF) is based on CSMA with *collision avoidance* (CA).

CSMA/CA was developed to overcome the hidden node problem. It incorporates a handshake protocol in the original CSMA protocol. In CSMA/CA, a sender must first transmit a *request to send* (RTS) frame. RTS contains the identification of the receiver so that only the intended receiver will answer this message with a clear to send (CTS) frame. Other mobile nodes intercepting either RTS or CTS defer their transmission for the period specified by the network allocation vector (NAV) (Figure 4) in the handshaking frames RTS and CTS. Therefore, the number of hidden nodes is reduced by some degree. The protocol is described through pseudo-code in Appendix A.

Each frame in DFWMAC is separated with an interval called inter-frame space (IFS). There are four classes of IFS: short *IFS* (SIFS), *point*

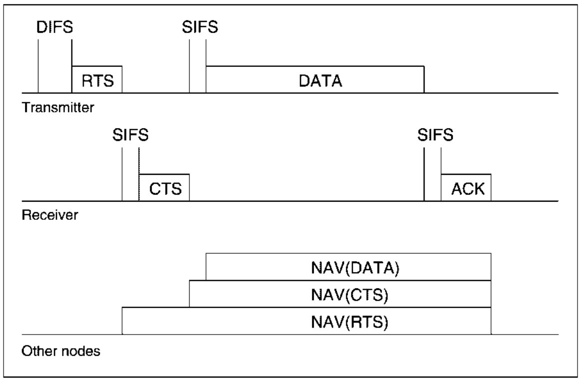


Fig. 4. Collision avoidance mechanism in IEEE 802.11.

coordination function IFS (PIFS) *distributed coordination function IFS* (DIFS) and *extended IFS* (EIFS), each of which is utilized in different situations [3]. Only when the medium is idle for more than IFS will the mobile realize that the medium is idle. Each mobile node is allowed to transmit whenever it realizes that the medium is idle.

Before sending an RTS frame, a mobile unit has to sense the medium for DIFS interval. If the channel is free, the mobile unit will start sending RTS. Upon receiving RTS, the receiver senses the channel for SIFS interval and sends a CTS frame if the channel is free. The transmitter and the receiver send DATA and ACK frames respectively, if the channel is free for SIFS interval (Figure 4).

If a mobile that has sent RTS or DATA does not receive CTS or ACK before timeout, it will initiate a back-off process. In this process, the mobile first generates the back-off time as a number of time slots uniformly distributed over [0, CW], where CW is the current contention window parameter. The mobile counts down the back-off time only if the medium is idle. Note that, in order to realize an idle medium a mobile has to sense that the medium is idle for DIFS. When the channel becomes idle for more than DIFS, the mobile counts down the back-off time and it stops counting down when the channel becomes busy again. The process is repeated until the back-off time reaches zero. Then the mobile will initiate transmission in the next time slot without waiting for DIFS.

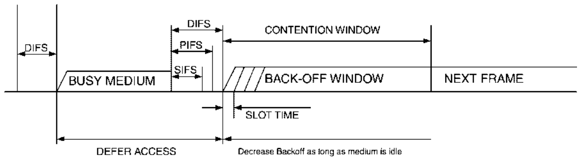


Fig. 5. Illustration of IFS and back-off.

Initially, the contention window starts with its minimum value ($CW_{\min}$). It is doubled after every collision and stops increasing after reaching its maximum value ($CW_{\max}$). The contention window is reset to $CW_{\min}$ once the transmission is successful.

Unlike the back-off time, NAV always decreases regardless of medium state, because it is the expected time required for data frame transmission. Mobile nodes aware of NAV can turn off their transceiver during this period to save battery power. DFWMAC sets shorter SIFS interval for transmission of the subsequent frames and, thereby, give priority to the mobile nodes which have acquired the channel. A mobile node that has sent RTS senses the channel for shorter time and has more probability to acquire the channel than others.

Unfortunately, CSMA/CA does not solve the hidden node problem entirely. For example, in Figure 6, when node C (which is hidden from the transmitter node A) moves into the communication area after the handshake procedure between node A and node B is completed and initiates data transmission, collision occurs at the active receiver.

CSMA/CA exacerbates the exposed node problem, because more mobile nodes are discouraged to send data. The exposed node problem prevents DFWMAC to work very well with transmission control protocol (TCP) [4]. In particular, only one TCP connection can exist in the same area.

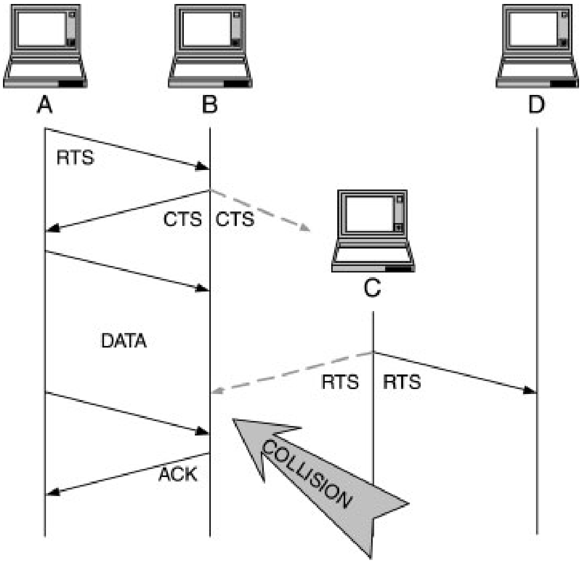


Fig. 6. Hidden node problem in CSMA/CA protocol.

5. Approaches for Designing MANET MAC Protocols

5.1. Using Handshaking Signaling

5.1.1. CSMA/CA and it's derivatives

To reduce the impact of hidden nodes, CSMA/CA was adopted in IEEE 802.11. It is recommended that a threshold should be setup, and the handshake signaling should be used only for comparatively long data [3]. However, in heterogeneous data traffic scenario where more than one class of traffic exist, the handshake protocol always gives higher throughput [5].

Multiple access with collision avoidance (MACA) [6] was proposed to resolve the hidden node and the exposed node problems and also to provide the ability to perform per-frame transmission power control. When a mobile node overhears some RTS/CTS frames corresponding to transmissions in other nodes, it does not need to remain completely silent and it can communicate with other neighboring nodes with lower transmission power.

MACA does not use carrier sensing. After waiting for a random number of time slots (where each time slot is equal to the time required for the RTS frame to reach the intended receiver), a mobile node transmits regardless of the medium status. In MACA, a three way handshake RTS-CTS-DATA is adopted. Note that, there is no ACK frame. It is assumed that the MAC layer frame loss will be taken care of by some higher layer protocols. Some mobiles overhearing RTS must defer their transmission until CTS is received by the sender. In order to avoid interference to data frame transmission, other mobiles intercepting CTS postpone their transmission until DATA is received by the receiver. This can be achieved by using a time parameter in the header of RTS and CTS. MACA uses the binary exponential back-off (BEB) algorithm and the back-off time always decreases regardless of the channel status [6].

Like MACA, the floor acquisition multiple access (FAMA) [7] employs RTS-CTS-DATA handshake signaling. FAMA allows multiple frame transmission by having a sender set *MORE* flag in data header if it has more data frames to send. The receiver replies with CTS immediately after obtaining a DATA frame with a *MORE* flag set. FAMA reduces the impact of hidden nodes by forcing a mobile node that has just turned on to sense the medium for some time. In FAMA, the back-off timer counts down only when the medium is idle. If transmission destined to other nodes is detected, new waiting time will be generated depending on the transmission type (as in Equation (1)). This waiting time decreases when the medium becomes idle.

$$T_{\text{wait}} = \begin{cases} 2 \cdot T_{\text{PROP}} + T_{\text{PROC}} + T_{\text{TR}} + T_{\text{DATA}}, \\ \text{initial waiting time} \\ 2 \cdot T_{\text{PROP}} + T_{\text{PROC}} + T_{\text{TR}} + T_{\text{CTS}}, \\ \text{RTS detected} \\ 2 \cdot T_{\text{PROP}} + T_{\text{PROC}} + T_{\text{TR}} + T_{\text{DATA}}, \\ \text{CTS detected} \\ 2 \cdot T_{\text{PROP}} + T_{\text{PROC}} + T_{\text{TR}} + T_{\text{CTS}}, \\ \text{more data} \\ 2 \cdot T_{\text{PROP}} + T_{\text{PROC}} + T_{\text{TR}}, \ \text{no more data} \\ 2 \cdot T_{\text{PROP}} + T_{\text{PROC}} + T_{\text{TR}} + T_{\text{DATA}}, \\ \text{unidentified} \end{cases} \tag{1}$$

where T_{PROP} = maximum propagation time, T_{PROC} = processing time, T_{TR} = turn-around time, T_{DATA} = time required to send DATA, T_{RTS} = time required to send RTS and T_{CTS} = time required to send CTS.

The frame length of CTS in FAMA is longer than that of RTS by at least $2 \times T_{\text{PROP}}$. After sending RTS, if a mobile detects any noise or an unidentified frame, it assumes a collision and initiates a back-off process. The back-off time is uniformly distributed between 1 and $10 \times T_{\text{CTS}}$. Again, the back-off timer is activated only when the medium is idle. As soon as the medium becomes busy, waiting time is replaced by the new timer value generated according to Equation (1). This approach is called FAMA with non-carrier sensing (FAMA-NCS).

