3/8/01

Implementation of Priority Schemes in MANETs Over the 802.11 MAC Layer

Leonard E. Miller
Principal Investigator

Xavier Pallot
Guest Researcher

Wireless Communications Technologies Group
Advanced Network Technologies Division
Information Technology Laboratory
National Institute of Standards and Technology
Gaithersburg, Maryland

Table of Contents

Table of Figures

ABSTRACT

Several mechanisms for giving preferential treatment of priority traffic are proposed for the IEEE 802.11 wireless local area network (WLAN) medium access control (MAC) layer. The mechanisms include (1) using separate queues for different priority levels of packets, and servicing the higher priority queues first; (2) giving priority traffic a shorter waiting (deferral) time for access to the medium; and (3) giving priority traffic a nonuniform distribution of backoff times when contending for access to the medium, such that the average backoff time of priority traffic is less, thereby giving the priority traffic an advantage in the contention.

Results are given for an OPNET simulation implementing the proposed mechanisms for two levels of priority in a network radio node model that includes the dynamic source routing (DSR) mobile ad hoc network (MANET) routing protocol for multihop operation. The scenario chosen for the simulations is designed to induce congestion at a central node that must carry all the traffic originating from four other nodes. The simulation results indicate that the priority mechanisms can be effective in providing smaller delays to the priority traffic. Further studies are proposed for determining best values of parameters for the mechanisms under typical WLAN and MANET operational configurations.

This work was supported by a grant from the Technology and Programs Division, National Communications System.

Implementation of Priority Schemes in MANETs Over the 802.11 MAC Layer

1. Introduction

This report summarizes work in the Wireless Communications Technologies Group (WCTG) of the National Institute of Standards and Technology (NIST) with the objective of promoting the improved performance of wireless mobile ad hoc networks (MANETs) for possible emergency communication applications. Supported by a grant from the Technology and Programs Division of the National Communications System, the work during the period covered by this report focused on the investigation of how certain modifications to a standard medium access control (MAC) layer protocol, the IEEE 802.11 wireless local area network (LAN) standard, can enhance the performance of MANETs when the traffic over the network includes priority traffic.

1.1 Background on MANETs

In the next generation of wireless communication systems, there is a need for the rapid deployment of independent mobile users operating together as an *ad hoc* network [1, 2]. Significant examples include establishing survivable, efficient, dynamic communication for emergency/rescue operations, disaster relief efforts, and military networks. Advances in information technology for these important types of situations are envisioned for future wireless communications. Such network scenarios cannot rely on centralized and organized connectivity, and can be termed wireless *mobile ad hoc networks* (MANETs). A MANET is an autonomous collection of mobile users (nodes) that communicate over wireless links with relatively low bandwidth because of the limited frequencies allocated from the available spectrum and because of the impairments typically present on the mobile wireless propagation channel. Due to nodal mobility, the network topology may change rapidly and unpredictably over time. The network is *decentralized*, where network organization and message delivery must be executed by the nodes themselves, *i.e.*, routing functionality will be incorporated into mobile nodes. Nodes must also contend with the effects of radio communication, including multiuser (multiple access) interference, multipath fading, and shadowing. A MANET may operate in a stand-alone manner, or be connected to a larger network, *e.g.*, the fixed Internet.

The design of network protocols for MANETs is a complex issue [29]. These networks need efficient *distributed* algorithms to determine network organization (connectivity), link scheduling, and routing. An efficient approach is to consider routing algorithms in which network connectivity is determined in the process of establishing routes. Message routing in a decentralized environment where network topology fluctuates is not a well-defined problem. While the shortest path (based on a given cost function) from a source to a destination in a static network is usually the optimal route, this idea is not easily extended to MANETs. Factors such as power expended, variable wireless link quality, propagation path loss, fading, multiuser interference, and topological changes, become relevant issues. The network should be able to adaptively alter routing paths to alleviate any of these effects. Moreover, in a military environment, preservation of security, latency, reliability, intentional jamming, and recovery from

failure are significant concerns. Military networks must maintain a *low probability of intercept* and/or a *low probability of detection*. Hence, nodes prefer to radiate as little power as necessary and transmit as infrequently as possible, thus decreasing the probability of detection or interception. A lapse in any of these requirements may degrade network performance and dependability.

1.1.1 Properties of MANET Routing Protocols

Various protocols have been proposed in the Internet Engineering Task Force (IETF) for performance of routing functions in a MANET [3]. The set of applications for MANETs is diverse, ranging from small, static networks that are constrained by power sources, to large-scale, mobile, highly dynamic networks. It is unlikely that a single routing protocol will be optimal for all scenarios. A given protocol, for performing the routing or any other function, will execute efficiently in those networks whose characteristics are in accord with the mechanisms used by the protocol. However, any protocol must efficiently handle several inherent characteristics of MANETs [29]:

- *Dynamic topology*: Mobility of nodes lends to unpredictable network topology.
- *Variable capacity wireless links*: Wireless links are bandwidth-constrained. Moreover, since wireless links have lower capacity than hardwired links, traffic congestion is typical rather than atypical. However, as a MANET is often an extension of a fixed network, the same services and demands must be accommodated. These demands will increase as multimedia computing and networking become more mainstream.
- *Power constrained operation*: Power conservation is *crucial* in mobile wireless systems since these networks typically operate from power-limited sources, which dictate whether a network is operational or not.
- *Physical security*: Mobile networks are more vulnerable to physical security threats such as eavesdropping and jamming attacks.

The merit of a routing protocol is judged with performance metrics, both qualitative and quantitative. Desirable qualitative properties of a MANET routing protocol include the following [29]:

- *Distributed*: The decentralized nature of a MANET requires that any routing protocol execute in a distributed fashion.
- *On demand operation*: Since a uniform traffic distribution can not be assumed within the network, the routing algorithm must adapt to the traffic pattern on a demand or need basis, thereby utilizing power and bandwidth resources more efficiently.
- *Loop-free*: To ensure proper message delivery and efficient network operation, a routing protocol must be loop-free.
- *Security*: Since MANETs are more vulnerable to physical security threats, provisions for security must be made, e.g., the application of Internet Protocol (IP) security techniques.
- *Entering/Departing nodes*: A routing protocol should be able to quickly adapt to entering or departing nodes in the network, without having to restructure the entire network.

- *Bidirectional/Unidirectional links*: Since the condition of a MANET is dynamic, a routing protocol should be able to execute on both bidirectional and unidirectional links.

1.1.2 MANET Routing Protocol Selected for Study

For the studies described in this report, the MANET routing protocol used was the Dynamic Source Routing (DSR) protocol [3, 4], as implemented as an OPNET simulation model by NIST [5]. DSR was designed especially for MANET applications. Its main feature is that every data packet follows the source route stored in its header. This route gives the address of each node through which the packet should be forwarded in order to reach its final destination. Each node on the path has a routing role and must transmit the packet to the next hop identified in the source route.

Each node maintains a Route Cache in which it stores every source route it has learned. When a node needs to send a data packet, it checks first its route cache for a source route to the destination. If no route is found, it attempts to find one using the route discovery mechanism. A monitoring mechanism, called route maintenance, is used in each operation along a route. This mechanism checks the validity of each route used.

DSR route discovery works as follows: If node *S* wants to communicate with node *D*, it needs to find a route on demand by using the route discovery mechanism. Node *S* broadcasts a Route Request packet in the network. This Route Request contains the address of the initiator, the address of the target, a field sequence number (set by the initiator and used to identify the request), and a route record. The latter is the field where a record of the sequence of hops taken by the Route Request is accumulated dynamically. Each node in the network maintains a table in order to detect a duplicate Route Request packet received.

A node propagates the Route Request if it is not the target and if it is the first time it receives this packet. The first node receiving this Request that has a valid route in its route cache for node *D* initiates a Route Reply packet back to node *S*. This Route Reply contains the list of nodes along the path from node *S* to node *D*. The first part is the information gathered along the path of the Route Request (that is, from node *S* to the node replying); the rest of the list is the information found in the route cache of the replying node. Moreover, it may occur that destination node *D* itself receives a Route Request packet, *e.g.* no node along the way before node *D* has an accurate route from itself to node *D* in its route cache. In this case, node *D* sends a Route Reply packet containing the path just created dynamically from source *S* to destination *D*, *i.e.*, the path traversed by the first Route Request packet received by node *D*, as illustrated in Figure 1.1. This path is the minimum delay route from node *S* to node *D* in the sense that it is the first route found. Node *D* discards all Route Request packets corresponding to the same route discovery process after the arriving of the first one.

The route maintenance mechanism ensures that the paths stored in the route cache are valid. If the data link layer of a node detects a transmission error, the node creates a Route Error packet and transmits it to the original sender of the data packet. This Route Error packet indicates which link is "broken", *i.e.*, the node that detected the error and the node it was trying to reach. When a node receives a Route Error packet, it removes the link in error from its route cache and for each route containing this link, truncates the route from the hop before the broken link.

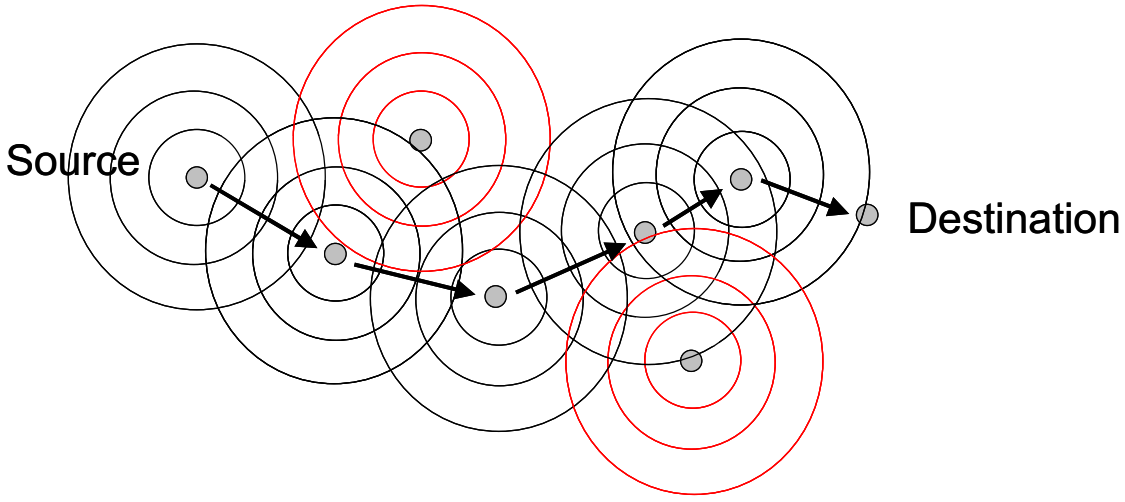

Figure 1.1 Route discovery in a wireless ad hoc network.

In order to have feedback on the status of each hop, several acknowledgement mechanisms may be used, *e.g.* acknowledgement at the MAC layer level, request of an explicit acknowledgement from the next-hop receiver in the data packet header, or passive acknowledgement (that is, a node overhears the next node forwarding its packet).

1.2 Background on IEEE 802.11

The IEEE 802 committee [6] has established the standards for the LAN industry for the past two decades, including 802.3 Ethernet, 802.5 Token Ring, and 802.3z 100BASE-T Fast Ethernet. In 1997, after seven years of work, the IEEE published 802.11, the first internationally sanctioned standard for wireless LANs [7]. In September 1999, they ratified the 802.11b "High Rate" amendment to the standard, which added two higher speeds (5.5 and 11 Mbps) to 802.11.

