Investigation of Media Access Control Protocols for Mobile Ad-hoc Networks

Mort Naraghi-Pour, Ph.D.

Final Report Submitted to:

National Institute of Standards and Technology

Gaithersburg, MD 20899-0001

Request no. Q-912364 Purchase request no. 892-9326

Contents

Abstract

This report includes the results of our investigation of media access control (MAC) protocols suitable for integration of emergency, real-time and non-real-time traffic in a mobile ad-hoc network. We first evaluate several existing protocols in their ability to provide access priority and differentiated quality of service to such integrated services. We then propose a new dynamic, distributed time division multiple access protocol which provides prioritized access to emergency, real-time and non-real-time traffic, in that order. In this protocol data packets do not suffer from collisions with any other transmission. Furthermore, when operated in a fully connected ad-hoc wireless LAN (with no hidden terminals), the protocol guarantees bounded delay for emergency traffic.

1 EXECUTIVE SUMMARY

This report contains the results of our investigation of the capabilities of media access control (MAC) protocols in supporting prioritized access and/or differentiated quality of service (QoS) for mobile ad-hoc networks.

In mobile ad-hoc networks, in addition to non-real-time data traffic, the human-to-human communication plays an important role. This communication generally produces real-time traffic such as voice and video. In addition, many applications of ad-hoc networks such as military tactical communication, law enforcement and emergency response efforts require transmission of emergency messages. Such messages require assured access with the least possible delay. The integration of emergency, real-time and non-real-time traffic creates different priority classes each with their own QoS requirements.

Provisioning of QoS and support of prioritized access in a mobile ad-hoc network is a challenging task due to the so called hidden-terminal problem. In this report we describe several MAC protocols (e.g., carrier sense multiple access, multiple access collision avoidance (MACA), MACAW, floor acquisition multiple access, and black burst contention) that have been recently presented in the literature. We highlight the advantages and disadvantages of these protocols and evaluate their capabilities in supporting different priority classes and differentiated quality of service. Among these protocols, only the black burst contention protocol can provide prioritized access and bounded delay for real-time traffic in a fully connected network with no hidden terminals.

Next, we present a new MAC protocol based on a dynamic slot assignment, distributed time division multiple access system which combines the concepts of interframe spacing of IEEE

802.11 and black burst contention with the request-to-send (RTS) and clear-to-send (CTS) exchange of MACA. In this protocol data packets are not subject to collisions and priority access is provided to emergency, real-time and non-real-time traffic, in that order. When operated in a fully connected ad-hoc wireless LAN (with no hidden terminals), the protocol guarantees bounded delay for emergency traffic. Unfortunately, no such guarantees can be given in a mobile ad-hoc network with hidden terminals. However, the average delay of emergency traffic is smallest, followed by that of real-time and non-real-time traffic.

2 Introduction

Ad-hoc networks are self-organizing wireless networks composed of geographically dispersed mobile radio units that communicate with each other without a fixed infrastructure. Because they do not require an existing infrastructure, these networks can be rapidly deployed to provide robust communication in a variety of applications from civilian (e.g., collaborative or distributed computing, electronic classrooms, convention centers, construction cites) to disaster recovery (e.g., fire, flood, earthquake), law enforcement (e.g., crowd control, search and rescue) and military (tactical communication in the battle field).

Radio units in an ad-hoc network generally have a small range of coverage due to their limited transmit power. If the network is limited to a small geographic area, it may be possible for every station to receive transmission from every other station. In this report we refer to such a fully connected network as an ad-hoc wireless local area network or ad-hoc WLAN. These are also referred to as single-hop networks. A network in which transmissions from some stations can not be heard by every other station is referred to as a Mobile Ad-hoc NETwork (MANET). These are also referred to as multi-hop packet radio networks in the literature.

The self-organizing and decentralized nature of ad-hoc networks coupled with the large number of users, their mobility and the need to support multimedia applications create many technical challenges for network designers. The requirement to support multimedia communication stems from the fact that in mobile scenarios, in addition to non-real-time data traffic, the human to human communication plays an important role. This communication generally creates real-time traffic such as voice and video. Consequently, real-time traffic support with quality of service (QoS) guarantees is essential. For real-time traffic, QoS is expressed in terms of delay

and delay jitter bound.

In addition, many applications of ad-hoc networks such as military tactical communication, law enforcement and emergency response efforts require transmission of emergency messages. Such messages require assured access with the least possible delay. The integration of emergency, real-time and non-real-time traffic creates different priority classes each with their own QoS requirements. We should point out that provisioning of QoS and support of prioritized access, although related, are not the same. Clearly a MAC protocol capable of providing QoS guarantees can support prioritized access. However, the converse is not true. A MAC protocol may be capable of providing prioritized access without QoS guarantees.

One of the principal challenges of ad-hoc networks is the design of an efficient media access control (MAC) protocol in order to coordinate the transmission of mobile nodes over the radio channel and support the QoS requirements of different priority classes. Of necessity, such a protocol has to be distributed. In addition, due to the scarcity of bandwidth in the wireless environment, the MAC protocol must ensure that a high level of statistical multiplexing gain and efficient channel utilization is achieved.

Designing a MAC protocol capable of providing QoS guarantees to different priority classes in an ad-hoc network is a challenging task. To provide performance guarantees, some networks have adopted MAC protocols based on fixed channel assignment schemes such as frequency-division multiple access (FDMA) or time-division multiple access (TDMA). Fixed channel assignment techniques however are not very efficient in their use of channel bandwidth for bursty traffic sources as stations will not use their allocated bandwidth efficiently.

Some new protocols have been recently proposed which can provide delay guarantees to real-time traffic in a fully connected ad-hoc WLAN [16][13]. However, except for simple FDMA

and TDMA protocols we are aware of no other such protocols for mobile ad-hoc networks which may not be fully connected.

The remainder of this report is organized as follows. In section 2 we review several protocols that have recently appeared in the literature, compare their features and evaluate their capabilities in providing QoS guarantees to multimedia applications and supporting different priority classes. In Section 3 we present a new protocol which allows prioritized access in any ah-hoc network and can provide QoS guarantees in a fully connected ah-hoc WLAN. Finally conclusions are drawn in Section 4.

3 Review of Existing Protocols and Evaluation of Their Capabilities

3.1 Carrier Sense Multiple Access

Carrier sense multiple access (CSMA) protocols have been suggested for ad-hoc wireless networks following their great success as a multiple access scheme in wired LAN's [8]. CSMA is a simple distributed protocol whereby nodes regulate their packet transmission attempts based on their perception of the state, busy or idle, of the common radio channel. A station transmits if it finds the channel to be idle (no carrier) and defers transmission if it finds it to be busy (carrier detected).