All of the above protocols use *non-persistent* channel access for the transmission of collision-avoidance handshake frames, that is, in the case of collision during handshake signaling, a mobile backs off for a random amount of time and attempts to transmit at a later time. Instead, *persistent* channel access would be more efficient under light traffic loads. It was shown that introducing a limited window of persistence in collision-avoidance handshake signaling could provide much higher throughput even when the average offered load is moderately high [8].

In limited-persistence carrier sensing (FAMA-LCS), when a mobile wants to transmit RTS and finds the medium busy, it will keep monitoring medium for *persistence time* [8]. If the medium becomes idle before the mobile gives up, the mobile will transmit RTS. Otherwise, a back-off procedure will be initiated.

MACAW is another derivative of the CSMA/CA protocol which uses RTS-CTS-DS-DATA-ACK

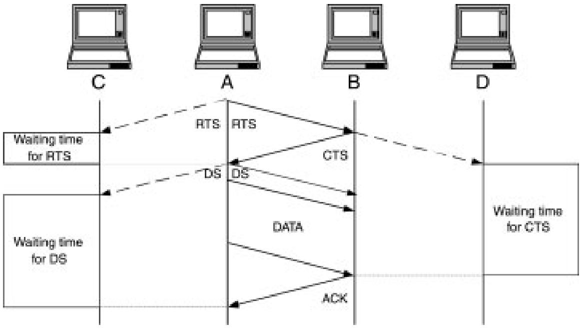


Fig. 7. The normal operation of MACAW.

handshake signaling (Figure 7) [9]. The ACK frame (which is not used in MACA) is used to ensure the MAC level reliability. Before a DATA frame is transmitted, a *data sending* (DS) control frame containing the duration of DATA and ACK frames is sent to inform the mobile nodes in the range of a sender not to transmit until the ACK is received by the sender.

The waiting time in this case is determined based on the receiver. For example, in Figure 7, node A sends an RTS to node B and node C overhearing this RTS defers transmission until node A receives CTS. Overhearing the CTS frame, node D defers all transmission up to the point at which node B receives all DATA. Again, upon receiving CTS, node A transmits a *DS* frame to notify other mobiles to defer their transmission until it receives the ACK.

As shown in Figure 8(a), if a receiver cannot reply to a RTS frame because of the current waiting time, it transmits a *Request for RTS* (RRTS) frame as soon as the waiting time expires. The sender then initiates normal collision avoidance handshake signaling again. RRTS enables the sender to know when the receiver is ready to transmit CTS.

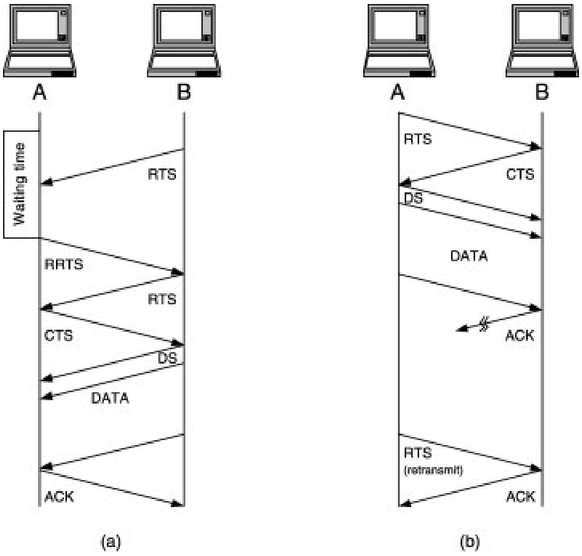


Fig. 8. Solutions of CTS blocking and loss of ACK in MACAW.

In the case that DATA is received successfully at the receiver but the ACK is lost (Figure 8b), the sender will initiate retransmission, and the receiver will reply with ACK rather than CTS. The back-off time is not changed in this case. Nevertheless, if CTS or ACK is not received by the sender before time-out, the sender will initiate the back-off process and retransmit when the timer reaches zero. In MACAW, the back-off time starts with the minimum value, increases by 1.5 times for every frame loss until it reaches the maximum. In case of successful transmission the back-off value is decreased by 1. Mobile nodes also broadcast their back-off values so that all nodes in the same vicinity can contend for the channel in fair manner.

The above protocol can be improved by sporadically omitting ACK while leaving the level of MAC reliability intact. In practice, an *MAC service data unit* (MSDU) is fragmented into several *MAC protocol data units* (MPDUs). That all MPDUs from the same MSDU are destined to the same receiver enables the MAC protocol at the receiver to omit sending ACK for all but the last MPDU. *Negative acknowledgement* (NACK) can be used as the notification of the transmission failure. In Figure 9(a), after the receiver B sends out a NACK or CTS, it expects a DATA frame. If B does not receive the DATA frame correctly, it sends an NACK after it receives the next RTS. Otherwise, it sends a CTS frame. Upon receiving an NACK,

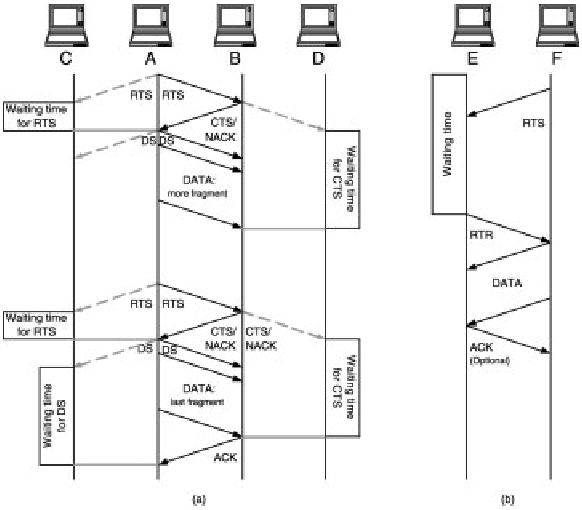


Fig. 9. Possible improvement of the MACAW protocol.

node A retransmits the previously transmitted frame. Node A will send the next MPDU, if it receives CTS instead. Except for the last MPDU, node C can send RTS or NACK/CTS out as soon as the waiting time for RTS expires (Figure 9a). The reduced waiting time would presumably improve channel utilization. For each lost acknowledgement message, this approach can reduce unnecessary transaction either by $(T_{\mathrm{RTS}} + T_{\mathrm{CTS}} + T_{\mathrm{DATA}} + T_{\mathrm{ACK}} + 3 \cdot \mathrm{SIFS} + \mathrm{DIFS})$ or $(T_{\mathrm{DATA}} + T_{\mathrm{ACK}} + 2 \cdot \mathrm{SIFS})$ depending on whether there are more frames to be transmitted to this receiver or not (Figure 4).

Another improvement is to omit RTS and CTS in the presence of RRTS. This concept is similar to the polling concept used in receiver initiated channel hopping with dual polling (RICH-DP). In this case, RRTS is analogous to ready to receive (RTR) in RICH-DP (to be described later in this article).

5.1.2. Multichannel CSMA

The original CSMA protocol was developed for a single-channel system. Multichannel CSMA was introduced in order to increase the number of simultaneous active users. This approach was applied to both direct-sequence spread-spectrum (DSSS) and frequency-hopping spread spectrum (FHSS) systems [10].

Common-transmitter-based multiple access with collision avoidance (MACA-CT) and receiver-transmitter-based multiple access with collision avoidance (MACA-RT) are two DSSS approaches [11]. In these approaches, a set of orthogonal codes is predefined and known to every node. In MACA-CT, control and data frames are sent using common code and transmitter code respectively (Figure 10a). The code to be used for data frame transmission is contained in the RTS. Upon receiving RTS the receiver tunes itself to the transmitter code. Other mobile nodes (e.g. node C) which want to initiate transmission use the common code and, therefore, do not interfere with the ongoing transmissions. Collision can occur only in the common control channel and the vulnerable period is the double of the duration of a control frame.

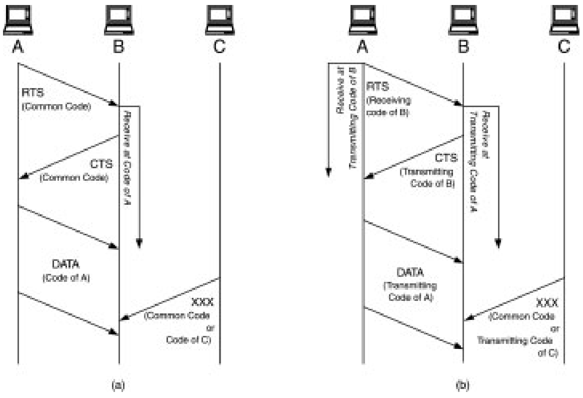


Fig. 10. Operation of (a) MACA-CT and (b) MACA-RT protocols.

With comparatively high persistence probability, throughput of MACA-CT can be more than six times that in the conventional CDMA approach [11]. The system requires $n + 1$ codes: n codes for the n mobile nodes and one for the common control channel.

In the MACA-RT approach (Figure 10b), each mobile node has two codes: transmitting code and receiving code. First, RTS is sent using the destination's receiving code. Next, CTS is sent back using the transmitting code of the destination. Finally, the data transmission is conducted using the transmitting code of the originator. This approach totally eliminates collision because each mobile uses different codes to send control and data frames. This approach requires $2 \times n$ codes—two codes for each mobile. In addition, every mobile has to recognize codes of each other so as to initiate transmission using a receiver's code. When the number of mobiles increases, available codes might not be enough for every mobile and dynamic code allocation should be used in this situation.