Like all IEEE 802 standards, the 802.11 standards focus on the bottom two levels of the ISO model, the physical layer and data link layer. Any LAN application, network operating system, or protocol, including TCP/IP and Novell NetWare, will run on an 802.11-compliant WLAN as easily as they run over Ethernet. The basic architecture, features, and services of 802.11b are defined by the original 802.11 standard. The 802.11b specification affects only the physical layer, adding higher data rates and more robust connectivity.

1.2.1 IEEE 802.11 Operating Modes

802.11 defines two kinds of terminal, a *wireless station*, which is usually a PC equipped with a wireless network interface card (NIC), and optionally an *access point* (AP), which acts as a bridge between the wireless and wired networks. An access point usually consists of a radio, a wired network interface (*e.g.*, 802.3), and bridging software conforming to the 802.1d bridging standard. The access point, when used, acts as the base station for the wireless network, aggregating access for multiple wireless stations onto the wired network.

The 802.11 standard defines two operating modes: *infrastructure mode* and *ad hoc mode*. In the infrastructure mode, the wireless network consists of at least one access point connected to the wired network infrastructure and a set of wireless end stations. This configuration is called a Basic Service Set (BSS). An Extended Service Set (ESS) is a set of two or more BSSs forming a single subnetwork. Since most corporate WLANs require access to the wired LAN for services (file servers, printers, Internet links) they will operate in infrastructure mode.

The ad hoc mode (also called an Independent Basic Service Set, or IBSS) is a set of 802.11 wireless stations that communicate directly with one another without using an access point or any connection to a wired network. This mode is useful for quickly and easily setting up a wireless network anywhere that a wireless infrastructure does not exist or is not required for services, such as a hotel room, convention center, or airport, or where access to the wired network is barred (such as for consultants at a client site). While designed for fully connected network (LAN) configurations, the ad hoc mode can be adapted to multihop use in MANETs.

1.2.2 The IEEE 802.11 MAC Layer

The 802.11 MAC layer is designed to support multiple users on a shared medium by having the sender sense the medium before accessing it at a random "backoff" time following a "distributed interframe space" (DIFS) during a period of contention for the channel, as illustrated in Figure 1.2. If there is a collision due to two users transmitting simultaneously, the users must re-enter contention. In 802.11, after a collision, the size of the contention window is increased, repeatedly if necessary, to make a collision-free transmission by one of the terminals more likely. Figure 1.3 shows the increase in contention window size (number of slots) as a function of the number of retries that are necessitated by collisions.

For 802.3 Ethernet LANs, the Carrier Sense Multiple Access with Collision Detection (CSMA/CD) protocol regulates how Ethernet stations establish access to the wire and how they detect and handle collisions that occur when two or more devices try to simultaneously communicate over the LAN. In an 802.11 WLAN operating in the ad hoc mode, direct detection of collisions is not possible since, to detect a collision, a station must be able to transmit and listen at the same time, which cannot be done in a radio system. However, an indirect detection of a collision is possible for a radio system using the scheme illustrated in Figure 1.4, in which an acknowledgement of correct reception is required to be returned from a receiver after waiting a "short interframe space" (SIFS) following a transmission. If an acknowledgement is not received, it is taken as an indication that a collision of two or more transmissions at the receiver has destroyed the information in both messages.

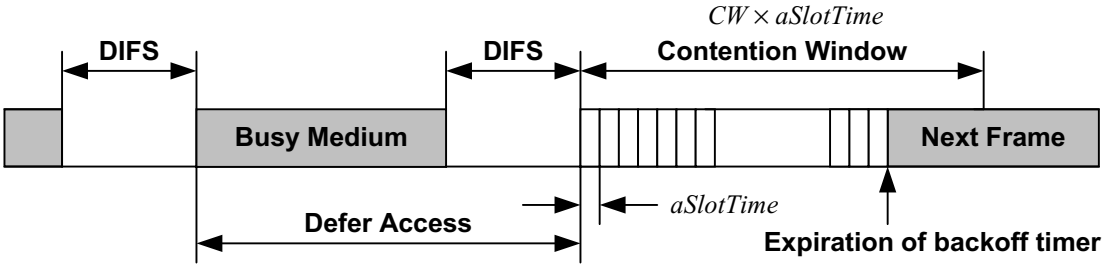

Figure 1.2. Carrier-Sense Multiple Access (CSMA)

Figure 1.3. Backoff timer adjustment

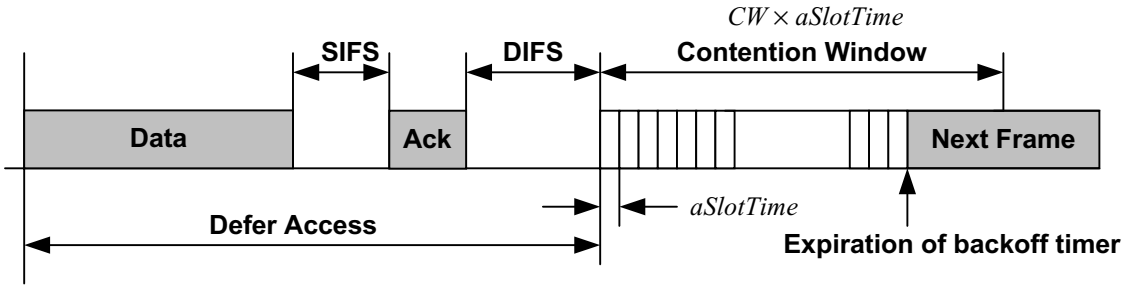

Figure 1.4. CSMA with collision detection

The "defer access" interval shown in Figures 1.2 and 1.4 is implemented in the ad hoc mode of 802.11 as "virtual carrier sense mechanism" [7] of a distributed coordination function (DCF) that provides information at the beginning of data messages concerning their duration; listening mobile terminals defer their contention for the next transmission opportunity until after the projected end of the current transmission (including the ACK message), plus a DIFS interval. In effect, the channel is "reserved" for the duration of the message by the transmitter that successfully competed for the transmission opportunity.

Although a wireless LAN generally is a network in which all the nodes can hear each other (*i.e.*, a fully connected network), in the ad hoc mode it is possible for nodes on opposite ends of the wireless LAN's coverage area to be unable to hear each another. (In a MANET application, this situation is assumed.) In that case, a "hidden" node on one side might not hear the beginning of the message of a node on the other side, in which the channel was reserved, and for that reason begin transmitting, cause a collision at receiving nodes in the middle of the network. To solve this problem, 802.11 specifies an optional "collision avoidance" feature that is implemented using a Request to Send/Clear to Send (RTS/CTS) protocol at the MAC layer. As illustrated in Figure 1.5, when this feature is in use, a sending node (*A*) transmits an RTS and waits for the receiving node (*B*) to reply with a CTS. All stations in the network that can hear either *A* or *B* now know to defer their access attempts, and a collision at *B* has been avoided.

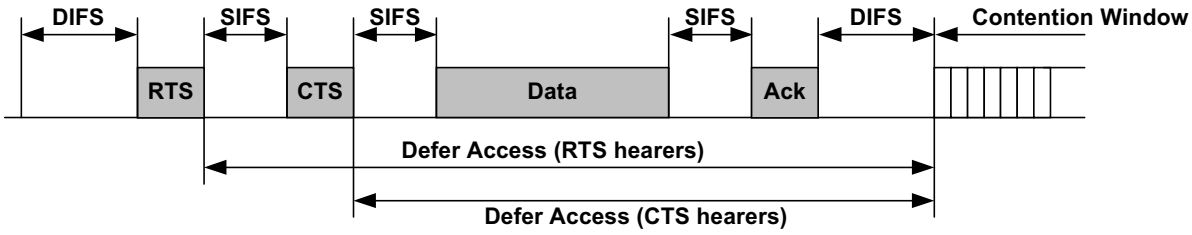

Figure 1.5. CSMA with collision avoidance

To save on overhead, there is an option in 802.11 to omit the RTS/CTS exchange when the data message length is shorter than some threshold, thereby trading off the risk of repeating a short message with the certainty of using up air time with more signaling traffic than data traffic.

1.2.3 Support for Time-Bounded Data in IEEE 802.11

Time-bounded data applications such as voice and video are supported in the 802.11 MAC specification through a point coordination function (PCF). As opposed to the DCF, where control is distributed to all stations, in the PCF mode a single access point controls access to the media. If a BSS is set up with PCF enabled, the available transmission opportunities are divided into portions in which the system is in the PCF mode and portions in which it is in the DCF mode. During the periods when the system is in PCF mode, the access point will poll each station for data, and after a given time move on to the next station. No station is allowed to transmit unless it is polled, and stations receive data from the access point only when they are polled. Under PCF, a mobile station with high priority or with a high priority message to or from it can be polled more frequently, thereby providing that station a higher proportion of the channel bandwidth.

While the nodes in a MANET can conceivably organize themselves into clusters and appoint certain nodes to act as access points to control and to prioritize transmissions within each cluster, in this report we are interested in considering modifications to the IEEE 802.11 MAC layer procedures to implement priority in the DCF mode.

2. Priority in MANETs

In many respects, MANETs have characteristics that are similar to those of the packet radio networks that were studied extensively in the 1970s and 1980s (e.g., [9, 10, 11]). Many analytical and experimental studies were conducted to determine the best methods for providing multiple access to packet users in terms of maximizing network throughput while maintaining acceptable delay, that is, maximizing the statistical multiplexing efficiency of the network. In recent years, there has been a renewed interest in packet networks that are configured as LANs or as mobile versions of the Internet, and instead of throughput *per se* the performance of the network in terms of quality of service (QoS) issues has been primary. This is not really a new development, but does reflect the fact that packet radio networks and MANETs are being called upon to carry different classes of traffic, some of which (such as digital voice) require timely

delivery of their packets, while other types of packets on the same network are more tolerant of delays. Since the limited bandwidth of the mobile radio channel makes it impossible to give every class of traffic the same quality of service except when the network is very lightly loaded, some means for providing each class a different quality of service must be implemented that involves assigning priority to one class over another in terms of allocating network resources. Thus the linkage between QoS and "priority" is a common one in the literature, and the two terms are almost synonymous.

However, there is also a distinct issue of priority *per se* in MANET and other wireless networks; recently, the National Communications System (NCS), through the regulatory authority of the Federal Communications Commission (FCC), established Priority Access Service (PAS) that gives certain national security and emergency preparedness (NS/EP) personnel special access to commercial mobile radio services [12]. Priority in this context is the immediate use of resources (non-blocking service) and the issue is not so much the bandwidth but the policy for managing this type of access.

In what follows, we conduct a brief review of the literature to gain an understanding of the different approaches that have been taken for implementing priority in packet radio networks and MANETs, then concentrate on the approaches that are suitable for the distributed network management that is typical of MANETs. Against this background, we then introduce the MAC layer priority schemes that are the subject of this report.

2.1 Approaches to Implementing Priority in Wireless Networks

For a static situation, in which the traffic flows originating at the network nodes are known and are fixed, it is possible using various algorithms to calculate a "schedule" of transmissions that maximizes network throughput while providing the desired QoS to the various traffic sources. As long as the total traffic does not exceed the capacity of the channel(s) allocated for the transmissions, the schedule guarantees performance in a non-contention access mode. The extension of these procedures to a mobile network, whose topology in general is changing with time, requires periodic or adaptive recalculation of the schedules.