Packet collisions are intrinsic to CSMA due to the fact that each node only has a delayed perception of the other nodes' activity. Packet collisions result in wasted channel bandwidth and lowered network throughput. In wired networks nodes can listen to the channel while

transmitting. Thus they can detect a collision, if it occurs, and abort their transmission and thereby avoid wasting channel bandwidth. They will then schedule the retransmission of their packets to a random time in the future, in order to avoid another immediate collision. This scheme is referred to as CSMA with collision detection (CSMA/CD) and has been a popular MAC protocol for wired LAN's.

The radio units operating on a single carrier frequency can not transmit and listen to the channel simultaneously. Even if they could, detection of collisions will not be feasible as collisions occur at the receiver, whereas at the transmitter where collision is to be detected, the interferers signal is often significantly weaker than the transmitter's own signal. These obstacles preclude the use of collision detection protocols in the wireless environment. Consequently, as discussed in Section 3.2, many of the recently proposed MAC protocols rely on a collision avoidance procedure.

In a mobile ad-hoc network the performance of CSMA is further limited by the so called hidden and exposed terminals. CSMA attempts to avoid collisions by detecting carrier in the vicinity of the transmitter. Collisions however occur at the receivers, not the transmitter, where two or more transmitted signals interfere with each other. Consequently, CSMA does not provide an appropriate mechanism for collision avoidance. Consider the network of Figure 1 consisting of four nodes. In this and other figures in this report, links are assumed to be symmetric, i.e., if node A can hear node B's transmission, then node B can hear node A's transmission as well. Furthermore, nodes that can hear each other are connected by a solid line. Nodes not connected by a solid line can not hear each other.

Consider the case where node A is transmitting to its neighboring node B (Figure 1). A node C that is a neighbor of B but can not hear A's transmission finds the channel to be idle

Figure 1: The hidden and exposed terminal problem in CSMA.

and may start transmission simultaneously with A. This causes a collision at node B. This is the so-called hidden terminal problem. The exposed terminal occurs when B is transmitting to A. Now if C intends to transmit to some other node D, it will detect carrier and therefor defer transmission. However, there is no reason for C to defer as its transmission will not cause a collision at A. Due to these problems the performance of CSMA degrades substantially to a point close to that of Aloha [1]. Clearly the hidden terminal problem does not exist in an ad-hoc WLAN where all the nodes can hear each other. Moreover, the exposed terminal problem is not relevant in an ad-hoc WLAN as a node detecting carrier should in fact defer transmission.

Conclusion

Carrier sense multiple access, due to the problems described above, does not provide satisfactory performance in the wireless environment. Furthermore, CSMA can not provide QoS guarantees to multimedia applications. All traffic sources are treated equally and performance guarantees can not be provided. In addition, CSMA can not support different priority classes in ad- hoc WLAN or in ad-hoc mobile networks.

3.2 Multiple Access Collision Avoidance

Split-channel reservation multiple access (SRMA) is one of the first protocols introduced for wireless networks that avoids collisions of data packets by introducing a control-signal handshake between the transmitter and the receiver [14]. When node A wishes to send a packet to node B, using Aloha or CSMA, it sends a Request-to-Send (RTS) packet to B. Upon receiving the

RTS packet, node B, if it is not currently deferring, transmits a Clear-to-Send (CTS) packet to A. Upon receiving the CTS packet, node A commences transmission of its data packet. SRMA used separate control channels for RTS and CTS packets.

Since SRMA was introduced, several other MAC protocols have been developed which utilize the RTS-CTS dialogue concept [2] [3] [9] [10] [11]. More recently Karn introduced the multiple access collision avoidance (MACA) protocol [6] which also relies on the RTS-CTS exchange. MACA is a single-channel protocol that uses the Aloha protocol for the transmission of RTS and CTS packets. When node A wishes to send a packet to B, it includes the length of the data packet in the RTS packet. Any station hearing the RTS packet will defer long enough for the associated CTS packet to be received by node A. Also, node B includes the length of the data packet in the CTS packet. Any node hearing the CTS packet will defer for the duration of the corresponding data packet.

After transmitting the RTS packet, if node A does not receive the CTS packet within a designated period, it will eventually time out, assume a collision occurred and schedule the packet for retransmission in the future. This is the backoff procedure which appears in many single-channel MAC protocol.

The RTS-CTS exchange relieves, but does not completely eliminate, the hidden and exposed terminal problem of CSMA (see Section 3.3). Nodes that hear the RTS but not the corresponding CTS (exposed terminals) can proceed with their transmissions since they are not in the range of the receiver and thus their packet transmissions will not collide at the receiver. Similarly, any node that hears the CTS (a potential hidden terminal) will defer for the duration of the data packet.

It should be pointed out that the RTS and CTS packets are themselves subject to collisions.

The RTS packet may collide with transmissions by the neighbors of the receiver. Similarly, the CTS packet may collide with the transmissions from neighbors of the transmitter. As these packets are generally much shorter than the data packet, the effect of their collision on network throughput is not significant compared with the effect of collision of data packets.

The MACA protocol as described above and in [6] is very brief and leaves many details out. In [2], the authors investigated in detail, various design and performance issues of this protocol and introduced several modifications to the original protocol. The authors referred to this new protocol as MACAW.

The first modification to MACA presented in [2] is in the backoff algorithm. MACA uses a binary exponential backoff algorithm which doubles the backoff after every collision and reduces it to its minimum valve after each successful RTS-CTS exchange. Noting that this results in large oscillations in the backoff counter causing some nodes to be completely backed off while others transmit at maximum rate, the authors propose a backoff algorithm in which the backoff counter is increased by a multiplicative factor when collisions occur and decreased linearly after each successful RTS-CTS dialogue. They demonstrate that, with this new algorithm, a more fair allocation of throughput can be achieved.

The next modification is the addition of acknowledgement packets (ACK's) to improve link reliability and enhance system throughput in the presence of channel errors.

In the original discussion of exposed terminal in Figure 1 we explained that the exposed terminal C is free to transmit since, even though it is in the range of the transmitting node B, it is out of the range of the receiving node A. While this is true, C's transmission of the RTS packet will be futile as it can not hear the returned CTS packet for as long as B is transmitting. Consequently, sending RTS packets causes C's back off counter to increase rapidly. In [2] the

authors propose an additional data sending (DS) packet to be sent by a node before transmission of data packets. The DS packet informs the neighbors of the sender of the impending packet transmission and the packet length. The neighbors of the sender must then defer for the duration of the packet and its corresponding ACK.