Receiver initiated channel hopping with dual polling (RICH-DP) utilizes the FHSS concepts [12]. This protocol is based on a receiver-initiated collision avoidance handshake. In RICH-DP, all the nodes employ the same hopping sequence and the duration of each hop is the amount of time needed for nodes to receive a control frame from a neighbor node.

A polling node (receiver) sends a ready to receive (RTR) control frame (which is analogous to RTS frame) over the current channel hop and upon successful reception of the frame the intended sender starts sending a data frame over the same channel hop (Figure 11). After the polled node completes transmitting the data frame (or it does not have data to transmit at all in which case it transmits a CTS frame), the polling node can start transmitting it's own data to the polled node over the same channel (dual polling). ACK frames are used to acknowledge successful data transmission (Figure 12). After all the data transmissions are complete, both the nodes can resynchronize to the common channel hop. The protocol is described through pseudo-code in Appendix B.

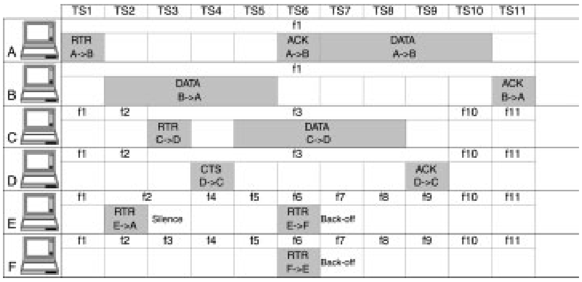
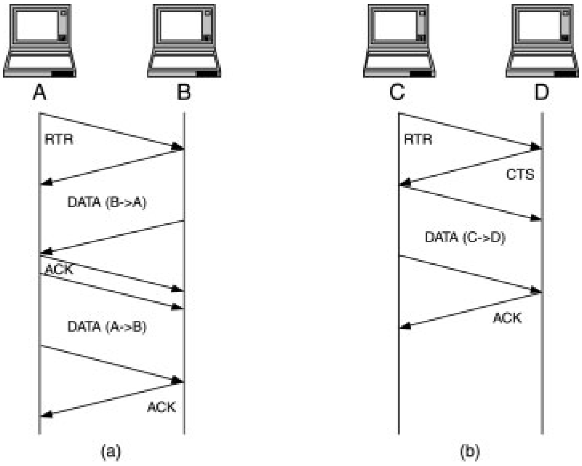


Fig. 11. Operation of the RICH-DP protocol.

This approach reduces the vulnerable period from two control frame periods to one. In other words, collision of control frames can occur only during transmission of RTR. The maximum throughput of RICH-DP was observed to be always higher than MACA-CT, the sender-initiated CDMA protocol, regardless of the size of the data frame length (under perfect channel conditions). However, when a minuscule fraction of mobile nodes are ready to send or receive data, polling inactive nodes leads to bandwidth inefficiency.

In DSSS-based packet radio networks, variable rate transmission can be achieved at the MAC level by using *orthogonal variable spreading factor* (OVSF) codes. Channel allocation (i.e. either frequency or code assignment) can be static or can be dynamic based on channel acquisition. Despite their higher utilization, the on-demand schemes based on channel acquisition are more complex and incur higher overhead.

Due to the limited wireless spectrum, high rate transmission in a spread-spectrum based ad hoc radio network can be achieved by using spectrally efficient higher order modulation schemes such as M-ary quadrature amplitude modulation (QAM). However, higher order modulation schemes require high signal-to-interference ratio (SIR) to maintain the link reliability, which would demand more battery-power at the portable mobile units. Again, with limited available bandwidth in a DSSS system, the *spreading factor*[¶] (SF) would be reduced with increased bit transmission rate. Then, even if the signal power is high, the throughput performance may degrade significantly due to multipath-induced interference. Therefore, multipath-intereference cancellers would be required to accommodate higher order modulation schemes in high-speed spread-spectrum ad hoc radio networks.



(a) (b)

Fig. 12. Data transmission signaling in RICH-DP protocol.

MAC level frame scheduling based on the radio link conditions experienced by each mobile (rather than individual user demand) can provide higher throughput in a spread-spectrum ad hoc radio network. But this may aggravate the fairness problem since the mobiles with more favorable link conditions may be allocated higher transmission rates, and the others may not be able to transmit at all.

5.2. Common Control Channel Signaling (CCCS) Approach

In this approach, control and data traffic are transmitted in different channels. One of the early implementations of this concept was by broadcasting a busy tone as in the busy tone multiple access (BTMA) protocol [13]. BTMA utilizes an out-of-band channel to transmit the busy tone signal. Every mobile node which is less than two hops away from the source and hears the busy tone transmits more busy tone to the channel, thus expanding busy region of the channel.

The key idea behind this approach is the fact that the channel holding time of control traffic in CSMA/CA is very short. In conventional CSMA/CA, as shown in Figure 13, the exposed node C can send neither RTS nor CTS to node D. After the handshake period, transmission from node D to node C will not interfere with the reception at node B. Spatial reuse (e.g. data transmission from node D to node C) can be greatly increased if a separate channel is dedicated to transmit control frames.

5.3. Using Adaptive Transmission Range/Power Control

When the source and the destination nodes are far apart, a multihop end-to-end protocol would be

---

[¶] It is defined as the ratio of the chip rate to the information bit rate.

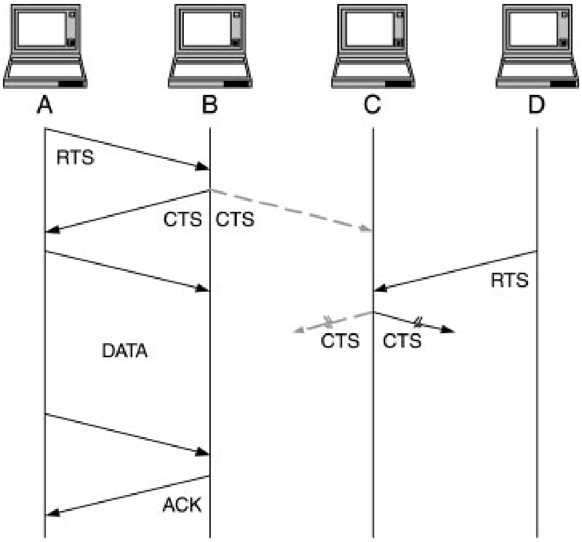


Fig. 13. The exposed node problem in CSMA/CA.

required for data transmission. The use of a longer transmission range at the MAC level increases the likelihood of finding the destination mobile node in a single hop or at least reduces the number of hops between the source and the destination (and hence the transmission failure probability due to some intermediate node or link failure). However, it can also lead to interference in wider area resulting in more exposed nodes [14].

Several strategies based on the transmission range control can be adopted for MAC. The range of transmission can be expressed in terms of the number of mobile nodes within the transmission radius of a sender and an optimal value of this range which maximizes the expected one-hop progress in the desired direction may exist.

In the case of the most forward with fixed radius (MFR) strategy, where the intended recipient is the node whose location gives the longest projection upon the displacement between source and destination, the optimal transmission range with slotted ALOHA-based access was observed to be *eight* mobile nodes on the average under high traffic condition [14]. In the case of nearest with forward progress (NFP) strategy, a mobile node adjusts it's transmission power to be just enough to reach the neighbor which results in nearest forward progress [15]. The most forward with variable radius (MVR) approach is similar to MFR approach except that the transmission radius is adjusted to be equal to the distance between the transmitter and the receiver. The NFP strategy was observed to provide the best performance among all in terms of both throughput and normalized average progress and MVR performed better than MFR. NFP was found to have the best normalized average progress when the transmission range was greater than eight, while for transmission range less than eight, NFP and MVR achieved almost the same normalized average progress.

When CSMA is used instead of ALOHA, the performance improves. It was observed that the optimized expected progress with non-persistent CSMA under the MFR approach was 16% higher than that for slotted ALOHA (for zero propagation delay) [14].

Collision avoidance signaling can be combined with the concept of transmission range/power control to improve the performance further. The power controlled multiple access (PCMA) protocol [16] uses handshake signaling, common control channels and power control. The concept of variable bounded power was introduced to relax the constraints due to a single channel. The sender transmits a request-power-to-send (RPTS) frame (which is analogous to the RTS frame) to the receiver. Upon receiving RPTS, the destination calculates the minimum acceptable power, and sends it using an APTS frame (which is similar to the CTS frame) back to the source (Figure 14).

The source node then transmits the data frames to the sink node at the power level as advertized by the receiver. While receiving the data, the sink node calculates the tolerable noise and interference power level, and then broadcasts it using a busy tone in the common control channel. Other mobile nodes transmit data with power level as calculated from the power level advertised in the busy tone. As in

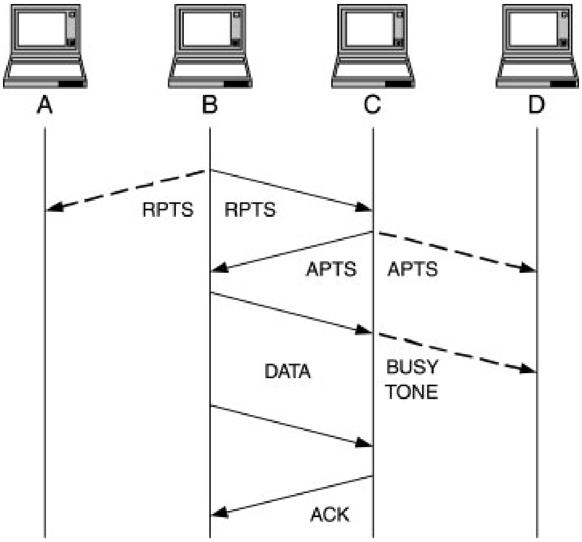


Fig. 14. Operation of the PCMA protocol.