2.1.1 Transmission Scheduling

In [13] a transmission schedule generated by exploiting the properties of Galois fields is discussed. In [14], the authors describe a centralized algorithm for assigning transmission slots in a multihop wireless network in a non-contention access scheme such that traffic requirements (including traffic forwarded by a node) are met. The algorithm is suggested as a means of comparing the performance of distributed scheduling algorithms and several examples are given. Metrics are given for rating the fairness of such algorithms. In [15], two algorithms are given for scheduling unicast conversations in a spread spectrum radio network, making use of the code division multiple access properties of spread spectrum communications. The algorithms seek to find sets of links that can be simultaneously active while satisfying end to end traffic demands in a multihop network.

In [16], Ephremides and Truong describe the general problem of scheduling transmissions in a multihop radio network under the constraint of no conflicting transmissions from nodes up to two hops away from a node that is scheduled to transmit in a given slot, and

assuming the same amount of transmit traffic at each node. The general centralized problem is shown to be NP complete, while the problem can be solved in polynomial time if the modulation scheme will tolerate interference from simultaneous transmissions. A distributed version that does not allow conflicts is shown, requiring two-hop connectivity information at each node for developing a "skeleton" transmission schedule of only one node per timeslot to be proposed to surrounding nodes, followed by a distributed assignment of non-interfering simultaneous transmissions that is resolved in part by postulating a certain order of priority for the different nodes' transmissions.

As alluded to in the previous paragraph in connection with [16], the generation of transmission schedules is computationally intensive and is usually conceived of as being performed at a central node in a network. Along with the assumption of a central node, it is often assumed that there is a separate channel available for signaling, whose capacity is not subtracted from the capacity allocated to the traffic. In [17] the authors describe an "inhibit sense" multiple access (ISMA) scheme that assumes a central station and duplex links between it and all mobiles. The downlink is used to indicate reception of a packet by one mobile and so to inhibit colliding transmissions by other mobiles. A new version of ISMA is described for handling traffic with different priorities: the transmitted idle signal changes in time while the channel remains idle, so that lower priority mobiles are inhibited longer. The paper includes analysis of arrival rate, idle duration, and throughput. Numerical results indicate an improvement of high priority throughput without decreasing overall throughput.

The system proposed in [18], described as the Limited Sensing Random Access Algorithm (LSRAA), pertains to clusters of radios that are connected via a backbone network of designated radios that are in the centers of the respective clusters, similar to a cellular system. Each cluster has its own channel, and non-priority users must use the channels assigned to their respective clusters. Designated "marginal users" can access any channel, and the algorithm is a way to select the channel on which to transmit, based on feedback reflecting the traffic in the various channels, specifically whether a collision has occurred in a channel during the current timeslot. When a collision occurs in a channel, transmissions are controlled until the collision is resolved by allowing the users who collided (and no others) to retransmit during a collision resolution interval (CRI), or window. However, the marginal users can participate during this process. An analysis is presented for a two-channel system that shows low delays for marginal users even when local users are beginning to be blocked.

In [19], it is assumed that an out-of-band signaling channel is available for requesting or reserving slots. Continuation priority (for ongoing, segmented data) is higher and requests can piggyback the data packets. In [20], priority is defined in terms of waiting time and the system studied assumes the availability of a slotted ALOHA out-of-band signaling reservation access (RA) channel for mobile users to request from a base station the assignment of packet transmit channel slots, under a priority based multiple access scheme. Under a modified scheme with throughput control, when a collision is detected on the RA channel, the base polls the requesting users on the RA in order of priority in order to reschedule the transmission.

The authors in [21] report on experiments to determine latency of digital voice over an 802.11 network using the point coordination function (PCF) to prioritize the voice signal in the presence of other signals.

2.1.2 Distributed Operation Using Mechanisms That Favor Priority Access

As we have mentioned in Section 1.1, *ad hoc* packet radio networks and MANETs are assumed to be decentralized, with distributed network management functions. These assumptions do not necessarily imply that the channel access is based on contention, nor that all signaling is done on the common channel, but typically algorithms developed for MANETs are usually based on these latter assumptions as well. Therefore, whatever schemes are proposed for scheduling transmissions and giving preference to priority users, our interest is primarily in those schemes that are distributed and do not require a central station or a separate signaling channel.

As alluded to previously in connection with [16], the generation of transmission schedules is computationally intensive but can be cast in the form of a distributed algorithm. However, most distributed schemes for implementing priority in packet radio networks utilize various "mechanisms" that are designed to give an advantage of some kind to priority users, and thereby deliver to them a better grade of service.

In [22], A "part and try" multi-priority random multiple access (RMA) algorithm is analyzed for a slotted packet radio channel. In the priorityless version of the algorithm, when there is a collision in a timeslot, a conflict resolution "session" begins in which an attempt is made to discover the colliding user that transmitted first and to permit that user to retransmit without contention; then normal contention access resumes. The version of the algorithm with priority treats collision resolution in the same way, but allows the user with higher priority to transmit first rather than the one which transmitted first. An analysis is given for the performance of this scheme when there are two levels of priority.

The authors in [23] show an analytical method for bounding the delays achieved by window RMA algorithms, which permit distributed control of contention access to a common channel for improved throughput and delay. The authors note that the distribution of delay has "significant mass" in the tails of the distribution, so that analysis or design based on mean values can be misleading.

Two scheduling algorithms are analyzed in [24] for a General Packet Radio Service (GPRS) application in a cellular system: static priority scheduling (SPS) and modified earliest deadline (MED). Under SPS, each priority class of traffic has its own first-in first-out (FIFO) queue, and the queues are served in order of delay priority, first in first out, as suggested in Figure 2.1. Under MED, each block in a message to be scheduled is assigned a transmission deadline, and the blocks are placed in a queue in deadline-priority order; when congestion occurs, as evidenced by deadlines being missed, blocks are placed in prioritized "late queues" for service in order of importance. Simulation results are presented in terms of normalized packet delay vs. number of mobile users contending for the channel, and statistical multiplexing efficiency (in terms of network utilization for unit normalized packet delay). Two congestion metrics are defined: queue lengths in an SPS scheduler (directly proportional to packet delay) and frame scheduling advance for MED (inversely proportional to packet delay).

Also in connection with GPRS, [25] contains a rather comprehensive list of scheduling methods. The paper considers FIFO, SPS, and EDF (earliest deadline first) for the case of three delay-service classes. Numerical studies included different traffic models for email, fleet

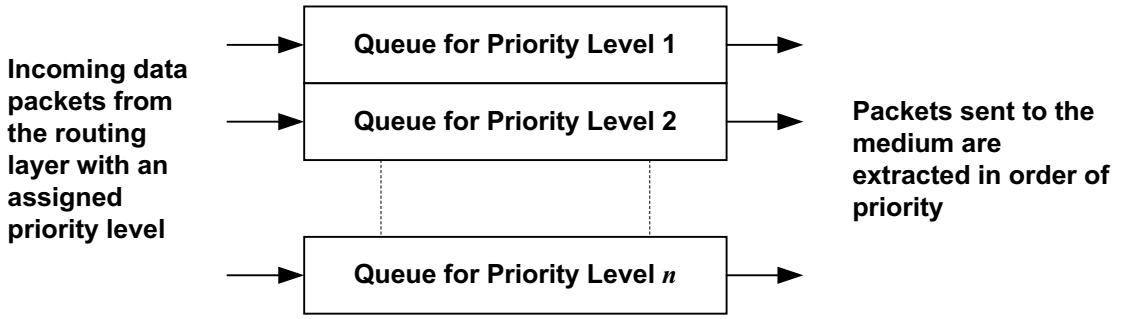

Figure 2.1. Priority queuing approach

management, and Internet browsing. The simulations showed that, of the techniques considered, only EDF meets the delay requirements when the link utilization is 90% in the presence of the assumed bursty traffic and packet length variations.

Similarly, in [26] three access priority schemes are proposed: random chip delay (RCD), random backoff based (RBB), and variable logical channel (VLC) based access priority schemes. The system in [27] gives higher priority to acknowledgement packets in a receiver-directed system in order to reduce blocking.

2.2 Mechanisms for Implementing Priority at the 802.11 MAC layer

Taking the approach that the MAC layer procedures of IEEE 802.11 can be modified to create mechanisms that implement prioritization of traffic in a MANET, we have developed new procedures and corresponding mechanisms, and have undertaken a study of the performance of MANETs with and without such mechanisms.

2.2.1 Node-Level Priority Queuing

The first MAC layer mechanism is to include a priority feature at the node level, in which each node tries to send the important packets before the others. This mechanism is in fact the SPS approach that was discussed previously and illustrated in Figure 2.1. It is easily implemented in 802.11 by replacing the single queue used to store the packets waiting to be transmitted with several different sub-queues representing different levels of priority. Packets awaiting transmission are extracted for transmission first from the highest priority level queue, then from the second one, and so on until the lowest priority queue is reached. For instance, the oldest packet located in the queue with the priority level 3, must wait until the level 1 and level 2 queues are empty, *i.e.* that every more important packet has been transmitted, before being extracted and emitted.

For the purpose of this study, we chose to implement only two priority levels, on the basis that two levels of priority are enough to indicate the effectiveness of the concept of priority queuing. Therefore we replace the single 802.11 packet buffer with two queues: one containing the "standard" packets (⇔ queue priority level 2) and the other the packets with "priority" (⇔ queue priority level 1), with the priority queue always being served first if it is not empty.

2.2.2 Prioritized Waiting Time Mechanism

The second mechanism for implementing priority at the MAC level is to permit a node wanting to send the highest priority packet to be the first to transmit on the channel. As noted above, in the 802.11 specification the "waiting" mechanism is divided in two cases, depending on whether the medium is free or not.

The first case provides that if the medium is free for a specified time (DIFS) then the transmitter is allowed to contend for an opportunity to transmit its packet. Thus, as illustrated in Figure 2.2, if we want to implement some priority mechanism at this level it will be necessary to define different DIFS times (each one separated by at least one small timeslot to follow the 802.11 specification), the shortest one being attributed to the packet with the highest priority and the longest to the packet with lowest priority.

The waiting time "traditionally" used in 802.11 is the shortest one possible (DIFS). To implement the prioritized waiting time mechanism in our simulation models, we assign this shortest waiting time to packets with "priority" and for the "standard" packet we add one more timeslot to it.

The second case provides that, when the medium is busy when a node wishes to transmit its packet, the node must first defer and wait for completion of the current transmission under the virtual sense mechanism discussed previously. At the end of a current transmission, a node that is waiting to transmit must then continue to wait for the DIFS time plus a random backoff time, defined as a number of small timeslots according to the 802.11 specification, before accessing the medium. In the next subsection, we propose a modification to the backoff time required following the deferral period.

2.2.3 Prioritized Backoff Time Distribution

It is possible to adapt the idea of different DIFS times for different priorities proposed for the free medium case, and to implement a priority mechanism by calculating the random backoff time depending on the priority level of the packet that the node wants to transmit. For example, it is possible to use different distributions (depending on the packet priority) to choose this random backoff time, as illustrated in Figure 2.3. The node having the highest priority packet to transmit will be most likely to access to the medium first.