Another improvement to the MACA protocol proposed in [2] is the addition of the Request-for-Request-to-Send (RRTS) packet. A node B receiving an RTS packet will not be able to reply with its CTS packet if it is deferring at the time (e.g., due to transmissions from its neighbors). This causes the backoff counter of the sender to increase and may result in significant reduction in its throughput. To alleviate this problem if node B receives an RTS packet to which it can not respond immediately, due to deferral, node B transmits an RRTS packet to the sender at the first opportunity it gets, thereby soliciting from the sender an RTS packet. The sender will then immediately transmit its RTS packet.

Many simulations are presented in [2] comparing the throughput and fairness of MACAW under different scenarios. However, it should be clear that this protocol does not distinguish between different priority classes. All nodes have equal access to the channel and no distinction is made for the different types of traffic.

Conclusion

MACAW improves the performance of the original MACA protocol in terms of network throughput and access fairness for mobile nodes. However, MACAW does not provide performance guarantees to emergency or real-time applications in a fully connected ad-hoc WLAN or a mobile ad-hoc network (with hidden terminals). Furthermore, MACAW treats all nodes equally and does not support prioritized access.

3.3 Floor Acquisition Multiple Access

In MACAW, if the transmission time of the control packets, RTS and CTS, is small compared with the propagation delay in the channel, it is possible for two nearby nodes to finish their RTS-CTS exchange before the sender can hear the transmission that has already started from some distant node. The authors in [4] have proposed the floor acquisition multiple access (FAMA) protocol to solve this problem in the case of fully connected networks with no hidden terminals (ad-hoc WLAN in the terminology of this report). The main requirement in FAMA is that the RTS and CTS packets have a transmission time that is larger than the maximum one-way propagation time between any two nodes in the network. The authors prove the correctness of FAMA showing that the data packets will not collide with any other transmission.

Although MACA and MACAW were originally intended to solve the hidden terminal problem of CSMA, it is easy to show that they can not eliminate all scenarios in which packet collisions occur. Consider the example in Figure 1. At time t_1 node A sends and RTS to node B. At time t_2 node B replies with a CTS packet. However, at t_2 node C also starts transmitting an RTS packet to node D. Due to this transmission, node C does not hear the CTS packet from B. Now if C receives a CTS reply from D, it will initiate its packet transmission and this packet will collide with the packet sent from A to B. The four way handshake of MACAW, namely RTS-CTS-data-ACK does not eliminate the packet collision described in this example; it only detects it after it occurs.

Floor acquisition multiple access (FAMA) was modified in [5] in order to protect the data packets from collisions in a mobile ad-hoc network with hidden terminals. This protocol, which is referred to as floor acquisition multiple access, non- persistent (FAMA-NPR) in [5], is briefly

described in the following

In FAMA-NPR, non-persistent carrier sensing is used in conjunction with RTS-CTS dialogue in order to obtain the state of the channel. The transmission time of the CTS packet is longer than the aggregate transmission time of the RTS packet, the maximum channel propagation delay, the transmit to receive turn-around and processing time. This is intended to enable the CTS packets to dominate the RTS packets in the channel. Consequently, if a node initiates transmission of an RTS packet simultaneously with a CTS transmission, that node will be able to hear at least a portion of the CTS packet and will back off.

Consider the example of Figure 1. Let the maximum propagation delay be denoted by τ. Suppose node A sends an RTS packet to B to which B replies with its CTS packet. Also suppose that node C initiates an RTS packet transmission simultaneously with B's transmission. Since node C uses carrier sensing before transmitting its RTS packet, its transmission can begin no earlier than τ seconds before B initiates its CTS transmission. The choice of the lengths for CTS and RTS packets implies that node C will hear the tail end of the CTS packet and would back off. On the other hand the latest that C can begin its RTS transmission is τ seconds after B begins its CTS transmission. Again C will be able to hear a portion of the CTS packet and would back off. Correctness of FAMA-NPR is demonstrated in [5] by proving that data packets will not collide with any other transmission.

In [5] the authors do not provide any detail on how a node would detect the front or the tail end of a CTS packet (and consequently back off). It is merely stated that if an RTS packet overlaps a CTS transmission, since the CTS "dominates" the RTS packet, the node transmitting the RTS will hear the front or the tail end of the CTS as "noise".

The modified FAMA protocol eliminates packet collisions in a network with hidden terminals.

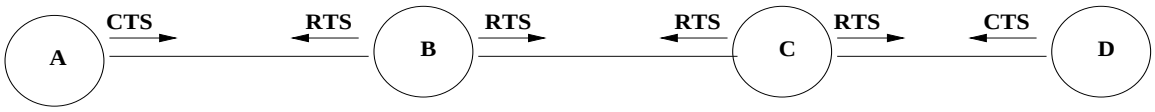

Figure 2: The exposed terminal problem in FAMA.

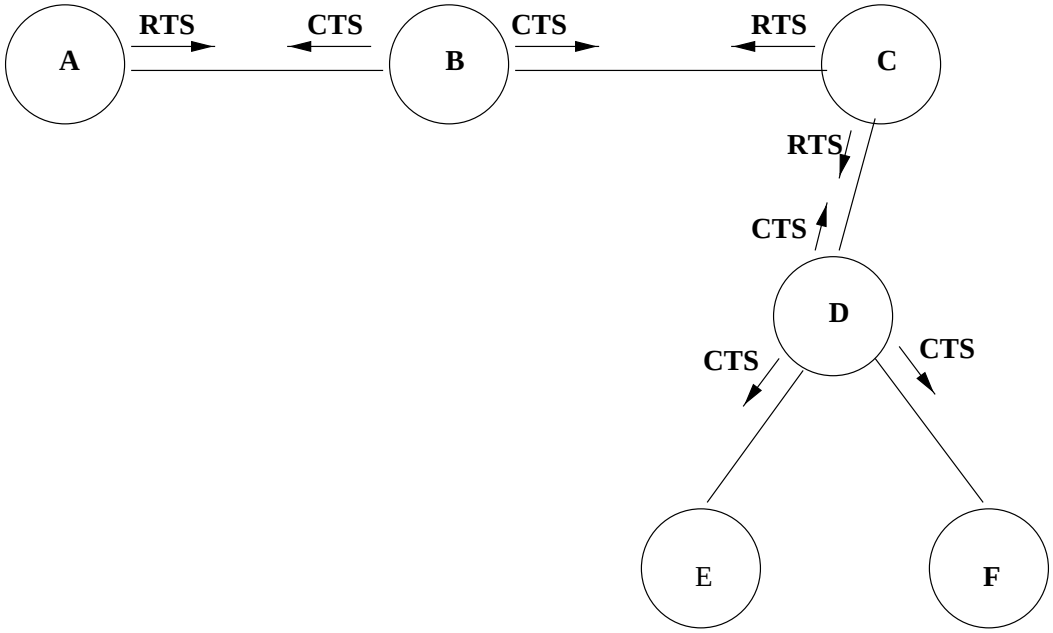

Figure 3: The waiting terminal problem in FAMA.