BTMA, a channel is dedicated to busy tone transmission. The busy tone is periodically transmitted to the common channel to prevent the new nodes moving into the active floor from transmitting with unacceptably high power level.

Since only receivers are responsible for sending the busy tone, this approach avoids signal collision at the receivers without causing exposed node problem. This approach was shown to increase the system throughput by as much as 50% compared to that for IEEE 802.11 when the flow rate of each user was high and their spatial distribution was non-uniform.

5.4. Using Directional/Adaptive Antennas

The impact of hidden node problem can be mitigated to some extent by using directional antennas instead of an omnidirectional antenna. A mobile node equipped with several antennas to cover all directions can receive multiple data frames simultaneously.

Protocols similar to CSMA/CA can be used in a directional MAC protocol, but on a per-antenna basis. If an antenna, say, T at node X, has received an RTS or CTS corresponding to data transfer between two other nodes, then it will be 'blocked' for the duration of that transfer. To increase the chance of finding the destination node, RTS can be sent initially using all antennas. If any one of the antennas in a mobile node is blocked, RTS can be sent from another antenna. After the RTS-CTS signaling, both nodes are aware of the location of each other and thus only one directional antenna is sufficient to send DATA and ACK. Each antenna performs operations such as carrier sensing or back-off independently. Since transmitted data does not spread out throughout the region, wider area is safe from signal contamination. Compared with that of 802.11 WLAN, the TCP throughput increases by about 31.43% when 4 directional antennas are implemented per node [17].

5.5. Fairness-Based Approaches

A fair wireless MAC protocol ensures fair competition among the mobile nodes for the wireless channel bandwidth in the presence of time and location dependent wireless channel errors. When the different mobile nodes require different levels of service, the resources need to be allocated in proportion to the 'weight' of each node to provide a service differentiation. From the perspective of MAC protocol *fairness* in a mobile ad hoc network, the following issues are important:

- *Node fairness and flow fairness*: Node fairness is achieved when the mobile nodes can share the channel bandwidth according to their corresponding weights. Flow fairness is ensured when the different flows get fair share of the channel bandwidth in an *end-to-end* basis. To achieve flow fairness, mobile nodes should provide fair services to all the flows passing through them. Therefore, to ensure flow fairness the MAC protocols will need to utilize the end-to-end flow information. Per-node (common) queuing and back-off mechanisms may treat the different flows unfairly, because in these cases the flows in a node cannot be differentiated. Per-flow queuing and back-off mechanisms would be required to improve fairness.
- *Spatial distribution of mobile nodes*: Flows traversing rather sparsely populated node areas would presumably achieve better performance than those passing through 'congested' regions. Therefore, when the distribution of the mobile nodes is non-uniform, the asymmetry in distribution may incur node unfairness as well as flow unfairness.
- *Trade-off between fairness and channel utilization*: Providing fairness may lead to decrease in channel utilization. For example, in Figure 15, if *Flow3* stops sending, more channel utilization can be achieved at the expense of reduced flow fairness.
- *Unreliability of radio channel*: Wireless channel impairments, such as noise and interference, aggravate the fairness problem. During the bad channel state, if a mobile node stops sending data so that the other mobile nodes can send, more utilization can be achieved, but node fairness and flow fairness are penalized.

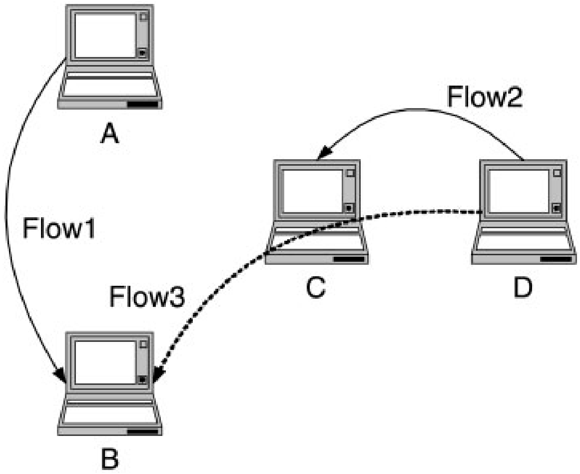


Fig. 15. Trade-off between fairness and channel utilization.

- *Short-term and long-term fairness*: In wireless environment, short-term fairness is difficult to attain due to radio channel unreliability and node mobility. Better short-term fairness can be provided when link failures are scarce. For the mobile nodes experiencing adverse channel conditions, long-term fairness can be provided by using some compensation mechanism.

5.5.1. Utility-based MAC approach for flow fairness

In general, the principle of utility-based approaches is to maximize the total *utility* subject to link capacity. Utility for a flow f (U_f) refers to the performance achieved by this flow in a network. The utility function can be different from flow to flow. The design objective is always to maximize total utility or *efficacy* (V). Utility function depends on the design objective. For example, if the objective is to maximize the channel utilization, the utility function for each mobile node can be defined as a linear function of throughput and then the efficacy is the total throughput.

Utility can be a function of several parameters. If we assume that it is a function of bit rate (x_f), and the general objective is to:

$$\begin{aligned} \text{maximize} \quad & V = \sum_{f \in F} U_f(x_f) \quad \text{subject to} \\ & \sum_{f \in F_l} x_f \le c_l \quad \text{and} \quad x_f \ge 0, \\ & \forall l \in L, \quad x_f \in X_F \quad \forall f \in F \end{aligned} \tag{2}$$

where c_l is the capacity of link l, L is the set of links in the network, F is the set of all the flows in the network, F_l is the set of flows traversing link l and X_F is the set of bit rates corresponding to the flows in the set F.

Then, the fairness problem reduces to a constrained optimization problem. The solution in terms of a rate adaptation algorithm can be achieved by using a subgradient-based algorithm [18] as follows:

$$x_f^{(n+1)} = \begin{cases} x_f^{(n)} + \alpha_n U'(x_f^{(n)}), & e^{(n)} = 0 \\ x_f^{(n)} - \beta_n e^{(n)}, & e^{(n)} > 0 \end{cases} \tag{3}$$

where $e^{(n)}$ is the number of congested links corresponding to flow f in nth iteration. If the sequences α_n and β_n satisfy the following conditions, the solution will converge within a finite number of iterations.

$$\lim_{n\to\infty} \alpha_n = 0 \qquad \sum_{n=1}^{\infty} \alpha_n = \infty \tag{4}$$

$$\lim_{n\to\infty} \beta_n = 0 \qquad \sum_{n=1}^{\infty} \beta_n = \infty \tag{5}$$

$$\lim_{n\to\infty} \frac{\alpha_n}{\beta_n} = 0 \tag{6}$$

Alternatively, the *penalty method* can be used to modify the objective function as follows:

$$\begin{aligned} \text{maximize} \quad & \alpha \cdot U(x_f) - \beta \cdot p_{loss} \cdot x_f \ \text{subject to} \\ & \sum_{f \in F_l} x_f \le c_l \quad \text{and} \quad x_f \ge 0, \\ & \forall l \in L, \quad x_f \in X_F \quad \forall f \in F \end{aligned} \tag{7}$$

where α and β are 'utility constant' and 'penalty constant', respectively, and p_{loss} is the contention loss probability [19]. The solution of the above problem is to adapt the persistence probability according to the following:

$$p_f^{(n+1)} = \begin{cases} p_f^{(n)} + \alpha_n - \frac{\beta_n}{\frac{\partial}{\partial x_f} U(x_f)}, \\ \quad \text{if collision occurs} \\ p_f^{(n)} + \alpha_n, \quad \text{otherwise} \end{cases} \tag{8}$$

Therefore, the adjustment in persistence probability is one solution to the fairness problem when a fairness objective is identified in terms of a utility function.

Note that, this general framework for utility-based fairness is very useful in that once the fairness objective is established, the implementation can be easily achieved by the algorithms using local parameters at each mobile node. When $U(x_f) = \lg x_f$, the above framework results in proportional fairness [20] and with this utility function the variance of flows experiencing different contention environment reduced from 0.2003 to 0.1617 in a in IEEE 802.11 WLAN [19].

5.5.2. Link-state dependent fairness

Link-state dependent approaches try to achieve fairness through scheduling mechanisms that utilize the channel state information. The server based fairness approach (SBFA) [21], which was originally designed for a centralized wireless system, can be applied in MANETs when a pseudo base station (PBS) is assigned for a cluster of mobile nodes. In this approach, each mobile node checks the channel condition before transmission and it postpones the transmission if the channel is bad. A centralized scheduler keeps track of transmissions by each mobile, and compensates the deferred mobiles by allocating extra bandwidth later when the channel condition becomes better. The identities of the deferred mobile nodes are kept at the long time fairness server (LTFS). The scheduler treats the LTFS as an additional flow in the network.