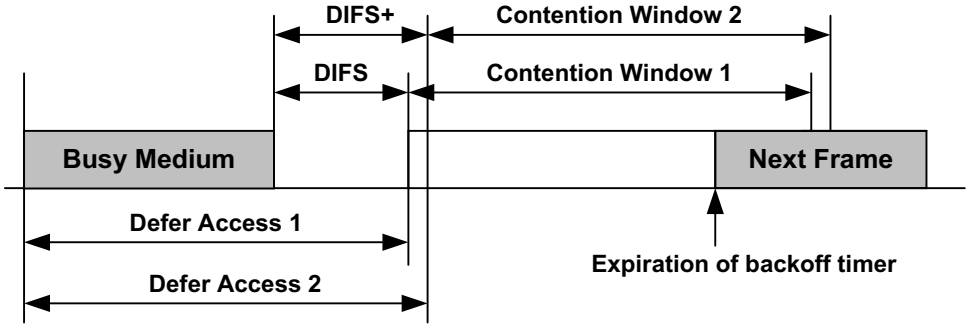

Figure 2.2. Prioritized waiting time mechanism

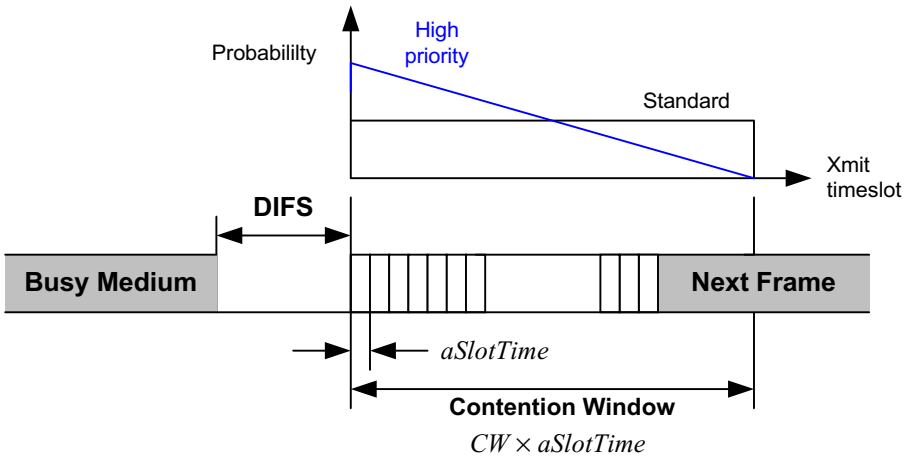

Figure 2.3. Prioritized backoff time distribution mechanism.

In Figure 2.3, the standard backoff time calculation procedure is represented by the "Standard" or uniform probability distribution in the graph in the upper part of the figure. The concept behind this distribution is that each node contending for the channel randomly selects a time to transmit from a set of equally likely, discrete values, with the spacing between the discrete values providing sufficient time to allow the carrier-sense mechanisms in later-scheduled nodes to keep them from transmitting on top of a node that had randomly selected an earlier time slot for transmission. Also in Figure 2.3, a prioritized backoff time calculation procedure is represented by the "High priority" probability distribution in the graph in the upper part of the figure. The concept behind this distribution is that the high-priority user has the same set of potential discrete-valued transmission times, but the times are not equally likely; instead the earlier times are more likely, making the high-priority user more likely to "win" in contending for the channel during the backoff period.

This concept of a prioritized backoff time distribution can be implemented in combination with the concept of prioritized waiting times, as suggested in Figure 2.4. The distribution of the backoff time for a standard packet is not only uniform, it is delayed relative to the contention window for a prioritized packet, which has a backoff time distribution that makes selecting an early transmission time more likely.

Our implementation of the prioritized backoff time distribution concept is based on the OPNET exponential distribution model given by the exponential probability density function (pdf)

$$p_t(x) = \lambda e^{-\lambda x}, \qquad x \geq 0$$

Recall that the standard 802.11 interval for choosing the backoff number of slots is variable from a small interval [0, *contention_window_min_size*] to a large interval [0, *contention_window_max_size*], depending on the number of medium access attempts as illustrated previously in Figure 1.3. For the priority backoff distribution, we decided to implement a similar mechanism by using a variable mean $1/\lambda$, where

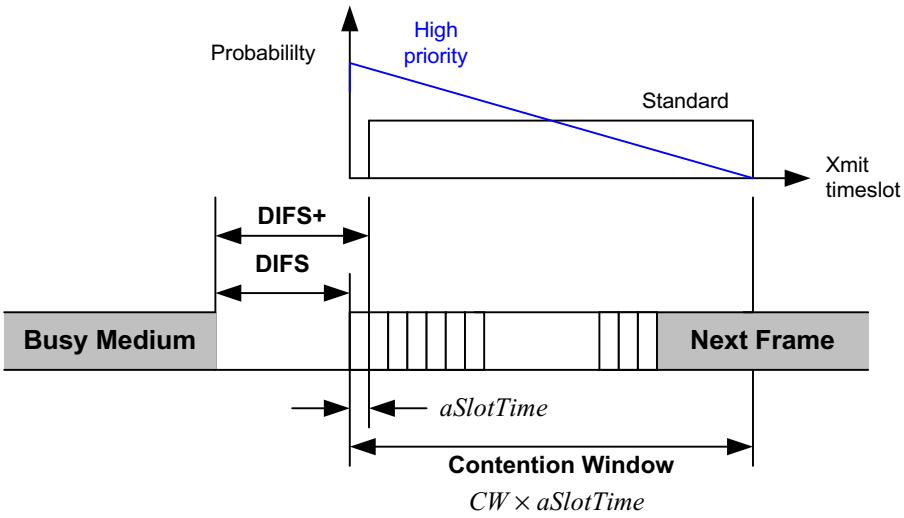

Figure 2.4. Combined priority mechanisms

$$\lambda = 0.1 + \left(\frac{CW\_\max\_size - CW\_current\_size}{CW\_\max\_size - CW\_\min\_size} \right) \times 0.3$$

and "*CW*" denotes Contention Window. Example distributions are shown in Figure 2.5. When the contention access has been collision-free and the window size is minimal, the distribution with $\lambda = 0.4$ is used, and when collisions have forced the window size to its maximum value, the distribution with $\lambda = 0.1$ is used.

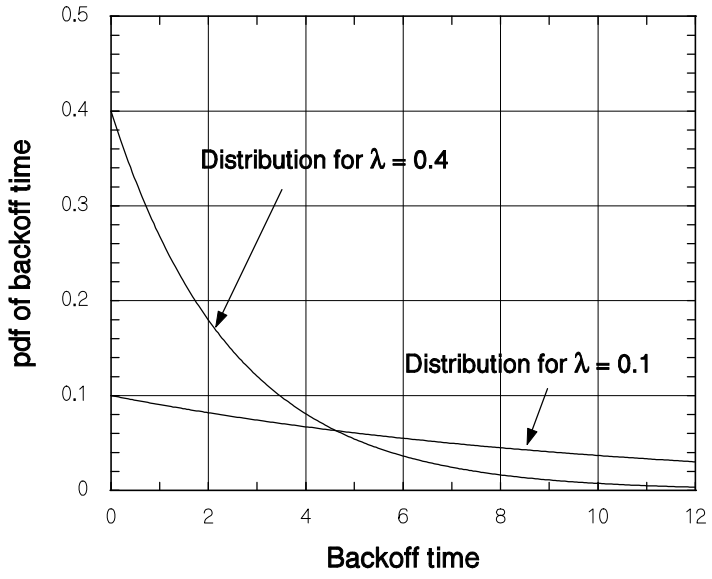

Figure 2.5. Backoff time probability distributions for high priority users.

3. Simulation of the Proposed Priority Schemes

In order to evaluate the potential advantages and disadvantages of the IEEE 802.11 MAC layer priority schemes introduced conceptually in the previous section, we have begun a series of simulations using the OPNET simulation software package. In this report, we present an overview of this ongoing effort and some preliminary results.

3.1 Description of Simulation Model

The OPNET simulation software package is structured so that each network is modeled as a configuration of *nodes* that are connected in user-specified ways, and in turn each node in the network is modeled as an interconnection of specific *modules* representing the various processes that take place in the actual communications equipment. In the case of a wireless network, the interconnections among the nodes are automatically determined during the progress of the simulation as a function of user-supplied propagation parameters, such as the effective transmission range of a node.

3.1.1 Node Model

An OPNET node model for the wireless network node incorporating the proposed priority mechanisms is diagrammed in Figure 3.1. As shown in that figure, the upper layers of the node model include the simulation modules for generating source traffic (*src*), receiving destination traffic (*receiver_sink*), and controlling the movements of the mobile node (*mobil*).

The network layer contains modules for implementing the DSR multihop routing protocol, and includes a module to interface the routing protocol with the traffic source/ destination and a module to perform the routing. This OPNET model was developed at NIST and is available to the public for downloading. It is described in detail in [6].

The link and physical layers contain simulation modules for the operation of the IEEE 802.11 MAC and physical layers. These modules are part of the OPNET simulation package version 7.0, but were modified by NIST for multihop MANET simulations as described in [6]. The feedback from the receiver and transmitter to the MAC layer includes the means for performing the CSMA scheme with collision detection and collision avoidance, as described above.

3.1.2 MAC Layer Process Model

A diagram of the OPNET process model for the 802.11 MAC layer is given in Figure 3.2. The features of this model are described as follows:

- A packet can be received from the upper layer in any state of this state machine. Actually, when such an event happens, the *wlan_higher_layer_data_arrival*(…) function is called via the general interrupt function *wlan_interrupts_process*(…). The packet is then stored in a queue through the *wlan_hlpk_enqueue*(…) function. In this latter function, a new queue is created in order to assign separate queues for the standard packets and for the priority ones, according to the first priority mechanism. If a packet arrives and is stored in the priority queue when the node is in the standard packet