However, the protocol introduces other problems which, we believe, may result in a significant reduction in network throughput. In the following we give two examples of such problems. The first is that of an overlapping RTS packet in the exposed terminal scenario. Consider the network depicted in Figure 2 where the connectivity of nodes is shown by the solid lines. Nodes B and C send RTS packets to A and D respectively, with C's transmission delayed with respect to that of B. After sending its RTS packet, B detects noise in the channel due to the tail end of the RTS packet sent by C. Consequently B will back off. Clearly this is unnecessary. One remedy to this problem is to include a flag at the beginning and the end of each CTS packet in order to distinguish CTS packets from RTS packets. We must then ensure that the flag does not appear as part of an RTS packet (e.g., using bit stuffing when necessary).

The second problem which we refer to as the waiting terminal problem is due to the fact that backoff occurs at the transmitter without any notification of the corresponding receiver. Consider the network of Figure 3. Node B replies to the RTS sent by A with its CTS packet. Node C also sends an RTS to D which overlaps the CTS sent by B. C hears this CTS packet partially and will back off. However, since D has received the RTS from C it will reply with its CTS packet. This causes D and all its neighbors (E and F in the figure) to remain silent for the duration of the would be transmitted packet from C. Clearly this is unnecessary and causes a reduction in the overall network throughput. Furthermore, this difficulty can not be easily overcome by modifying the RTS and CTS packets.

Simulation results are presented in [5] showing that FAMA-NPR achieves the same throughput as MACAW. However, the network examples considered there are very simple and do not include such scenarios as discussed above.

Conclusion

Although FAMA protocols ensure that data packets will not collide with any other transmission, in certain cases their throughput may be significantly lower than MACAW. Furthermore, as in the case of MACA and MACAW, FAMA protocols only support one priority class and can not provide QoS guarantees to multimedia applications.

3.4 The MAC Protocol of IEEE 802.11

The basic access method in the 802.11 MAC protocol is the "Distributed Coordination Function" (DCF) which is best described as the Carrier Sense Multiple Access with Collision Avoidance (CSMA/CA). In addition to the DCF the 802.11 also incorporates an alternative access method

known as the "Point Coordination Function" (PCF) - an access method that is similar to "polling" and uses a point coordinator to determine which station has the right to transmit. Further, an optional "Distributed Time Bounded Service" (DTBS) may be provided by the DCF. DTBS is the "best effort" service that provides bounded delay and delay variance. In the following we describe the DCF access method (PCF requires a central controller and is not suitable for ad-hoc networks.).

When using the DCF, a station, before initiating a transmission senses the channel to determine if another station is transmitting. The station proceeds with its transmission if the medium is determined to be idle for an interval that exceeds the "Distributed InterFrame Space" (DIFS). In case the medium is busy the transmission is deferred until the end of the ongoing transmission. A random backoff interval is then selected which is used to initialize the backoff timer. The backoff timer is decremented only when the medium is idle; it is frozen when the medium is busy. After a busy period the decrementing of the backoff timer resumes only after the medium has been free longer than DIFS. A station initiates a transmission when the backoff timer reaches zero.

Acknowledgements are employed to determine the successful reception of each data frame. This is accomplished by the receiver initiating the transmission of an acknowledgement frame after a time interval "Short InterFrame Space" (SIFS), that is less than DIFS, immediately following the reception of the data frame. Note that the acknowledgement is transmitted without the receiver sensing the state of the channel. In case an acknowledgement is not received the data frame is presumed lost and a retransmission is scheduled (by the transmitter). This access method is referred to as "Basic Access" in the following.

The DCF also provides an alternative way to transmitting data frames that involve trans-

mission of special short Request To Send (RTS) and Clear To Send (CTS) frames prior to the transmission of the actual data frame. As in MACA, etc., a successful exchange of RTS and CTS frames attempts to reserve the channel for the time duration needed to transfer the data frame under consideration. The rules for the transmission of an RTS frame are the same as those for a data frame under basic access. On receiving an RTS frame the receiver responds with a CTS frame (the CTS frame acknowledges the successful reception of an RTS frame), which can be transmitted after the channel has been idle for a time interval exceeding SIFS. After the successful exchange of RTS and CTS frames the data frame can be send by the transmitter after waiting for a time interval SIFS. In case a CTS frame is not received within a predetermined time interval, the RTS is retransmitted following the backoff rules as specified in the basic access procedures outlined above.

The RTS and CTS frames contain a duration field that indicated the period the channel is to be reserved for transmission of the actual data frame. This information is used by stations that can hear either the transmitter and/or the receiver to update their "Net Allocation Vectors" (NAV) - a timer that is always decreasing if it value is non-zero. A station is not allowed to initiate a transmission of its NAV is non-zero. The use of NAV to determine the busy/idle status of the channel is referred to as the "Virtual Carrier sense" mechanism.

The different interframe spacings of IEEE 802.11 allows for differentiation between transmissions. For example since SIFS is shorter than DIFS, an acknowledgement always has access priority over a new data frame. This idea is used in the black burst contention scheme of [13] to provide access priority for real-time traffic. Provisioning of QoS however, is more difficult and is only present for Point Coordination Function.

3.5 Black Burst Contention

In the previous sections, we showed the inability of CSMA, MACA, MACAW and FAMA protocols to provide QoS guarantees for multimedia applications as well as to support different priority classes of traffic. Furthermore, we showed this to be the case for a fully connected ad-hoc WLAN as well as mobile ad-hoc networks with hidden terminals. A new protocol referred is as black burst contention (BBC) was recently introduced in [13] which provides delay-bound guarantees for real-time traffic in an ad-hoc WLAN. In this scheme real-time packets are not subject to collisions and have access priority over non-real-time data packets. We briefly describe this protocol in the following.