Although this approach increases channel utilization, it might have an adverse impact on the frame transmission delay. Therefore, rather than postponing transmission in the case of bad channel condition, frames can be transmitted with lower data rate. This is the dynamic link adaptation concept which can be implemented through dynamic channel coding and/or adaptive modulation. In case of dynamic channel coding, the length of the error correcting codes can be determined based on the channel condition and the required QoS.

5.5.3. *Back-off-based fairness*

High variance among the back-off timer values of different mobile nodes may contribute to unfairness. The channel contention becomes more fair when the mobiles have similar back-off timer values. This is the reason why in MACAW the mobile nodes broadcast their back-off values [9].

Derived from the 802.11 DCF, the distributed fair scheduling (DFS) mechanism proposed in Reference [22] can be used to allocate bandwidth to different flows in proportion to their 'weights'. Each frame is tagged with 'start' and 'finish' tags and at the current time the frame with the minimum finish tag is scheduled to be transmitted next. When a frame arrives, the start-tag (S) for the frame is set as the current time and the finish tag (F) is calculated as follows: $F = S + SF \cdot \frac{L}{W}$, where SF is a scaling factor, L is the frame length and W is the weight of the corresponding flow. Therefore, a bigger weight results in a smaller finish tag. The back-off interval for a frame is chosen to be proportional to the finish tag as follows:

$$B = \lfloor F - Current\_Time \rfloor = \lfloor SF \cdot \frac{L}{W} \rfloor \qquad (9)$$

Therefore, flows with larger weights will achieve higher throughputs. The protocol is described through pseudo-code in Appendix C.

In absence of collision, DFS achieves fairness while in the presence of collision there will be short-term unfairness. Again, assigning appropriate weights to different flows with bursty traffic arrival patterns may not be simple at all.

5.6. Energy-Efficiency-Based Approaches

Efficient use of battery power at each mobile node strongly impacts the overall performance of a mobile ad hoc network. Inefficiency in energy usage not only shortens the useful period of a mobile node, but also leads to failure in the routing process. To retain network connectivity, a routing mechanism is generally designed in such a way that all the mobile nodes have quite the similar power levels. The design objective would be to maximize the network performance, while minimizing the energy consumption. In other words, the objective is to maximize energy efficiency (η_e), where

$$\eta_e = \frac{\text{Essential energy dissipation}}{\text{Total energy dissipation}} \qquad (10)$$

Since retransmission of frames leads to unnecessary energy consumption, techniques such as frame transmissions based on channel state sensing and reducing the number of collisions can lead to efficient battery-power usage. Turning off the transceiver during idle period and during period when the transmission is forbidden or not likely to be successful (e.g. the NAV period in IEEE 802.11) may lead to better energy efficiency. Several approaches for energy-efficient MAC design are described below.

5.6.1. *Power level adjustment*

Even though developed mainly to increase the system throughput, the MAC mechanisms based on power level adjustment ([14–16]) also lead to efficient battery-power usage. Note that, although by using directional antennas throughput can be improved significantly, it requires more power than using an omnidirectional antenna.

5.6.2. *Periodic hibernation*

In this approach, the wireless transceiver is turned off during particular periods when the mobile node can neither transmit nor receive. For example, in IEEE 802.11, a mobile node hearing NAV may go to *sleep mode*, turning off the transceiver during this period. Note that, the NAV timer always decreases regardless of the channel status as opposed to the back-off timer which decreases only when medium is idle. In IEEE

802.11, every mobile node must wake up during an announcement traffic indication message (ATIM) period during which transmitters inform their destination not to go to *power save mode*. If no notification is received, the mobile can go to power save mode and wake up in the next ATIM period [3].

Again, a transmitting mobile node can defer its transmission (or at least reduce the transmission rate) when channel quality is bad and may try to compensate the loss when the channel becomes better.

In power aware multi-access with signaling (PAMAS) protocol [23], the receiving mobile nodes transmit a busy tone (in a separate control channel) when they start receiving frames so that other mobile nodes know when to turn off. When a mobile node does not have data to transmit, it should turn itself off if a neighbor begins transmitting to some other node. A node should power off even if it has data to transmit if at least one of it's neighbor-pairs is communicating. A mobile node, which has been powered off when one or more of it's neighbor-pairs started communicating, can determine the length of time that it should be powered off by using a *probe* protocol. In this protocol, the node performs a binary search to determine the time when the current transmission will end. However, the loss of probe frames may cause significant power wastage. Simulation results showed that power saving in the range from 10% (for sparsely connected networks) to almost 70% (for fully connected networks) could be achieved without affecting the delay-throughput behavior.

The periodic hibernation of a mobile node can be *actively* supported by another mobile node. The *High Performance LAN* (HIPERLAN) distributed MAC protocol uses the *power-saver* and *power-supporter*-based approach [24]. A mobile which wants to save power (power-saver node) must locate a power-supporter node before shutting down. During the power saving period, all data frames destined to the power-saver are buffered in the power-supporter. Only when the power-saver reactivates itself, the stored frames are forwarded from the power-supporter.

Periodic hibernation can be assisted through pseudo-centralized control, where the mobile nodes may periodically select a PBS as the coordinator for channel access. The PBS allocates the uplink and the downlink time slots to each mobile node resulting in much lesser number of collisions and hence reduced energy loss. Mobile nodes which do not receive the allocation can turn off their transceiver and wake up to listen in the next frame. Figure 16 shows the general transmission frame of a pseudo-centralized protocol.

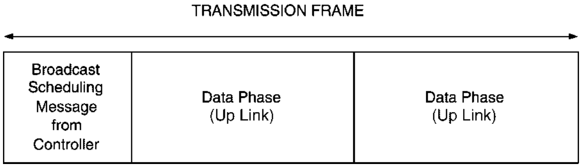


Fig. 16. General frame format for pseudo-centralized MAC protocols.

While evaluating the performance of a PBS-based approach, the loss incurred in battery-power and channel bandwidth for the discovery of a new PBS must be taken into account. The energy of the PBS is more likely to drop faster than that of other mobiles. Therefore, a PBS must be re-elected when its power drops below certain threshold [25].

5.6.3. Batch transmission

For transmission of a certain number of MAC frames, better energy efficiency can be achieved by transmitting them in batches rather than independently, because batch-transmission reduces the number of transceiver turnarounds. Therefore, a mobile node may start contending for the channel when it has enough data to transmit. However, this approach cannot be adopted in the case of urgent and real-time data. Data compression is an alternative to increase the actual amount of information transmitted over the wireless interface.

5.7. Approaches Based On Multimedia/QoS Support

The basic principle of multimedia MAC protocols is to support both non-real-time and real-time traffic flows so that the QoS requirements (e.g. throughput, delay, loss, jitter) for all the admitted flows are satisfied. In a mobile ad hoc network it is difficult to achieve due to the lack of any centralized control.

Since for retransmissions may often be unnecessary for real-time traffic, the MAC level ACKs could be suppressed [5]. Priority buffer management and scheduling mechanisms can be used to provide prioritized channel access to the real-time traffic. The Random Early Drop with IN/OUT (RIO) buffer management mechanism marks every buffered data frame with an IN or OUT tag. During congestion, OUT frames are dropped with higher probability than IN frames [26]. Priority scheduling considers the aspect of arranging the frames in queues. Service differentiation can be achieved by assigning separate queues for each class of traffic. Frames with higher priority could be scheduled and transmitted earlier. The scheduler might transmit low priority frames only in the absence of high priority frames. An alternative approach is to transmit high priority traffic with higher probability and to transmit low priority traffic with smaller probability.

Since providing a guaranteed QoS would be prohibitively hard, provisioning of differentiated QoS might be a better choice for a mobile ad hoc network. In the differentiated QoS approach, *soft QoS*[‖] is provided to each flow and all the flows in the same class (where the classes are predefined) will receive similar level of QoS.

In an ad hoc network, the differentiated QoS approach in the MAC protocol can be realized by adjusting the back-off and the persistence parameters. For example, in the 802.11 MAC protocol, the values of DIFS, $CW_{\min}$ and $CW_{\max}$ can be predefined, and priority-dependent random distributions can be used to generate the back-off times [27,28] for the different flows. The assignment of different IFS time and back-off time distribution corresponding to two different traffic classes can result in delay differentiation of the order of hundreds of milliseconds [27]. When the values of $CW_{\min}$ for two flows differ by 80, a delay difference of around 60 msec can be achieved. Therefore, effective service differentiation may be achieved by appropriately adjusting the contention window limits.

On top of the differentiated services MAC mechanism, a distributed admission control (or traffic control) mechanism would be required so that the relative assurances offered by the MAC mechanism can also meet service requirements at the application level [28]. Such traffic control mechanism can keep the offered load below a particular threshold under which throughput is proportional to offered load [29]. After this threshold, the throughput does not increase as the offered load increases. The admission control threshold can be designed as a function of delay or utilization. In any case, it should take channel condition into consideration and an adaptive admission control mechanism would be more efficient.

Implementation of a distributed admission control mechanism may involve significant computational power at the mobile nodes. In Reference [28], implementation of the distributed admission control algorithm in the mobile nodes was facilitated by using two passive radio channel monitoring algorithms, namely, the virtual MAC (VMAC) and the virtual source (VS) algorithms, to estimate the achievable level of service (at the MAC and at the application level) without actually loading the channel. Then the admission decision was made based on the application level QoS estimates (e.g. average packet delay).