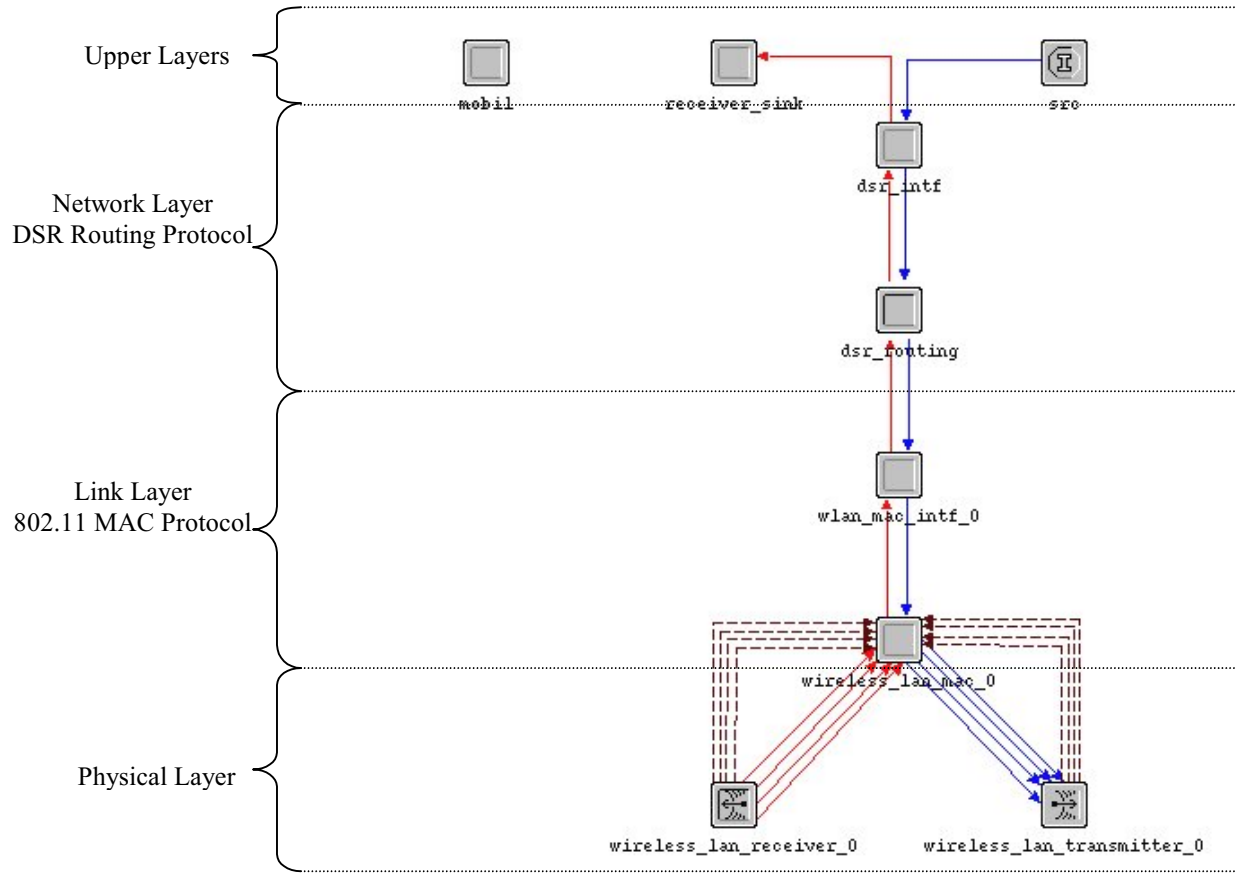

Figure 3.1. OPNET node model for mobile network node.

transmission mode, a flag is set to 1 in order to indicate that the node has switched to the "priority" packet transmission mode. At the same time some variables are also changed, like the DIFS duration, in order to reflect the new transmission mode.

- In the *Transmit* state, the *wlan_frame_transmit*(…) function is called in order to transmit or retransmit a packet on the medium. If it is the first transmission of a data frame, the packet is extracted from the priority queue, or if it is empty from the standard queue. In that way, the priority packets are always transmitted before the standard ones.
- In the *Bkoff_Needed* state, a new calculation method is implemented for the priority packets using the exponential distribution according to the second proposed priority mechanism. If the node is in the priority packet transmission mode, this new method is used to determine the number of backoff slots; otherwise the standard 802.11 technique with a uniform distribution is applied.
- In the *Backoff* state, if a data packet is received from the upper layer and if the priority queue flag indicates that the transmission mode has just switched to "priority," then the current standard packet transmission or retransmission is canceled. In this case, our new *wlan_retransmission_cancel* (…) procedure is activated and replaces the eventual data packet waiting for retransmission in the standard queue, and resets every variable and flag in order to cancel the retransmission process. Moreover a new number of backoff

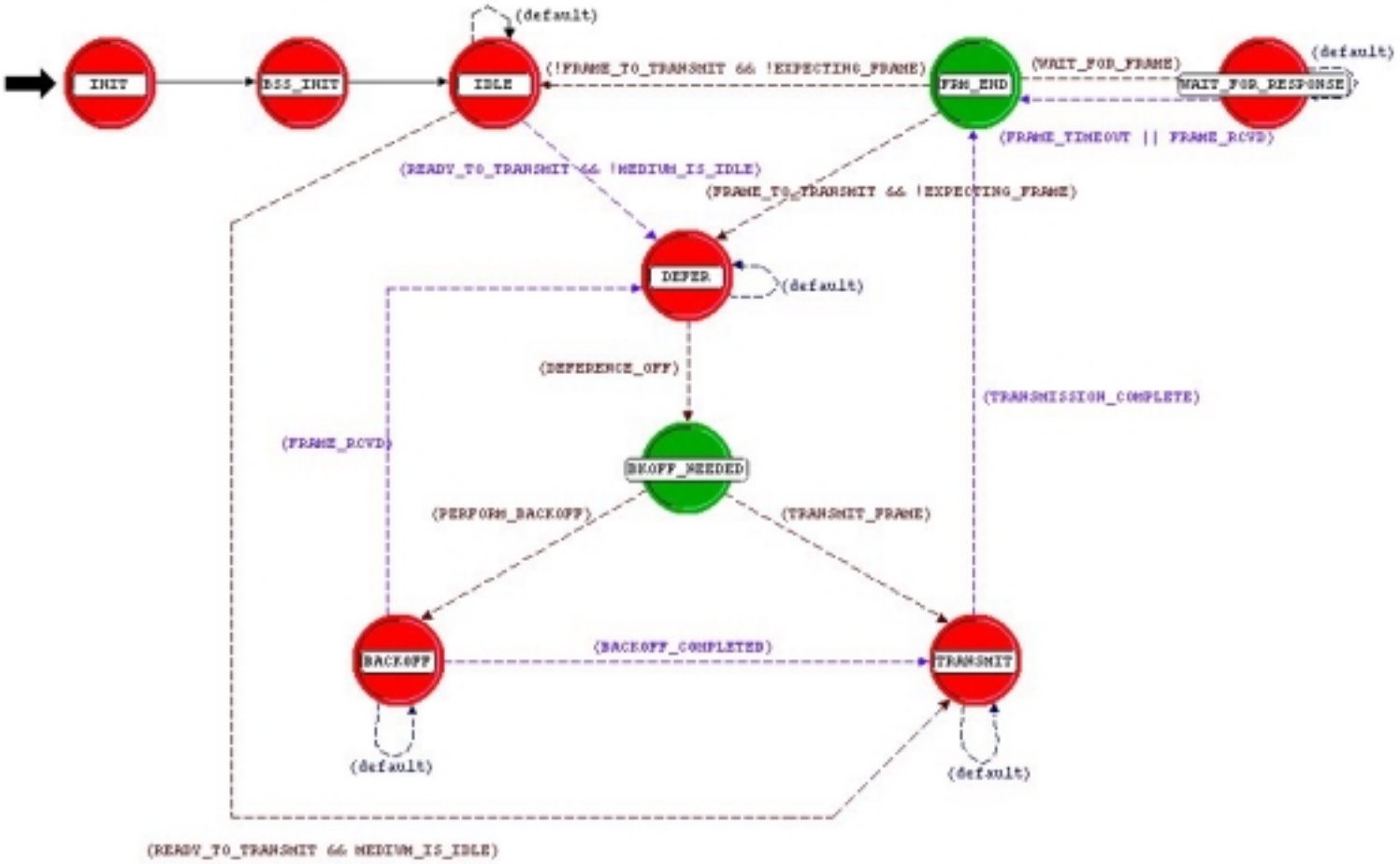

Figure 3.2. OPNET process model for IEEE 802.11 MAC layer

slots is calculated via the technique explained in the second mechanism. However, since the node has already waited for a certain amount of time for its retransmission, the new backoff timer is adjusted considering this time, which can result in an instantaneous emission.

- According to the third mechanism, in the *Transmit* and the *Wait_For_Response* states, if a data packet is received from the upper layer and if the flag indicates that the transmission mode has just switched to "priority," another flag indicating the interdiction of any future retransmission of the current packet is set to 1. In this way we give a last chance to the current node to receive the ACK or the CTS that it is waiting for. But if it does not receive this reply the model calls the *wlan_frame_discard*(…) function in order to plan a new retransmission. Thus we implement in this function a small test that looks at the value of the retransmission flag and that calls our function *wlan_retransmission_cancel*(…) if the flag is set to 1.
- A packet can be received from the physical layer in any state of the state machine. Actually, when such an event happens the general interrupt function *wlan_ interrupts_process*(…) calls the *wlan_physical_layer_data_arrival*(…) function. We only modified a few things in the function at the ACK reception in order to reset the mode of packet transmission to "standard" if there is no priority data packet waiting.

3.2 Description of Simulation Scenario and Parameters

We have concentrated our attention so far on a scenario that highlights the salient points showing how the proposed priority mechanisms work. The scenario studied is designed to analyze throughput performance with and without the priority mechanisms. As the study continues, other scenarios will be generated for the purpose of refining the priority mechanisms in the context of typical MANET operations.

3.2.1 Throughput Analysis Scenario

The throughput is the most difficult parameter to analyze, and a scenario was designed to examine the throughput performance of the proposed priority mechanisms. The network topology used in this scenario is depicted in Figure 3.3. The goal of this scenario is to increase the level of traffic until a saturation point is reached on a node in order to observe the effects of the priority mechanisms on it. In this topology, every packet generated by one of the four outer nodes must pass through the central node (node number 5) in order to reach its destination. Therefore, node 5 should saturate above a certain amount of traffic. After fixing equal percentages of standard and priority packets, we can find the saturation point by simply increasing progressively the amount of traffic. Thus we should see that the total throughput for the two priority levels, which must be equal at the beginning, should be more and more attributed to the priority packet as node 5 approaches to its saturation limit.

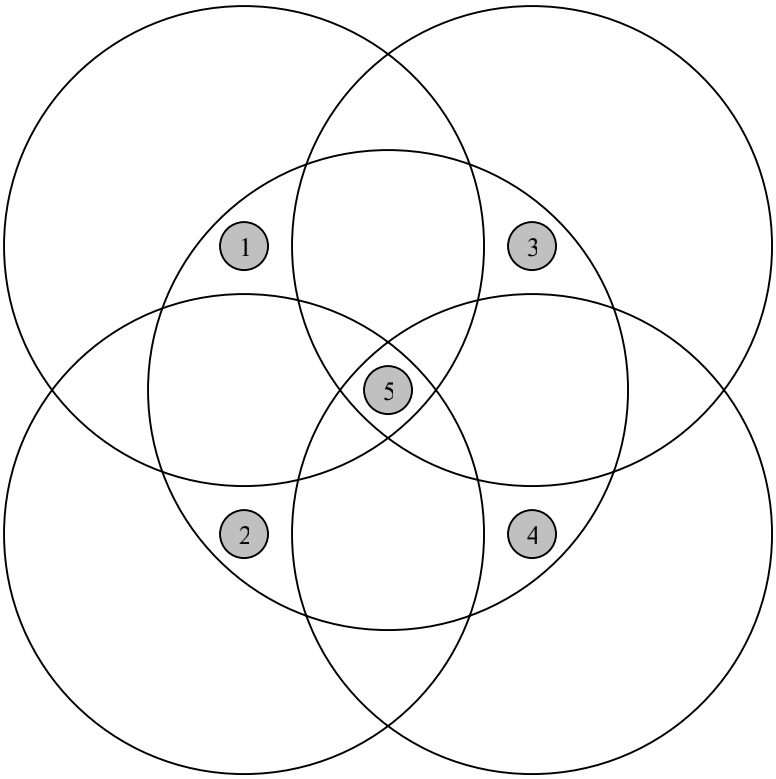

Figure 3.3. Scenario for throughput analysis

3.2.2 Parameters of Throughput Analysis Scenario Simulations

As indicated in Figure 3.4, which is a screen shot of the OPNET network model for this scenario, the five nodes indicated in the conceptual diagram of Figure 3.3 were located on a grid such that the four outer nodes are at the corners of a 300 m × 300 m area. The center node therefore is located $0.707 \times 300\,\text{m} = 212\,\text{m}$ away from each of the outer nodes, while the outer nodes are at least 300 m away from each other. With the range of the WLAN radios in the simulation specified as 250 m, the intent of the scenario in Figure 3.3 is implemented by the arrangement shown in Figure 3.4.