In the BB contention scheme of [13] two interframe spacings are defined, namely t_{med} and t_{long}. A real-time node contends for channel access after sensing the channel to be idle for a duration of t_{med}, whereas a non-real-time node contends for channel access after it perceives the channel to be idle for a period of t_{long}. t_{med} and t_{long} are chosen such that $t_{long} > t_{med} + 2\tau$ where τ is the maximum propagation delay in the channel. The different waiting times for real-time and non-real-time packets enables prioritized access for real-time traffic. It should be noted that this is similar to the different interframe spaces in the IEEE 802.11 MAC protocol.

In order to sort the access rights of real-time nodes based on their delay, real-time nodes are required to transmit a series of pulses after the period t_{med}. The number of these pulses which are referred to as black burst (BB) is proportional to the access delay the node has experienced. In particular, if d is the delay incurred by the node, the the number of black slots, $b(d)$, is given by

$$b(d) = 1 + \lfloor \frac{d}{t_{unit}} \rfloor$$

where $\lfloor x \rfloor$ is the largest integer no larger than x and where t_{unit} is the unit of time used to convert the delay into an integral number of black bursts.. After exhausting its BB transmission, the node waits for an observation period t_{obs} to see if any other node transmitted a longer BB, indicating it has had a longer delay. If the channel is perceived to be idle after t_{obs}, then the node transmits its packet. The length of the observation period is chosen to satisfy $t_{obs} < t_{med}$ so that another node may not initiate its BB transmission during this waiting period. By appropriately choosing the length of BB slots and a minimum packet length the authors show that real-time packets will be free of collisions [13]. Furthermore, they demonstrate that for a fully connected ad-hoc WLAN (without hidden nodes) the protocol guarantees bounded delay for real-time nodes.

Conclusion

Black burst contention uses carrier sensing to obtain the state of the channel. Access is initiated if the channel is found to be idle and the interframe spacing for the traffic class is satisfied. For a fully connected ad-hoc WLAN, this approach works well. By increasing the number of interframe spacings one can accommodate a larger number of priority classes.

Clearly for a network with hidden terminals this protocol suffers from the same shortcomings as the CSMA protocol. The lack of an RTS-CTS exchange implies that packet collisions will occur at the receiver. Due to these packet collisions bounded delay can no longer be guaranteed for real-time traffic. Moreover, while on average one may expect real-time nodes to experience

less access delay, this can not be guaranteed for every real-time node. Although the real-time node has access priority due to its shorter interframe spacing, due to collisions at the receiver, its actual delay may be higher than some non-real-time nodes.

In the following section, we present a new MAC protocol which combines the RTS-CTS handshake with the BB contention mechanism in a dynamic, distributed TDMA system. The new protocol is capable of multiplexing different priority classes and scheduling such traffic in appropriate order.

4 A Dynamic TDMA Protocol for Ad-hoc Networks

4.1 Introduction

In this section we present a new MAC protocol for mobile ad-hoc networks based on a dynamic time division multiple access (TDMA) scheme. The new protocol is a distributed MAC protocol which can flexibly accommodate different traffic rates through dynamic slot allocation. It accommodates different priority classes by providing lower access delay to traffic sources of higher priority. Furthermore, the protocol achieves fair delay and throughput performance for traffic sources within a given priority class.

In this report we have considered three priority classes, namely emergency traffic, real-time traffic and non-real-time traffic. However, the protocol can be easily extended to accommodate more levels of priority. This implies that the emergency traffic will experience the lowest average delay followed by the real-time traffic and non-real-time traffic, in that order.

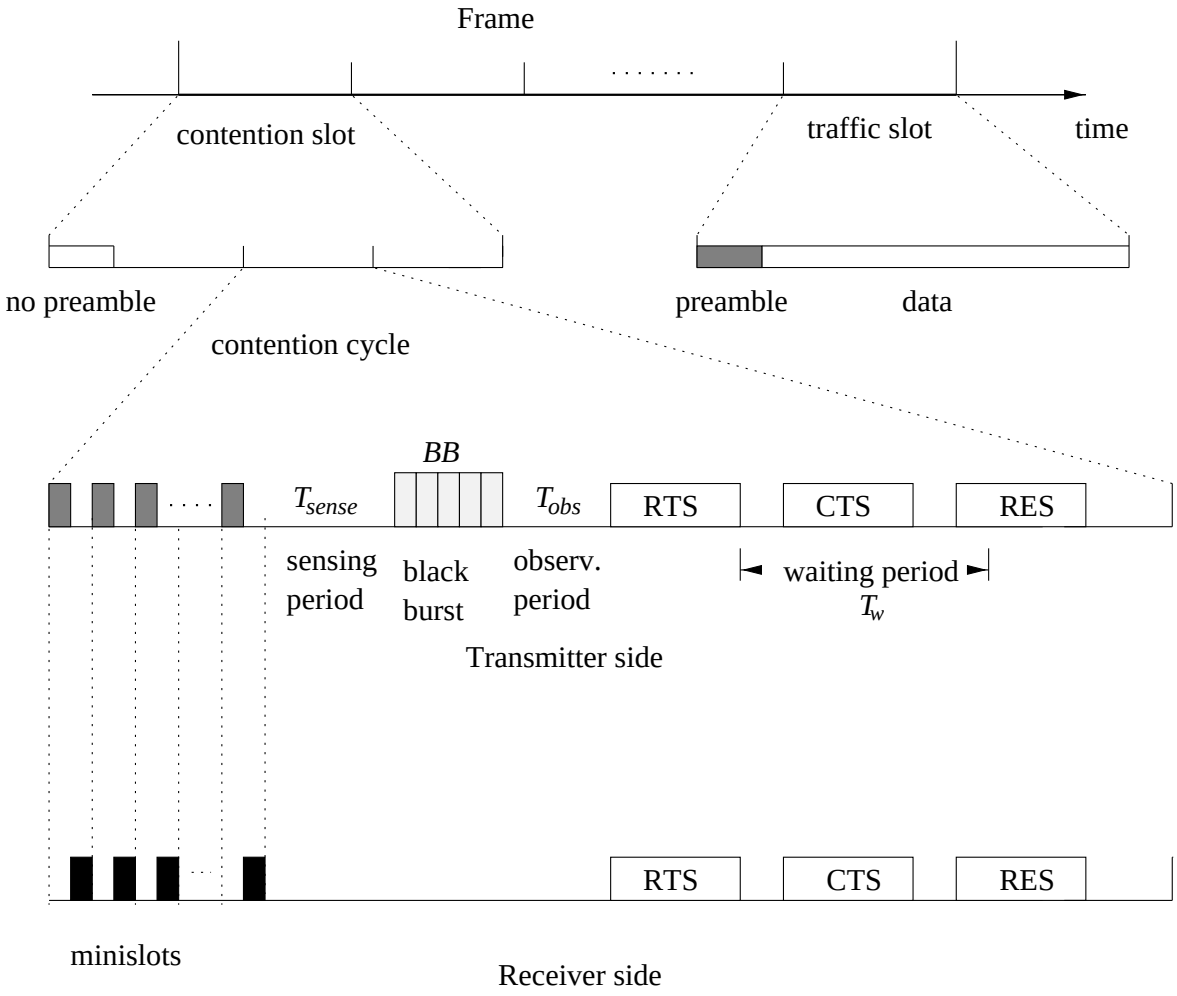

Figure 4: Frame Structure

4.2 Frame Structure

The frame structure for the MAC protocol is shown in Figure 4. Time is divided into fixed-size frames. Each frame is further divides into N slots. A slot is either a traffic slot or a contention slot as described in the following.