In dynamic bandwidth allocation/sharing/-extension (DBASE) approach [30], service differentiation was achieved by using Priority IFS (PIFS), back-off timer and some reservation mechanism. When a mobile wants to send a non-real-time frame, it senses idle channel for DIFS and generates data back-off time (DBT) as in IEEE 802.11 DCF. If a mobile wants to transmit a real-time frame, it waits for channel to be idle for more than PIFS and generates *real-time back-off time* (RBT). PIFS is necessarily less than DIFS so that a real-time flow can initiate a back-off process before a non-real-time flow. By letting $(DIFS + DBT_{\min}) - (PIFS + RBT_{\max}) \geq Slot\_Time$, real-time flows are always allowed to send frames before non-real-time flows (Figure 17).

In DBASE, each mobile maintains the following data structures:

- Reservation table (RSVT): RSVT contains access sequence, MAC address, service type and required bandwidth corresponding to mobiles which have finished reservation successfully.
- *Sequence ID* (SID): SID identifies access order.
- *Active counter* (AC): AC is the number of the active real-time flows.

Let $D_{\max}$ be the smallest maximum acceptable delay among all real-time flows. Each mobile requiring to transmit a real-time flow must scan medium for $D_{\max}$. As can be seen in Figure 17, if there is at least one real-time flow in the network, reservation frame (RF) must be detected within $D_{\max}$. Therefore, if the mobile does not detect RF, it will become a scheduler responsible for generating RF.

In Figure 17, node A and node C do not hear RF within time $D_{\max}$. These nodes wait for PIFS and generate RBT. Assuming that RBT of node A is less than that of node C, node A initiates transmission earlier and becomes the scheduler. Node A transmits RF as well as its data. After node A finishes transmission, node C will transmit data without RF. After the first round, node A and node C will be in RSVT. Therefore, they do not have to send any handshake messages such as RTS or CTS before transmitting

---

[‖] This means that the quality measures can only be probabilistically guaranteed.

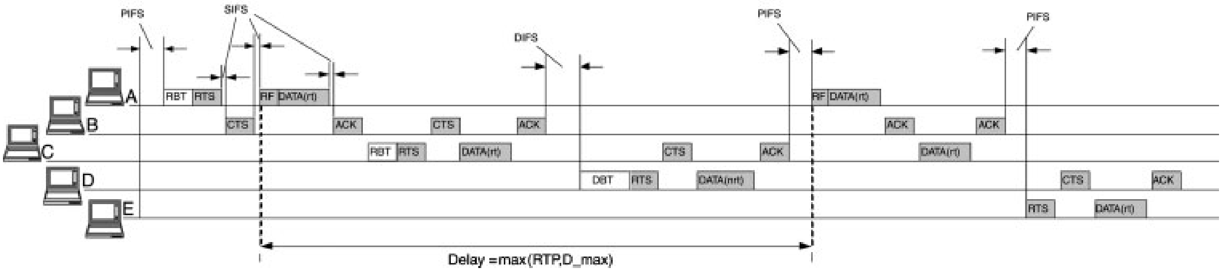


Fig. 17. Operation of DBASE protocol.

subsequent data frames. The period during which the RTS-CTS signaling is not involved for data transmission is called the contention free period (CFP). During this period, if the medium is idle for one time slot, the current session will be considered disconnected, and the other mobiles are shifted forward and access the medium according to the *SID*.

A new mobile (node E) detecting RF waits until the end of the current CFP and starts its transaction in the period called real-time contention period. In this period, when RTS is successfully sent, the mobile increases the *AC* by one and sets the *SID* to *AC*. Mobiles hearing CTS realize that a new mobile has successfully joined RSVT, and thus increase their *AC* by one. After mobiles in real-time mode finish transmission, the channel will be idle for more than DIFS and non-real-time data (say, in node D) can be transmitted.

A combination of CFP and real-time contention period is called real-time transmission period (RTP). RTP should be limited such that the repetition cycle is not larger than $D_{\max}$. A mobile starts monitoring the medium again after time RTP since it last accessed the medium. Therefore, frame delay is bounded by $(RTP, D_{\max})$. As a scheduler, a mobile emits RF after the medium is sensed idle for PIFS. Others will wait for RF before entering into the CFP. Note that since the scheduler does not generate RBT, it has more priority than any new mobile initiating a real-time flow. The protocol is described through pseudo-code in Appendix D.

6. Comparative Performance

The performances of the MAC protocols based on the different approaches described above differ in the different performance metrics. Table I shows a qualitative performance comparison among the representative MAC protocols based on the different approaches described before.

Techniques such as dedicating a portion of the channel bandwidth for control traffic, reducing the impact of hidden nodes (say, by using collision avoidance signaling, adaptive transmission power control, directional antenna, or enhancing the handshake signaling messages) improve the throughput and the delay performances. MAC level fairness and

Table I. Performance comparison among different protocols ('+'/'−' denotes improvement/degradation in the corresponding performance measure, '0' denotes no impact).

Approaches	Throughput/ delay	Fairness	Multimedia/ QoS support	Energy eff.
Handshake signaling (e.g. MACAW)	+++	0	0	+
Multichannel (e.g. MACA-CT, RICH-DP)	++	0	0	+
CCCS (e.g. MACA-CT, RICH-DP)	+++	0	0	+
Power control (e.g. PCMA)	+++	−	0	+++
Directional antenna	+++	+	0	−
Persistence scheme (e.g. FAMA-LCS)	+ (loss in stability)	+	+	0
Utility-based fairness	−	++	+	0
SBFA/LTFS	++	++	++	+, −
Hibernation (e.g. PAMAS)	−	0	0	++
Differentiated services	−	+	++	0
Reservation (e.g. DBASE)	−	+	+++	−

service differentiation can be achieved by controlling the transmission persistence and the back-off parameters and using reservation for the real-time flows. Turning off transceiver in the period of NAV can improve the energy-efficiency performance. A differentiated services MAC technique coupled with a distributed admission control mechanism can be used to support real-time multimedia traffic along with asynchronous data traffic in a wireless mobile ad hoc network.

In general, as MAC level throughput increases, the energy efficiency also increases due to decrease in the number of frame retransmissions. However, the approach using directional antennas results in more power consumption. Persistence in transmission increases the maximum throughput, but the throughput stability may decrease. QoS assurance to the real-time flows can be provided through increased persistence.

In order to achieve better fairness, an amount of throughput may need to be sacrificed, as can be seen in the utility, differentiated services and reservation-based approaches. Nevertheless, SBFA/LTFS enhances both fairness and throughput because it gives back allocation to mobiles deferring their transmission during bad channel condition. The implementation of SBFA requires a PBS and the battery-power of a PBS may become the bottleneck.

7. Research Issues

The following are some of the MAC design issues which are needed to be addressed to design high performance wireless ad hoc mobile systems for the fourth-generation mobile communications (4G mobile):

- *Application of advanced signal processing techniques*: Most of the MAC protocols do not attempt to separate the colliding frames. Application of signal processing techniques in multi-user detection (user separation) in CDMA systems has gained much research interest recently. These techniques require extra diversity in the form of spreading codes and generally involve substantial computational complexity and require expensive hardware. Techniques with low computational complexity are to be developed for MANETs.
- *Dynamic link adaptation*: To match the data transmission rate to the time-varying channel and interference conditions, the mobile nodes may adopt some adaptive coding and modulation scheme so that higher channel utilization is achieved. Information regarding link condition such as mean carrier-to-interference ratio (C/I) or signal-to-noise ratio (SNR) required for link adaptation can be communicated through the handshake signaling messages. Efficient and implementation-friendly dynamic radio link adaptation mechanisms need to be developed for wireless ad hoc networks.
- *Distinguish between congestion (collision) and channel error*: When the network is not congested, but transmission from a mobile node suffers from channel error, the node should retransmit as soon as the channel becomes better. In this case the node does not need to increase the back-off time aggressively. But when a mobile node experiences collision, it can increase the back-off time and wait for retransmission until the network becomes less congested. Therefore, if a mobile node is able to distinguish between these two cases, throughput and channel utilization at the MAC level could be increased.
- *Mobility and location-aware MAC design*: The mobility patterns of the mobile nodes and the resulting spatial distribution is one of the most consequential issues for channel access protocol design in ad hoc networks. As mobile speed increases, the channel errors (due to fading) generally become more uncorrelated. Therefore, even if a mobile experiences frame error in the current time slot, retransmission of the frame in the following time slot may be successful. This has implication on both the fairness and energy-efficiency performance of a MAC protocol. Again, a mobile node in a sparse neighborhood can choose to adjust the transmission power and the back-off parameters differently than a node with a relatively denser neighborhood. In fact, the mobility and location information along with the information on channel quality can be used for dynamic link adaptation.
- *Interaction among transport, routing and MAC layer protocols*: The performance of a transport layer protocol heavily depends on the MAC protocol performance. TCP performance in an IEEE 802.11 DFC suffers due to the exposed node problem. Therefore, a transport-aware MAC protocol in a mobile node may give higher priority to a flow closer to it's destination. Collision in the last hop is very undesirable in that all transmissions which might have already passed several hops have to be started over again. The MAC protocol can also use the end-to-end information provided by the transport layer protocol for back-off time adaptation (say, to deal with end-to-end fairness).