Other parameters of the simulations involving the scenario of Figure 3.4 are summarized as follows:

- *Physical Layer:* The frequency hopping mode of IEEE 802.11 was arbitrarily selected, resulting in the parameter values given by

$$T_{slot} = 50\,\mu\text{s}$$

$$SIFS = 28\,\mu\text{s}$$

  Also, instead of utilizing the higher data rates available, the simulation assumed that the

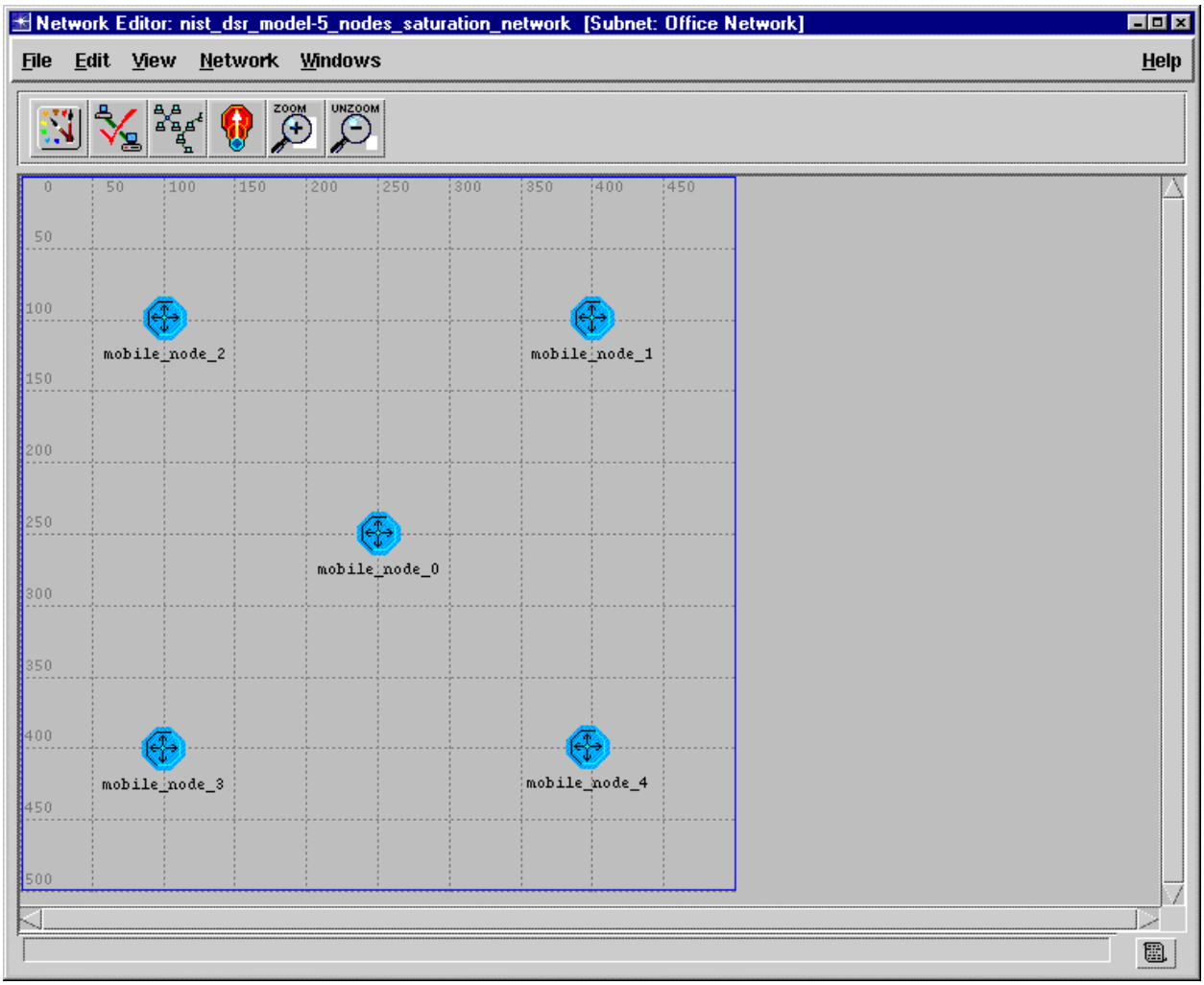

Figure 3.4. OPNET model for scenario depicted in Figure 3.3.

transmission rate on the medium is 1 Mbit/sec, resulting in a transmission delay of 1 μs per channel bit. Each data bit, including the contents of the signaling packets is transmitted on the channel without further modification, such as error-control coding. Therefore, the delay incurred by sending, for example, a 112-bit ACK message is

$$T_{ACK} = 112\,\mu\text{s}$$

- *MAC Layer:* The RTS threshold was selected as 256 bytes; if a packet offered to the MAC layer is shorter than 256 bytes (2048 bits), the RTS/CTS collision avoidance mechanism is not employed. To implement the priority access mechanisms discussed above,

$$\text{DIFS} = \begin{cases} \text{SIFS} + 2T_{slot} = 128\,\mu\text{s}, & \text{priority packet} \\ \text{SIFS} + 3T_{slot} = 178\,\mu\text{s}, & \text{standard packet} \end{cases}$$

In order to implement the proposed priority mechanisms, the MAC layer was modified to include two separate queues for packets awaiting transmission on the channel, one for

standard packets, another for priority packets. Also the backoff distribution for priority packets was made to be non-uniform, as discussed above in Section 2.2.3.

- *Upper Layers:* The packet generators indicated in the node model of Figure 3.1 were chosen to have 512-bit packets, prior to encapsulation at the DSR routing level and at the MAC level. The packet rate, in terms of mean interarrival time, was fixed for each simulation run. A random selection of priority status was made for each packet, with standard and priority packets being equally likely. The encapsulation of this data in a DSR data packet is illustrated in Figure 3.5, in which the shading represents modifications to the standard DSR packet for the purpose of this study (see also [28]). The total size of the DSR data packet equals the original data plus 184 bits. After further encapsulation at the MAC layer, the data packet gains another 224 bits, for a total of 920 bits per data packet for 512-bit source data packets, giving the data packet transmission time of

$$T_{data} = 920\,\mu\text{s}$$

Again, for implementing the priority mechanisms, the model for the node operations at the upper layers includes two queues for data awaiting routing, one for standard packets and another for priority packets. A limit on the size of these queues was not imposed.

The different routing packets (route request, reply, etc.) vary in size, with the average transmission delay for a routing packet being about 390 μs.

Based on these parameter values, an estimate of the capacity of the network in the scenario of Figure 3.3 can be obtained. Note that, for equal volumes of traffic generated at each of the five nodes, the amount of source traffic equals $5R$, using R for the source rate at one node. In addition, there is relay traffic: for randomly selected destination nodes, three out of four packets from each of the four outer nodes is destined for another outer node and has to be relayed (repeated) by the center node. Thus the relaying adds another $3R$ of traffic for a total of $8R$ that must be accommodated by the same medium, since, for the assumed network topology, successful transmission of data requires that only one node transmit at a given time.

Neglecting backoff time, an estimate of the time resource needed for each data packet is found by assuming not RTS/CTS exchange and expressing the total channel access time required by the equation

$$T_{total} = T_{data} + \text{SIFS} + T_{ACK} + \text{DIFS} = \begin{cases} 1.188\,\text{ms}, & \text{priority} \\ 1.238\,\text{ms}, & \text{standard} \end{cases}$$

SRC (8 bits)	DEST (8 bits)	RELAY (8 bits)	Seg_left (8 bits)	Size_Route (8 bits)	Type (8 bits)		
Node_0 (8 bits)	Node_1 (8 bits)	Node_2 (8 bits)	Node_3 (8 bits)	Node_4 (8 bits)	Node_5 (8 bits)	Node_6 (8 bits)	Node_7 (8 bits)
Priority (8 bits)	Data (inherited)						
TR_Source (16 bits)	Packet_ID (32 bits)			DSR Sent Time (16 bits)			

Figure 3.5. Format for DSR data packet.

From this equation, the maximum total data packet rate, again neglecting backoff time, is found as one-eighth the inverse of T_{total}, or about 105 packets/sec for priority traffic and about 101 packets/sec for standard traffic. Based on this estimate, we expect to observe the following:

- For source data rates significantly less than 100 packets/sec, the channel is not congested except for occasional bursts of traffic that are randomly generated by the simulator's Poisson traffic generators.
- For source data rates at 100 packets/sec and higher, the channel cannot service the data requests fast enough and the DSR routing queues will grow in size as backlogged traffic accumulates. With full routing queues, the behavior of the traffic being output by the MAC layer will reflect the operation of the MAC layer rather than the statistics of the traffic generators at the upper layers.

3.2.2 Performance Statistics Collected by Simulation

As the OPNET simulations execute, at the user's option various local or global statistics may be measured, calculated, and written to a file for post-simulation analysis. The statistics thus gathered during our simulations include the following global statistics:

- MAC layer statistics
  - *Average number of hops in a route*: the average internodal hop distance for the scenario in Figure 3.3 is 1.6 hops, and if the traffic between all node pairs is equal the average number of hops for a successful route will also be 1.6 hops.
  - *Average MAC delays for standard packets and for priority packets:* the difference between the time the packet is received from the DSR layer and the time that it is transmitted, averaged over the number of packets of each type.
  - *Average MAC backoff slots for standard packets and for priority packets:* the number of backoff slots produced by the random number generator at any node (different generators for standard and priority packets), divided by the total number of times the backoff calculation is performed for that type of packet.
- Averages of various quantities at the DSR routing layer:
  - *Average time in buffer for standard packets and for priority packets*
  - *Average hop delay for standard packets and for priority packets:* combination of routing layer buffering delay and MAC layer access/servicing delay.
- Running totals of various quantities at the DSR routing layer: data served, data transmitted, data sent directly, errors detected.
- Running totals of various quantities at the upper layer: total data transmitted, total data sent directly, total errors detected, total overhead

In what follows, we focus primarily on the MAC layer statistics, which are sufficient to indicate the behavior of the priority mechanisms that we are studying.

3.3 Simulation Results

3.3.1 Average Number of Hops

A graph of simulation results for the average number of hops on a per route basis is given in Figure 3.6 for node packet average interarrival times of 1.0, 0.1, 0.05, and 0.01 seconds (source rates: 1, 10, 20, and 100 packets per second). We observe in the figure that the average does indeed settle down to the expected value of 1.6 hops for the data rates shown that are less than 100 packets/sec, but for 100 packets/sec generated at each node, the average number of hops is smaller, around 1.35 or 1.4 hops. These results are consistent with our estimate of the capacity of the network and indicate that traffic originating from the center node is more successful at the high data rate, when there is congestion.