4.2.1 Traffic slot

A traffic slot consists of a preamble period followed by the packet transmission period. Both the sender and the receiver transmit a preamble during this period to inform their neighbors

that this slot is a traffic slot and should not be accessed for contention. Packet are assumed to be fixed length. The time needed to transmit a packet is equal to the length of the packet transmission period.

4.2.2 Contention slot

The absence of the preamble signal at the beginning of a slot indicates that the current slot is not being used for traffic and thus can be utilized as a contention slot. A contention slot is further divided into several contention cycles. A contention cycle consists of several time periods. These periods and their functions in the media access control protocol are described in the following

Mini-slots: Each contention cycle starts with a sequence of N mini-slots. Mini-slot i corresponds to slot i in the frame. A minislot is further subdivided into two intervals. In the first intervals the nodes that use the corresponding slot to transmit packets transmit a short "busy" signal, while in the second interval the nodes that use the corresponding slot to receive packets transmit their "busy" signal.

By observing the mini-slot sequence a node can determine which slots in the frame are being used by its neighbors. If a node intends to reserve a new slot for transmission, by observing the second interval of all the mini-slots it can determine which slots in the frame are being used for transmission to its neighbors (its neighbors are receiving packets in those slots). Clearly the node must not transmit in those slots. The remaining slots in the frame can be used for transmission by this node without causing a collision with any of its neighbors. Similarly, by observing the first interval of all the minislots, a node can determine which slots are not being used by its neighbors for transmission to other nodes. Only in these slots can this node receive

packets free of collision.

Channel sensing period: The sequence of mini-slots is followed by the *channel sensing period* which is denoted by T_{sense}. A node wishing to reserve a slot must sense the channel and only if it finds the channel to be idle for this period can it proceed to access the channel. Various channel sensing periods are used to differentiate between different priority classes. An example for the length of this period is shown in Table 4.2.2. It can be seen that the emergency traffic as a group have the highest level of priority followed by the real-time traffic and the non-real-time traffic.

Traffic type	T_{snese}
Emergency traffic	One unit
Real-time traffic	Two units
Non-real-time traffic	Three units

Black burst: After sensing the channel to be idle for a period of T_{sense}, the nodes in a priority class sort their access rights by jamming the channel with pulses of energy, referred to as black bursts [13], and denoted by BB . As in [13], the length of BB is an increasing function of the contention delay experienced by the node, measured, in number of slots, from the time that the first attempt to access the channel is made until the beginning of the current contention slot.

Request to send: If the channel is perceived to be idle after the observation period, the node assumes that it is the sole winner of the contention cycle. It now transmits a *request to send* (RTS) packet. The RTS packet is a mini-packet which contains the source ID, the destination ID, and a list of candidate slots that the source proposes to transmit in. This list is a subset

of the slots that are unused in the neighborhood of the source. As mentioned previously in the description of the mini-slots, the list of unused slots in the neighborhood of the source is obtained from the observation of the second interval of all the mini-slots. Since there are N slots in each frame, each candidate slot will be denoted by $\log_2 N$ bits.

Clear to send: Upon successfully receiving the RTS packet, the destination node extracts the list of candidate slots from the RTS packet and compares this list with the list of unused slots in its neighborhood. (Again, as discussed before in describing the mini-slots, by observing the first interval in the sequence of mini-slots, the node obtains the list of the unused nodes in its neighborhood.) If at least one common slot exists between the two lists, the node chooses one such slot and reports back to the source with a *clear to send* (CTS) packet. The clear to send (CTS) packet contains the source ID, the destination ID and the slot chosen by the destination. If a common slot can not be found between the two lists, the destination node remains silent and does not send the CTS packet.

Waiting time: The transmission of the RTS is followed by a waiting time T_w during which the *clear to send* (CTS) packet from the receiver is expected. If the CTS packet is received from the destination, the reservation process continues. Otherwise it is abandoned.

Reservation packet: If the CTS packet is received during the waiting period T_w, then the source node sends the RES packet which echoes the content of the CTS packet. The RES packet informs the destination node that the reservation has been successful and that transmission will start in the designated slot in the following frame. It also informs the neighbors of the source as to which slot has been chosen for the transmission of the impending packet.

From a transmitter's point of view, after the mini-slot sequence, the channel sensing period is followed by the black burst followed by the observation period. This is then followed by the

RTS-CTS-RES dialogue which indicate a successful reservation, or a single RTS packet, which indicates an unsuccessful reservation attempt.

From the receiver's point of view, the contention cycle also starts with a mini-slot sequence, followed by the RTS-CTS-RES dialogue for a successful reservation, or a single RTS, or RTS-CTS pair, for an unsuccessful reservation attempt.

We would like to point out that although all the stations are synchronized, the designation of a slot as a traffic or contention slot varies spatially. A particular slot in the frame may be a traffic slot in one part of the network and a contention slot in another.

4.3 State Transition Diagram

A node can be in one of four states, designated as Idle, Transmit, Receive and Backoff states. These are described in the following. Figure 5 shows the state transition diagram.

Idle state

The default state is the Idle state. While in this state the node continuously monitors the channel and collects the following information.

1. Whether the current slot is a traffic slot or a contention slot, by observing the preamble at the beginning of the slot.

2. Maintains a counter C_{frac} which indicates the number of reserved slots in the frame. If this counter exceeds a threshold, then this node can not contend for transmission of real-time traffic.

3. By observing the sequence of minislots determine which slots are being utilized for transmission/reception in its neighborhood and maintain the two lists.