Some power saving mechanism at the MAC protocol may impact the routing performance in a mobile ad hoc network. Mobile nodes should not go to power save mode, if it results in the disruption in network connectivity. Again, some intermediate nodes with lower energy may be avoided during routing (even though they are in the least-cost path) in order to maintain similar amount of battery power at all mobile nodes. This will lead to longer network connectivity. In such a case, the operation of a transmission power adaptation-based MAC protocol would be dictated by the routing protocol. Again, frames destined for a low power mobile node should be sent with more priority so that they will arrive the mobile before the battery power is depleted.

The interactions among MAC, routing and transport layer protocols need to be thoroughly investigated and this would be required for transmission protocol stack performance optimization in a wireless mobile ad hoc network.

References

1. Chhaya HS, Gupta S. Performance of asynchronous data transfer methods for IEEE 802.11 MAC protocol. *IEEE Personal Communications Magazine* 1996; **3**: 8–15.
2. Nelson R, Klienrock L. The spatial capacity of a slotted ALOHA multihop packet radio network with capture. *IEEE Transactions on Communications* 1984; **32**: 684–694.
3. IEEE Std. 802.11. Wireless LAN media access control (MAC) and physical layer (PHY) specifications. 1997.
4. Xu S, Saadawi T. Does the IEEE 802.11 MAC protocol work well in multihop wireless ad hoc networks? *IEEE Communications Magazine* 2001; **39**: 130–137.
5. Anastasi G, Lenzini L. QoS provided by the IEEE 802.11 wireless LAN to advanced data applications: a simulation analysis. *ACM Kluwer Journal of Wireless Networks* 2000; **6**(2): 99–108.
6. Karn P. MACA—a new channel access method for packet radio. *Proceedings of 9th Amateur Radio Computer Networking Conference*, 1990.
7. Fullmer CL, Garcia-Luna-Aceves JJ. Solutions to hidden terminal problems in wireless networks. *Proceedings of ACM SIGCOMM'97*, Cannes, France, 1997; 212–225.
8. Garcia-Luna-Aceves JJ, Tzamaloukas A. The effect of exerting adequate persistence in collision avoidance protocols. *Proceedings of IEEE MoMuC'99*, San Diego, CA, 1999; 328–337.
9. Bharghavan V, Demers A, Shenker S, Zhang L. MACAW—A media access protocol for wireless LANs. *Proceedings of ACM SIGCOMM'94*, 1994; 212–225.
10. Sousa S, Silvester JA. Sperading code protocols for distributed spread-spectrum packet radio networks. *IEEE Transactions on Communications* 1998; **36**(3): 272–281.
11. Joa-Ng M, I-Tai L. Spread sprctrum medium access protocol with collision avoidance in mobile ad-hoc wireless network. *Proceedings of IEEE INFOCOM'99*; 776–783.
12. Tzamaloukas A, Garcia-Luna-Aceves JJ. A receiver-initiated collision-avoidance protocol for multi-channel networks. *Proceedings of IEEE INFOCOM'01*; 189–198.
13. Tobagi FA, Klienrock L. Packet switching in radio channels: Part II—the hidden terminal problem in carrier sense multiple access and the busy tone solution. *IEEE Transactions on Communications* 1975; **23**: 1417–1433.
14. Takagi H, Kleinrock L. Optimal transmission ranges for randomly distributed packet radio terminals. *IEEE Transactions on Communications* 1984; **32**: 246–257.
15. Hou T, Li VOK. Transmission range control in multihop packet radio networks. *IEEE Transactions on Communications* 1986; **34**: 38–44.
16. Monks JP, Bharghavan V, Hwu WMW. A power controlled multiple access protocol for wireless packet networks. *Proceedings of IEEE INFOCOM'01*; 219–228.
17. Ko YB, Shankarkumar V, Vaidya NH. Medium access control protocols using directional antennas in ad hoc networks. *Proceedings of IEEE INFOCOM'00*; 13–21.
18. Kar K, Sarkar S, Tassiulas L. A simple rate control algorithm for maximizing total user utility. *Proceedings of IEEE INFOCOM'01*; 133–141.
19. Nandagopal T, Kim T, Gao X, Bharghavan V. Achieving MAC layer fairness in wireless packet networks. *Proceedings of ACM MOBICOM'00*, Boston, MA; 87–98.
20. Kelly FP, Maulloo AK, Tan DKH. Rate control in communication networks: Shadow prices, proportional fairness and stability. *Journal of the Operational Research Society* 1998; **49**(3): 237–252.
21. Ramanathan P, Agrawal P. Adapting packet fair queuing algorithms to wireless networks. *Proceedings of ACM MOBICOM'98*, Dallas, TX; 1–9.
22. Vaidya NH, Bahl P, Gupta S. Distributed Fair Scheduling in a Wireless LAN. *Proceedings of ACM MOBICOM'00*, Boston; 167–178.
23. Singh S, Raghavendra CS. PAMAS: power aware multi-access protocol with signaling for ad hoc networks. *Computer Communication Review* 1998; **28**(3): 5–26.
24. ETSI. HIgh PErformance Radio Local Area Network (HIPERLAN) Type 1 Functional Specification. *ETSI Standard*, Document number EN 300 652, 1998.
25. Jin KT, Cho DH. Optimal threshold energy level of energy efficient MAC for energy-limited ad hoc networks. *Proceeding of IEEE GLOBECOM'01*, San Antonio, TX.
26. Xiao H, Chua KC, Seah W, Lo A. On service prioritization in mobile ad hoc networks. *Proceedings of IEEE ICC'01*.
27. Pallot X, Miller LE. Implementating message priority policies over an 802.11 based mobile ad hoc network. *Proceedings of IEEE MILCOM'01*.
28. Veres A, Campbell AT, Barry M, Sun LH. Supporting service differentiation in wireless packet networks using distributed control. *IEEE Journal on Selected Areas in Communications* 2001; **19**(10): 2081–2093.
29. Wu HK, Chuang PH. Dynamic QoS allocation for multimedia ad hoc wireless networks. *ACM Baltzer Mobile Networks and Applications* 2001; **6**(4): 377–384.
30. Sheu ST, Sheu TF. A bandwidth allocation/sharing/extension protocol for multimedia over IEEE 802.11 ad hoc wireless LANs. *IEEE Journal on Selected Areas in Communications* 2001; **19**(10): 2065–2080.

Appendix A: (Pseudo-code for 802.11 MAC)

```
Procedure Passive()
begin
  while (!CD && !Local_Data) wait;
  if CD
    if (dest_id != mobile_id)
      if (frame_type == RTS || CTS || DATA)
        wait_type = NAV;
        timer = readtime(frame);
      else
        wait_type = none;
        timer = 0;
      WAITandSEND(wait_type, timer, blank_frame);
    else
      if (frame_type == RTS)
        WAITandSEND(IFS, SIFS, CTS_frame);
    else
      WAITandSEND(IFS, DIFS, RTS_frame);
end

Procedure WAITandSEND(wait_type, time, next_frame)
begin
  Do CASEs (wait_type}
    NAV: while (time != 0) reduce(time);
          Passive();
    IFS: temp = time;
        timeout = ReadTimeout(next_frame)
        while (!CD && time != 0) reduce(time);
        if (CD && timeout != 0)
          WAITandSEND(IFS, temp, next_frame);
        else
          SENDandWAIT(next_frame);
    BO:   while (time != 0)
            while (CD)
              if ((dest_id != mobile_id) && (frame_type == RTS || CTS || DATA))
                while (NAV) reduce(NAV);
              if ((dest_id == mobile_id) && (frame_type == RTS))
                WAITandSEND(IFS, SIFS, CTS_frame);
          residue_time = DIFS;
          while (!CD && residue_time != 0) reduce(residue_time);
          if (residue_time == 0)
            while (!CD && time != 0) reduce(time);
        Passive();
    default:
end

Procedure SENDandWAIT(frame)
begin
  Transmit(frame);
  Do CASEs (frame_type)
    RTS: timer = CTS_timeout;
        while (!receive_CTS && timer != 0) reduce(timer);
          if (timer == 0)
            BackOff();
        else
            Receive(in_frame);
            WAITandSEND(IFS, SIFS, DATA_frame);
```

```
CTS:   timer = DATA_timeout
       while (!receive_DATA && timer != 0) reduce(timer);
         if (timer == 0)
             Passive();
         else
             Receive(in_frame);
             WAITandSEND(IFS, SIFS, ACK_frame);
DATA:   timer = ACK_timeout
       while (!receive_ACK && timer != 0) reduce(timer);
       if (timer == 0)
         BackOff();
       else
           Receive(in_frame);
           CW = CW_min;
           Passive();
   ACK: Passive();
end

Procedure BackOff()
begin
  BO_time = rand(0, CW);
  CW = CW * 2;
  if (CW > CW_max)
    CW = CW_max;
  WAITandSEND(BO, BO_time, blank_frame);
end;
```