3.3.2 Average MAC Packet Delay

Example results for average MAC packet delay as a function of time are shown in Figures 3.7 and 3.8, for standard and for priority packets, respectively, and for node packet average interarrival times of 1.0, 0.1, 0.05, and 0.01 seconds (source rates: 1, 10, 20, and 100 packets per second). In each of these figures, we see that the delay increases with data rate, but only slightly for data rates up to 20 packets/sec at each node. For 100 packets/sec, the delay increases with

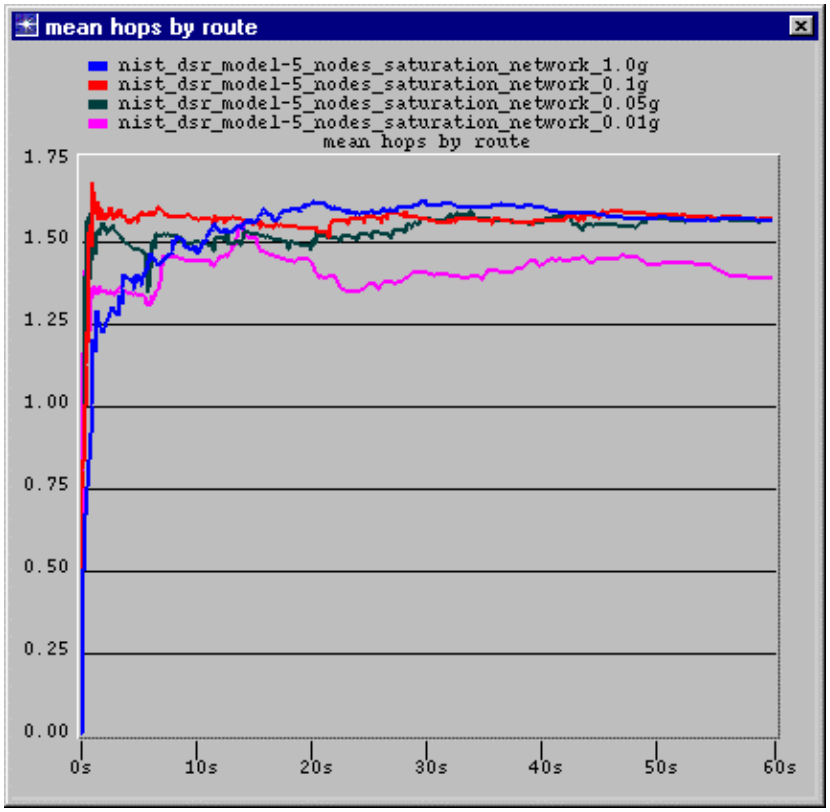

Figure 3.6. Average number of hops per route for source data rates of 1, 10, 20, and 100 packets/sec.

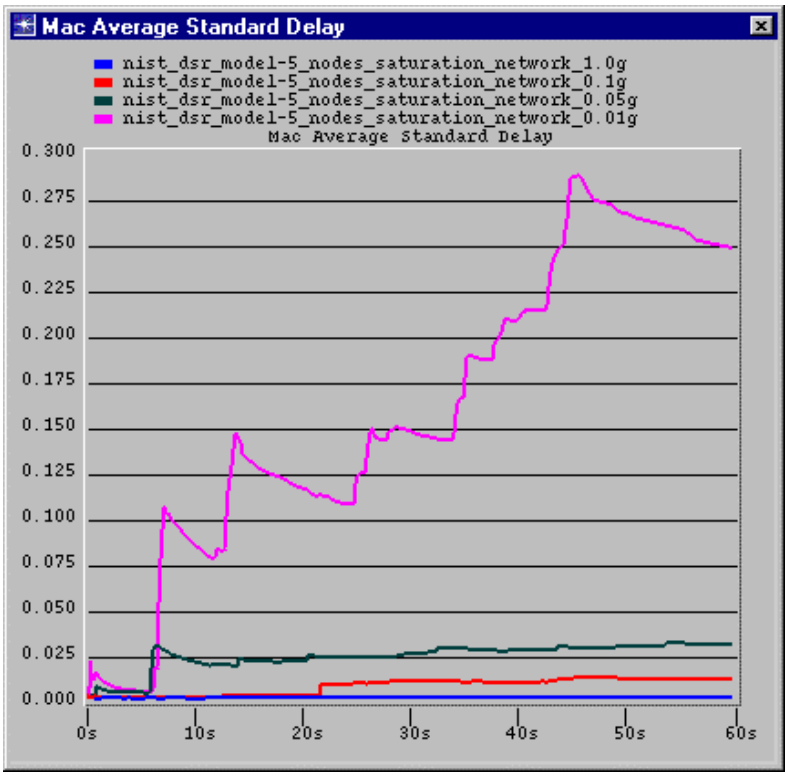

Figure 3.7. Average MAC packet delays for standard traffic.

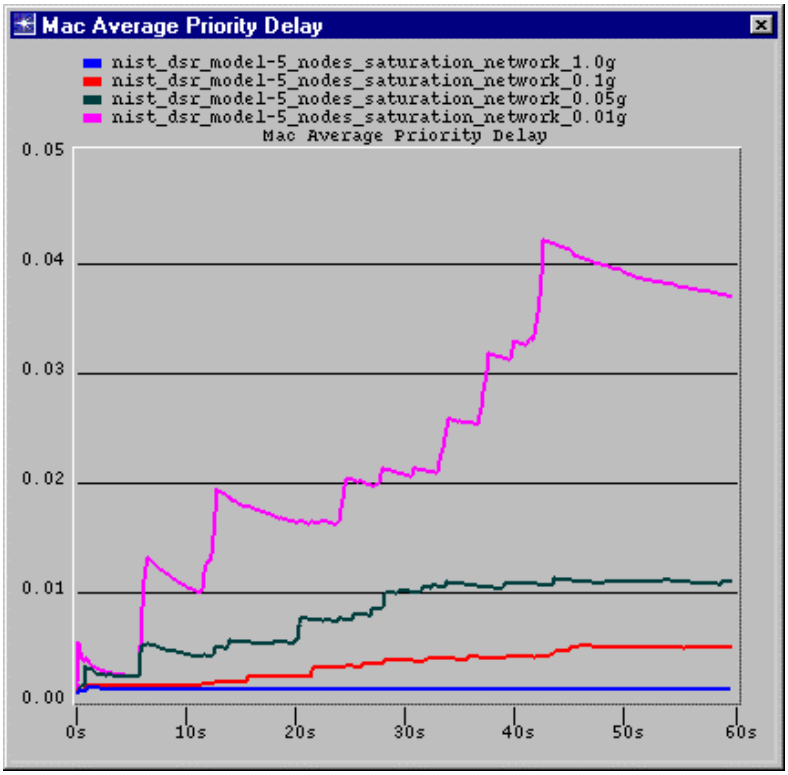

Figure 3.8. Average MAC packet delays for priority traffic.

time for both standard and priority packets in a kind of "staircase" fashion, indicating a backlog of messages to be sent that is periodically serviced during a lull in traffic but not sufficiently to keep up with the average amount of incoming traffic. Note, however, that the delays are much smaller for priority packets than for standard packets, which indicates that the priority mechanism is working in that it is definitely favoring the priority traffic. For example, when the packet rate is 20 packets/sec at each node, as shown in Figure 3.9, at the end of the simulation the average MAC layer delay for priority packets is about 0.011 seconds and stable, while for the same rate it is about 0.03 seconds and rising for standard packets.

3.3.3 Average MAC Backoff Slots

The small MAC delays seen in Figures 3.7 and 3.8 for the lower packet data rates are typical of little contention for the channel. This observation is borne out from an examination of Figures 3.10 and 3.11, in which the average number of backoff slots for standard and priority packets, respectively, is plotted vs. time for node packet average interarrival times of 1.0, 0.1, 0.05, and 0.01 seconds (source rates: 1, 10, 20, and 100 packets per second). In both figures, for the low data rate of 1 packet/sec, the average number of backoff slots awaiting transmission at the MAC layer is quite steady at the theoretical value of 2.5 slots for priority packets (corresponding to $\lambda = 0.4$ in Figure 2.5) and the theoretical value of 8 slots for standard packets (consistent with the default *CW*_max_*size* value of 15 that is implemented in the OPNET 802.11 model as a starting value, before any collisions). This behavior indicates that there are very few collisions for the 1 packet/sec data rate.

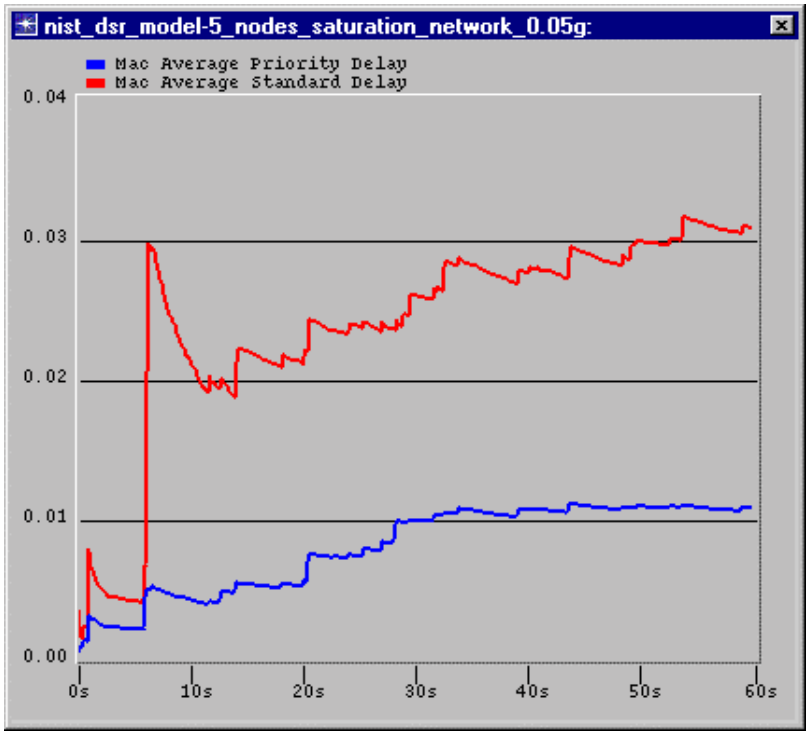

Figure 3.9. Average MAC packet delays for traffic generated at an average rate of 20 packets/sec.

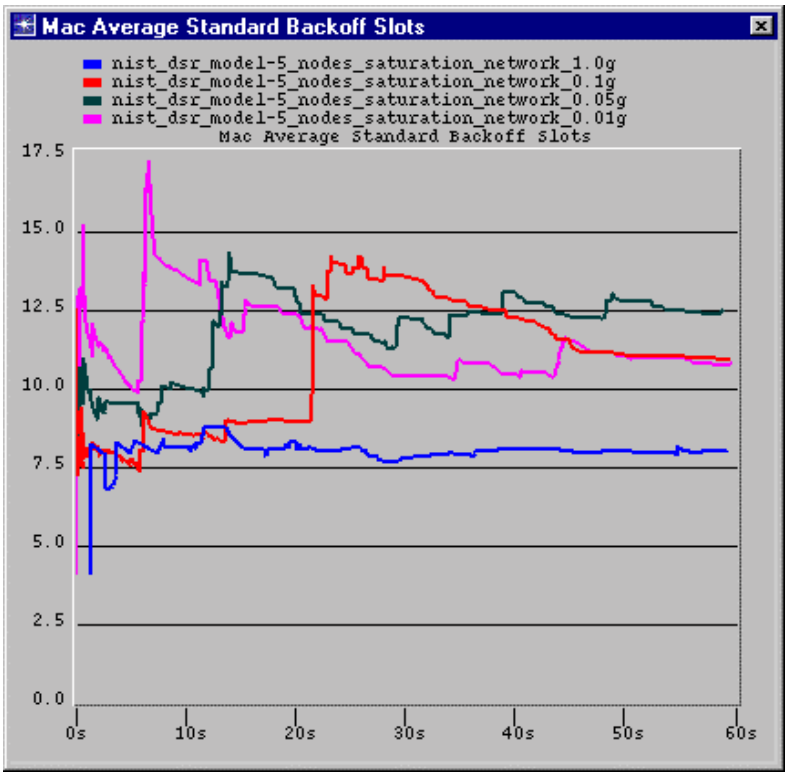

Figure 3.10. Average MAC backoff slots for standard packets.