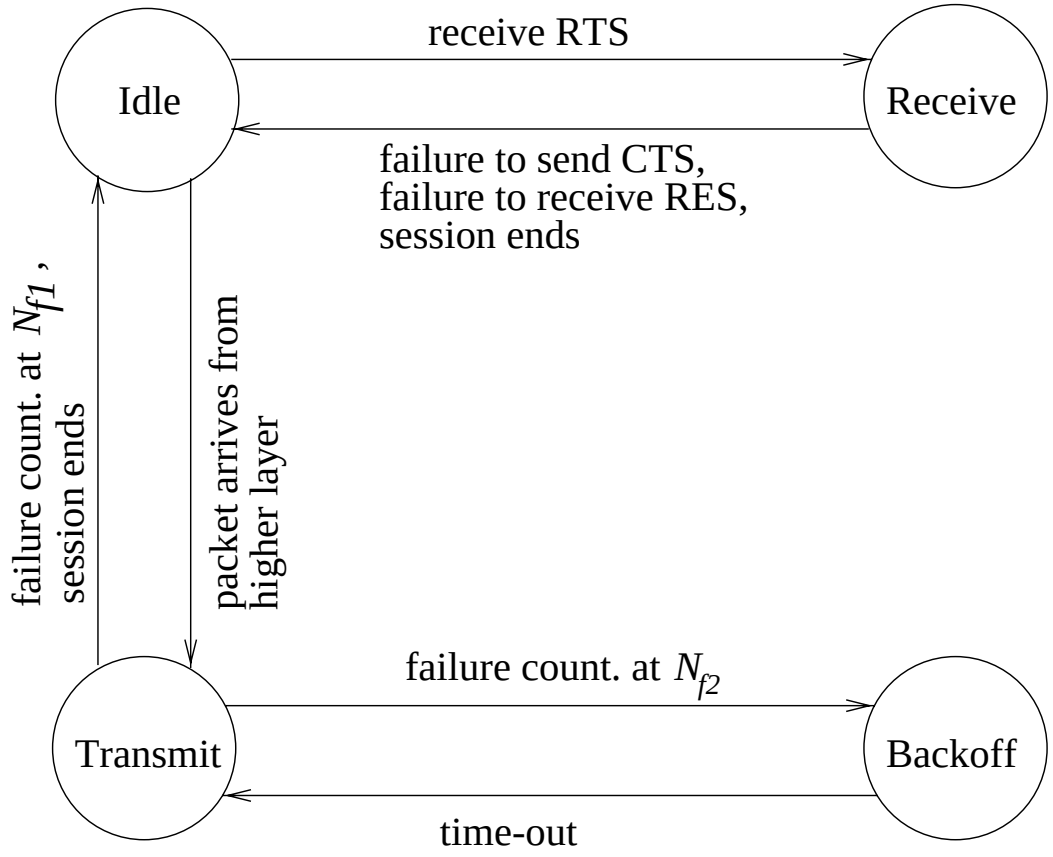

Figure 5: State Transition Diagram

Transmit state

When the MAC layer receives a packet from the higher layer, the node moves from the Idle state into the Transmit state. Depending on whether the packet is an emergency, real-time or non-real-time packet, the behavior of the node will be different. In the following we describe each case separately.

Non-real-time traffic

Nodes transmitting non-real-time traffic must make their reservation on a packet by packet basis. Therefore such nodes must wait for the next contention cycle in the next contention slot to make their reservation. The reservation procedure will be explained shortly.

Real-time traffic

Two modes of operation is considered for nodes transmitting real-time traffic. In the first mode these nodes operate in the same way as non-real-time nodes in that they make their reservation on a packet-by-packet basis. In the second mode, once they reserve a slot, real-time nodes maintain their reservation for future frames until they relinquish it. This is done by transmitting a last packet which indicates that the slot is released. Therefore, when a node receives packets from a higher layer, it is possible that the node already has a slot reserved. In that case the node will transmit its packets in that slot in the future frames. If the node does not have a reservation, it must make a reservation. To do so the node first compares the counter C_{frac} to the threshold Th_{frac}. If the counter does not exceed the threshold, the node will attempt to make its reservation. Otherwise it must abandon its reservation attempt. This is intended to leave a certain fraction of the slots in each frame for emergency and non-real-time traffic and prevent the real-time traffic from "hugging" the channel.

Emergency traffic

As in the case of real-time traffic, nodes transmitting emergency traffic also maintain their slot reservation for future frames until they relinquish it. Therefore, if the node already has a reservation, it will transmit its packets in those reserved slots in the future frames. On the other hand if the node does not have a reservation, it will proceed to make a reservation. Unlike the case of real-time traffic however, the node does not compare its counter C_{frac} to the threshold value Th_{frac}.

Reservation

To make a reservation, the node waits for the first contention slot. As discussed in Section 4.2, the contention slot is indicated by the absence of the preamble, followed by a sequence of mini-slots. By observing the second interval in all the minislots the node can determine which slots in the frame are not reserved and thus can by used for its transmission.

After the minislot sequence is over, the node listens to the channel for the channel sensing period T_{sense}. If the channel remains idle during this period, the node transmits its black burst BB followed by the observation period T_{obs} where it remains silent and monitors the channel. Again, if the channel remains idle for T_{obs}, the node proceeds with its transmission of the request to send (RTS) packet. Following the transmission of the RTS packet, the node remains silent for the period of T_w waiting for the reception of the clear to send (CTS) packet from the receiver. As in the MACA algorithm, all the nodes that hear the RTS packet must also remains silent for the period of T_w to allow the CTS packet to be successfully received by the source node [6]. If the CTS packet is received from the destination, the source node echoes the CTS packet back to the destination in the form of the RES packet. In the RTS-CTS-RES exchange the source and destination reserve the slot which will be used for their communication in the following frame(s). Note that while the RTS packet may be involved in a collision due to hidden terminal

transmissions, the CTS packet will not collide with any transmission.

The CTS packet may not arrive within the waiting period T_w. There are two possible reasons. The first which is similar to the case in MACA and MACAW, is when the RTS packet is involved in a collision at the destination due to a transmission from one of the neighbors of the destination node. The second reason which is unique to this MAC protocol is when the destination node does not find any common slots between the list of slots contained in the RTS packet and the list of slots that it finds to be unused in its neighborhood.

If the CTS packet does not arrive within the T_w period, a failure is declared and a counter denoted C_{fail} is incremented. C_{fail} keeps track of the number of access failures and is reset every time the node enters the Idle state. When C_{fail} exceeds the constant N_{f1}, the node enters the Backoff state. If C_{fail} exceeds the constant N_{f2}, the node enters the Idle state and reports the failure to the higher layer.