Appendix B: (Pseudo-code for RICH-DP protocol)

```
Procedure Passive() /* After turned on, a mobile starts this procedure */
begin
  continue hopping;
  while (!CD && !Local_Data) wait;
  stop hopping;
  if Local_Data
    if (!CD in current time slot)
      SENDandWAIT(RTR);
    else
      BackOff();
    else
      if ((frame_type == RTR) && (dest_id == mobile_id))
        if Local_Data
          SENDandWAIT(DATA);
        else
          SENDandWAIT(CTS);
end

Procedure SENDandWAIT(frame)
begin
  Transmit(frame);
  Do CASEs (frame_type)
    RTR:  if (DATA or CTS in next time slot)
             Receive(in_frame);
             if Local_Data
               Transmit(ACK);
               SENDandWAIT(DATA);
             else
               SENDandWAIT(ACK);
             else
               BackOff();
```

```
CTS:  if (DATA in next time slot)
        Receive(in_frame);
        SENDandWAIT(ACK);
       else
        Passive();
DATA:  if (ACK in next time slot)
        receive frame;
        if (more data is coming)
          Receive(in_frame);
          SENDandWAIT(ACK);
        else
          Passive();
        else
          BackOff();
   ACK: Passive();
end

Procedure BackOff()
begin
  continue hopping
  BO_time = rand(1, CW);
  while (no RTR coming for this mobile && BO_time != 0) reduce(BO_time);
  if (BO_time == 0)
    Passive();
  else
    if Local_Data
      SENDandWAIT(DATA);
    else
       SENDandWAIT(CTS);
end;
```

Appendix C: (Pseudo-code for DFS protocol)

```
Procedure Passive()
begin
  while (!CD && !Local_Data) wait;
  if CD
    if (dest_id != mobile_id)
      if (frame_type == RTS || CTS || DATA)
        wait_type = NAV;
        timer = readtime(frame);
      else
        wait_type = none;
        timer = 0;
      WAITandSEND(wait_type, timer, blank_frame);
    else
      if(frame_type == RTS)
        WAITandSEND(IFS, SIFS, CTS_frame);
  else
      WAITandSEND(IFS, DIFS, RTS_frame);
end

Procedure WAITandSEND(wait_type, time, next_frame)
begin
  Do CASEs (wait_type)
    NAV:  while (time != 0) reduce(time);
          Passive();
    IFS:  temp = time;
        timeout = ReadTimeout(next_frame)
        while(!CD && time != 0) reduce(time);
```

```
        if (CD && timeout != 0)
          WAITandSEND(IFS, temp, next_frame);
      else
        if (frame_type == RTS)
            wait_time = floor(scaling_factor * frame_length/flow_weight);
            WAITandSEND(BO, wait_time, next_frame);
          SENDandWAIT(next_frame);
BO:   while (time != 0)
      while (CD)
        if ((dest_id != mobile_id) && (frame_type == RTS || CTS || DATA))
            while (NAV) reduce(NAV);
        if ((dest_id == mobile_id) && (frame_type = = RTS))
          WAITandSEND(IFS, SIFS, CTS_frame);
      residue_time = DIFS;
      while (!CD && residue_time != 0) reduce(residue_time);
      if (residue_time == 0)
          while (!CD && time != 0) reduce(time);
      Passive();
end

Procedure SENDandWAIT(frame)
begin
  Transmit(frame);
  Do CASEs (frame_type)
    RTS:   timer = CTS_timeout;
      while (!receive_CTS && timer != 0) reduce(timer);
        if (timer == 0)
          BackOff();
        else
          Receive(in_frame);
          WAITandSEND(IFS, SIFS, DATA_frame);
  CTS:   timer = DATA_timeout
         while (!receive_DATA && timer != 0) reduce(timer);
          if (timer == 0)
            Passive();
          else
            Receive(in_frame);
            WAITandSEND(IFS, SIFS, ACK_frame);
  DATA:   timer = ACK_timeout
          while (!receive_ACK && timer != 0) reduce(timer);
            if (timer == 0)
              BackOff();
            else
              CW = CW_min;
              Passive();
      ACK:   Passive();
end

Procedure BackOff()
begin
  BO_time = rand(0, CW);
  CW = CW * 2;
  if (CW > CW_max)
    CW = CW_max;
  WAITandSEND(BO, BO_time, blank_frame);
end;
```

Appendix D: (Pseudo-code for DBASE protocol)

```
Procedure Passive()
begin
 RF_detected = false;
 while(!CD && !Local_Data) wait;
  if CD
   if (dest_id != mobile_id)
    if (frame_type == RTS || CTS || DATA)
     while (NAV != 0) reduce(NAV);
    else
     if (frame_type == RTS)
      if (idle_time < SIFS)
       Passive();
      else
       WAITandSEND(CTS_frame);
       Passive();
 else
    WAITandSEND(RTS_frame);
    Passive();
end

Procedure WAITandSEND(time, next_frame)
begin
 DO CASEs (frame_type)
  RTS:    timer = D_max;
       while (!RF_detected && !RF && timer != 0) reduce(timer);
       if RF
         update RSVT;
         RF_detected = true;
         wait(CFP);
       while (idle_time < PIFS) wait;
       BackOff();
       SENDandWAIT(next_frame);
  default:   while (!CD && (idle_time < SIFS))
       if CD
        Passive();
       else
       SENDandWAIT(next_frame);
end

Procedure SENDandWAIT(Frame)
begin
 Transmit(frame);
 Do CASEs (frame_type)
  RTS:  timer = CTS_timeout;
     while (!receive_CTS && timer != 0) reduce(timer);
       if (timer == 0)
        p = rand(0,1);
        if (p < persist_threshold)
          BackOff();
        else
         SENDandWAIT(RTS);
        else
         receive(frame)
         if (!RF_detected)
          Transmit(RF);
         WAITandSEND(DATA_frame);
 CTS:  timer = DATA_timeout
     while (!receive_DATA_or_RF && timer != 0) reduce(timer);
       if (timer == 0)
        Passive();
```

```
            else
              recieve(frame);
              if (type(frame) == RF)
                receive(frame);
              if (type(frame) == DATA)
                WAITandSEND(ACK_frame);
              else
                Passive();
DATA:   timer = ACK_timeout
        while (!receive_ACK && timer != 0) reduce(timer);
          if (timer == 0)
            BackOff();
          else
            min = CW_min;
            CFP();
            RF_detected = false;
    ACK:
end

Procedure BackOff(frame)
begin
  received_RTS = false;
  BO_time = rand(min, max);
  min = min + 1;
  if (min > max) min = max;
  while (BO_time != 0)
    while (!CD && (BO_time != 0)) reduce(BO_time);
  if CD
    DO CASEs (frame_type)
      RTS:   if (dest_id == mobile_id)
              BO_time = 0;
              received_RTS = true;
            else
              while (NAV != 0) reduce(NAV);
      CTS:     record dest_id of CTS to RSVT;
              while (NAV != 0) reduce(NAV);
      DATA:    while (NAV != 0) reduce(NAV);
      RF:       rf_detect = true;
               update RSVT;
if received RTS
  SENDandWAIT(CTS_frame);
end

Procedure CFP()
begin
  if (!Local_Data) Passive();
  else
    while ((RTP > 0 since sent last RF) && !RF)
      reduce(RTP);
      if (CTS_detected) update RSVT;
      if (RTS_detected && (dest_id == mobile_id))
        WAITandSEND(CTS_frame);
    RF_detected = true;
    while (!RF && (idle_time < PIFS)) wait;
    if (!RF) Transmit(RF);
    else
      while (CD && !time_to_transmit)
        if (idle_time > time_slot)
          shift access order forward;
          update RSVT;
    SENDandWAIT(DATA_frame);
end
```

Authors' Biographies

Teerawat Issariyakul is currently working towards his Ph.D. at the Department of Electrical and Computer Engineering at University of Manitoba, Winnipeg, Canada. He received his M.Eng. degree in Telecommunications Engineering from Asian Institute of Technology (AIT) in 2000. His research interests include radio resource management techniques and transport layer protocol design for mobile networks.

Ekram Hossain is an assistant professor at the Department of Electrical and Computer Engineering at University of Manitoba, Winnipeg, Canada. He received his Ph.D. in Electrical Engineering from University of Victoria, Canada, in 2000. He was a University of Victoria Fellow. He received his B.Sc. and M.Sc. both in Computer Science and Engineering from Bangladesh University of Engineering and Technology (BUET), Dhaka, Bangladesh, in 1995 and 1997, respectively. His research interests include radio link control and transport layer protocol design issues for the next-generation wireless data networks. Currently he is serving as an editor for the *IEEE Transactions on Wireless Communications*.

Dong In Kim received the B.S. and M.S. degrees in Electronics Engineering from Seoul National University, Seoul, Korea, in 1980 and 1984 respectively and the M.S. and Ph.D. in Electrical Engineering from University of Southern California (USC), Los Angeles, in 1987 and 1990, respectively. From 1984 to 1985, he was with the Korea Telecommunication Research Center as a researcher. During 1986–1988, he was a Korean Government Graduate Fellow in the Department of Electrical Engineering, USC. From 1988 to 1990, he was a research assistant at the USC Communication Sciences Institute. Since 1991 he has been with the University of Seoul, Seoul, Korea, leading the Wireless Communications Research Group. He is now an associate professor of the School of Engineering Science at Simon Fraser University, Burnaby, B.C., Canada. He was a visiting professor at University of Victoria, Victoria, B.C., Canada, during 1999–2000. He has performed research in the areas of packet radio networks and spread-spectrum systems since 1988. His current research interests include spread-spectrum systems, cellular mobile communications, indoor wireless communications and wireless multimedia networks.

Dr. Kim has served as an editor for the *IEEE Journal on Selected Areas in Communications: Wireless Communications Series* and also as a division editor for the *Journal of Communications and Networks*. Currently he is serving as an editor for the *IEEE Transactions on Communications* and the *IEEE Transactions on Wireless Communications*.