Receive state

When a node receives an RTS packet that is addressed to it, it moves to Receive state. Clearly a node can reside in both Transmit and Receive states during different slots. A node that has already made a reservation will listen for the preamble in every slot and if the preamble is not present it determines that the slot is a contention slot. In the sequence of minislots that are present in the contention slot, the node must transmit it busy signal in the second interval of the minislot corresponding to the slot which it has reserved (through the RTS-CTS-RES exchange). In the reserved slot the node first transmits its preamble. Following the preamble it begins to receive the packet. In addition, by observing the minislots, the node must determine the slots that are not being used in its neighborhood so that it can appropriately respond to RTS packets that it may receive.

After receiving an RTS packet, the node must reply with its CTS packet. If the node can not send the CTS packet immediately or within the waiting period T_w, it just ignores the RTS and moves back to Idle state. After sending the CTS, the node waits for the RES packet. If no RES packet arrives within T_w, it moves back to Idle state. On the other hand if a RES is received, the node will consider the reservation as successful and will behave as a node which has a reservation as described above.

4.4 Backoff State

A parameter $N_{backoff}$ is defined. The number of contention cycles that a node should backoff is a random number uniformly distributed in the range of $n * N_{backoff}$ where $n = 1$ for emergency traffic, $n = 2$ for real-time traffic, and $n = 3$ for data traffic. During the traffic slots the backoff counter is frozen. It is only decremented during the contention slots. When the backoff counter reaches zero, the node moves to the Transmit state.

Conclusion

In this protocol emergency packets have the highest access priority followed by real-time and non-real-time packets. In the case where real-time nodes make their reservation on a packet-by-packet basis, in an ad-hoc WLAN this protocol guarantees bounded-delay for emergency traffic. Furthermore, real-time traffic will have a smaller average delay than non-real-time traffic.

In a mobile ad-hoc network with hidden-terminals, no bounded delay guarantees can be provided. This is due to the collisions that the RTS packets may incur at the receiver. However, emergency traffic will have access priority and thus the lowest average delay followed by real-time and non-real-time-traffic.

5 Conclusions

In this report we have evaluated the capabilities of several existing media access control (MAC) protocols in supporting different priority classes and differentiated quality of service in a mobile ad-hoc network. Of the protocols considered, only black burst contention (BBC) protocol provides access priority for real-time traffic over non-real-time traffic. Furthermore, in an ad-hoc WLAN with no hidden terminals BBC provides delay bound guarantees to real-time traffic.

Since it relies on carrier sensing, black burst contention is not suitable for mobile ad-hoc networks with hidden terminals. We have presented a new MAC protocol based on a dynamic, distributed time division multiple access system. In this protocol data packets will not suffer collisions. Furthermore, this protocol provides access priority to emergency traffic over rea-time and non-real-time traffic. In an ad-hoc WLAN, emergency traffic will have bounded delay. Real-time traffic will have a lower average delay that non-real-time traffic. In a mobile ad-hoc network which may include hidden terminals, bounded delay can not be guaranteed due to the collisions of the RTS packets. However, emergency traffic will have the highest access priority followed by real-time and non-real-time traffic. This ensures the lowest average delay for the emergency traffic followed by real-time and non-real-time traffic.

References

[1] N. Abramson, "The Aloha System-Another Alternative for Computer Communications," *Proc. of the Fall Joint Computer Conference*, pp. 281-285, 1970.

[2] V. Bharghavan, A. Demers, S. Shenker and L. Zhang, "MACAW: A Media Access Protocol for Wireless LAN's," in Proc. ACM SIGCOMM'94, pp. 212-25, London, UK, Aug. 31-Sept. 2, 1994.

[3] A. Colvin, "CSMA with collision avoidance," Computer common, vol. 6, no. 5, pp. 227-35, 1983.

[4] C. L. Fullmer and J.J. Garcia-Luna-Aceves, "Floor Acquisition Multiple Access (FAMA) for Packet-Radio Networks," Proc. ACM SIGCOMM 95, Cambridge, MA, Aug. 28-Sept. 1, 1995.

[5] C. L. Fullmer and J. J. Garcia-Luna-Aceves, "Complete Single-Channel Solutions to Hidden Terminal Problems in Wireless LANs," Proc. of IEEE International Conference on Communication, 1997, pp. 575-579

[6] P. Karn, "MACA - A New Channel Access Method for Packet Radio," in ARRL/CRRL Amateur Radio 9th Computer Networking Conference, pp. 134-40, ARRL, 1990.

[7] L. Kleinrock and F.A. Tobagi, "Packet switching in radio channels: Part I - carrier sense multiple-access modes and their throughput-delay characteristics," IEEE Trans. common, vol. COM-23, no. 12, pp. 1400-1416, 1975.

[8] B. M. Leiner, D. L. Nielson and F.A. Tobagi, ads., *Proc. of the IEEE*, vol. 75, IEEE, Jan. 1987.

[9] W.F. Lo and H.T. Mouftah, "Carrier Sense Multiple Access with Collision Detection for Radio Channels," *IEEE 13th int'l Commun. and Energy Conf.* pp. 244-247, IEEE, 1984.

[10] R. Rom, "Collision Detection in Radio Channels," *Local Area and Multiple Access Networks*, pp. 235-249, Computer Science Press, 1986.

[11] G.S. Sidhu, R.F. Andrews and A.B. Oppenheimer, *Inside AppleTalk, Second Edition*, Addison Wesley Publishing Company, Inc., 1990.

[12] F.A. Tobagi and L. Kleinrock, "Packet switching in radio channels: Part II - the hidden terminal problem in carrier sense multiple-access modes and the busy-tone solution," IEEE Trans. Commun., vol. COM-23, no. 12, pp. 1417-1433, 1975.

[13] J.L. Sobrinho and A.S. Krishnakumar, "Quality-of-Service in Ad Hoc Carrier Sense Multiple Access Wireless Networks," *IEEE Journal on Selected Areas in Communications*, vol 17, no. 8, pp. 1353-1368, August 1990.

[14] F.A. Tobagi and L. Kleinrock, "Packet switching in radio channels: Part III- polling and (dynamic) split-channel reservation multiple access," IEEE Trans Commun., vol. COM-24, no. 8, pp. 832-845, 1976.

[15] C. Lin and M. Gerla, "Asynchronous multimedia multihop wireless networks, " in Proc. IEEE INFOCOM '97, Kobe, Japan, pp. 118-125.

[16] IEEE Standard for Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications, IEEE Standard 802.11, 1997.