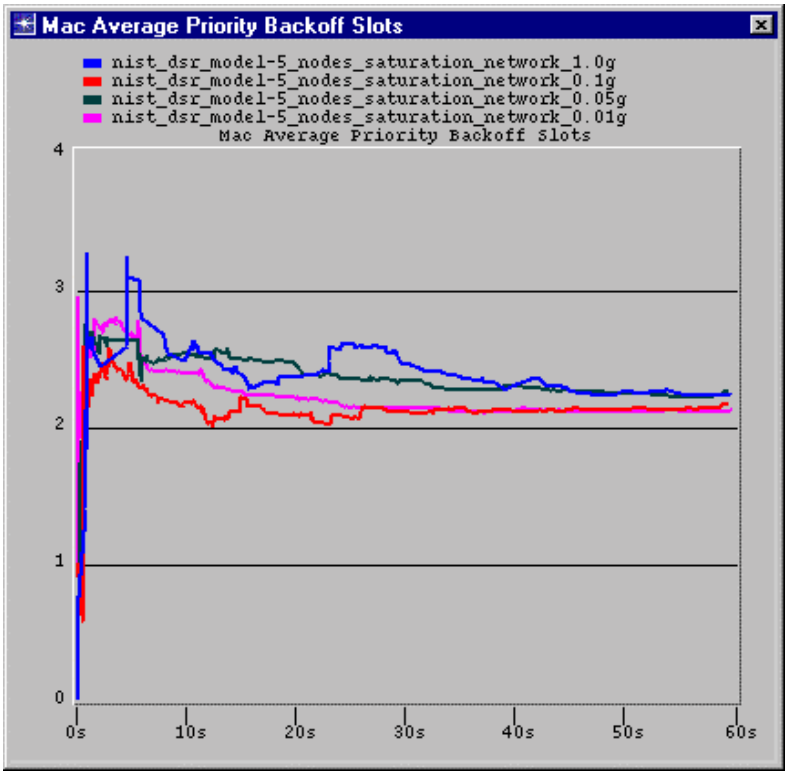

Figure 3.11. Average MAC backoff slots for priority packets.

Using the average time of the backoff slots, $T_{backoff} = 2.5 \times 50\mu s = 125\,\mu s$ for priority packets and $T_{backoff} = 8 \times 50\mu s = 400\,\mu s$ for standard packets, we can revise the estimates of T_{total}, the time on the channel used by the two different kinds of packet, to

$$T'_{total} = T_{data} + \text{SIFS} + T_{ACK} + \text{DIFS} + T_{backoff} = \begin{cases} 1.313\,\text{ms}, & \text{priority} \\ 1.638\,\text{ms}, & \text{standard} \end{cases}$$

which lead to estimates of the upper limit on (nonblocking) traffic sources rates of about 95 packets/sec for priority packets and about 76 packets/sec for standard packets. The actual limit can be expected to be lower because of collisions, which not only "waste" time but induce larger backoff times.

Now we consider the interesting behavior of the MAC average number of backoff slots for standard packets that is displayed in Figure 3.10 for packet source rates greater than 1 packet/sec. The nearly discrete "jumps" in the value of the averages apparently are the result of single collisions requiring sharp increases in the backoff time for retransmission attempts. The steady "leakage" of the values after the jumps (to use an analogy with the voltage step on a capacitor due to a current impulse) is consistent with the occurrence of collisions for a few packets, separated by non-colliding transmissions for other packets. The collisions appear to start earlier as the data rate is increased, probably because the absence of the RTS/CTS mechanism (due to the small size of the packets) leaves this particular network vulnerable to the hidden terminal problem.

3.3.4 DSR and Upper Layer Statistics

In connection with the capacity of the network at the MAC and physical layers, and the congestion that affects that capacity, it is interesting to look at some of the statistics recorded by the simulation pertaining to quantities at the DSR and upper layers.

The number of packets enqueued at the DSR layers of all five nodes, awaiting servicing by the MAC layer, is shown plotted vs. time for node packet average interarrival times of 1.0, 0.1, 0.05, and 0.01 seconds (source rates: 1, 10, 20, and 100 packets per second) in Figure 3.12 for standard packets and in Figure 3.13 for priority packets. Note even for the data rates 10 and 20 packets/sec that there is a perceptible "creeping up" of the number of packets in the queue, although the priority packets are not accumulating as fast as the standard packets.

For the source rate of 100 packets/sec at each node, both the standard and the priority DSR queues are seen in Figures 3.12 and 3.13, respectively, as accumulating at a rate much faster than the MAC layer can handle. This behavior indicates that the network capacity has been exceeded. The slope of these accumulations, expressed as a data rate, can be considered the excess of the aggregate source rate over the network capacity. For the standard packets, the slope in Figure 3.12 is approximately 7000 packets in 60 seconds, or 117 packets/sec, which divided by eight (to account for the five nodes plus relaying) becomes approximately a 15 packets/sec per node excess. Similarly, for the priority packets, the slope in Figure 3.13 is approximately 6000 packets in 60 seconds, or 100 packets/sec, which divided by eight becomes approximately a 13 packets/sec per node excess.

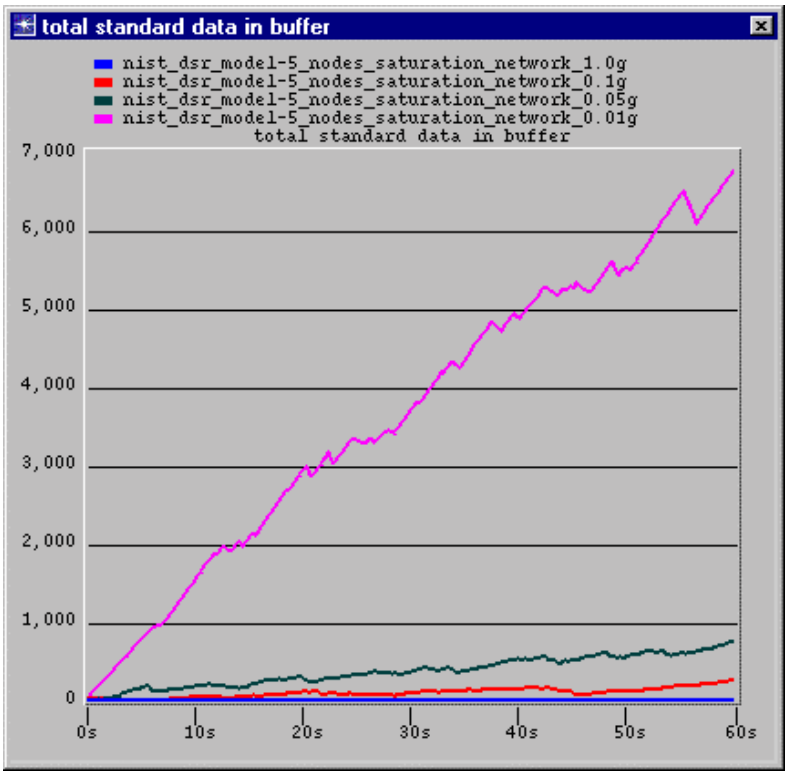

Figure 3.12. Total standard data in buffer at the DSR layer.

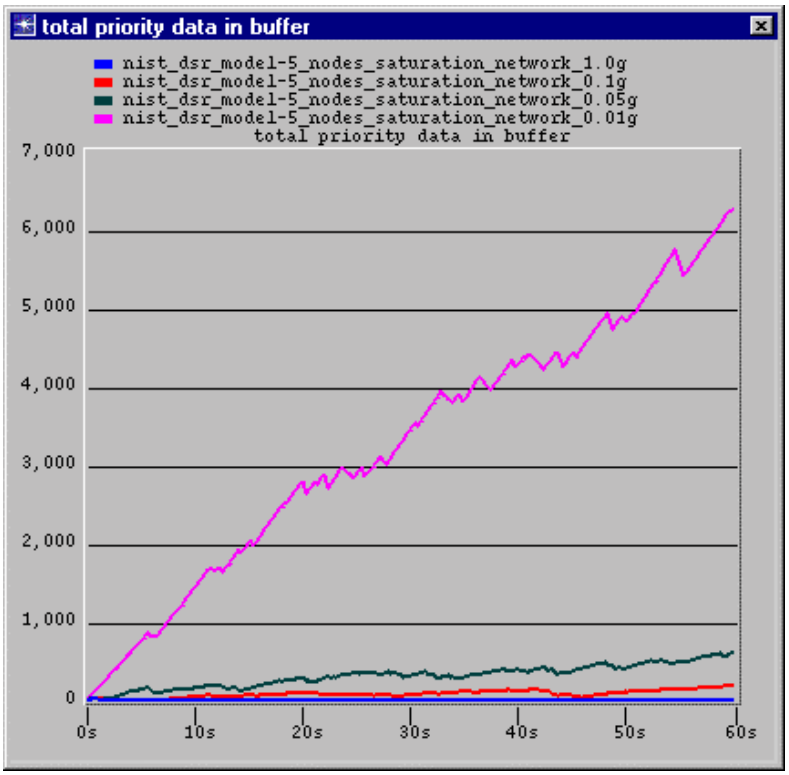

Figure 3.13. Total priority data in buffer at the DSR layer.

An example upper layer statistic is the upper layer packet throughput that is displayed in Figure 3.14 for standard packets and in Figure 3.15 for priority packets. For all packet data rates at each node, the throughput is calculated as the cumulative number of packets served divided by the cumulative (simulated) time since the start of the simulation. We observe from these two figures that the throughputs for the two types of packets are almost the same, with the numbers for priority packets being slightly higher. The throughput values for source data rates of 1 and 10 packets/sec in Figures 3.14 and 3.15 are consistent with the fact that half of the generated packets are standard packets and half are priority packets; thus the throughput for either type of packet is about 0.5 packets/sec when the source rate is 1 packet/sec, and the throughput for either type of packet is slightly less than 5 packets/sec when the source rate is 10 packets/sec. In fact, as displayed in Figure 3.16 and 3.17, the upper layer efficiencies for these rates, defined as the ratio of number of packets delivered to the destination upper layer to the number of packets transmitted from the source upper layer, are about 100% and 88%, respectively.

For a source data rate of 20 packets/sec, the throughput for either type of packet is around 7.5 packets/sec, which is about 75% of the 10 packets/sec that would be delivered for 100% efficiency. As Figures 3.16 and 3.17 show, the efficiency for the two packet types is a relatively steady 76% for standard packets and 81% for priority packets. A steady throughput with an efficiency of less than 100% is consistent with either a percentage packet loss or, as in our case, with the queuing backlogs we have observed.

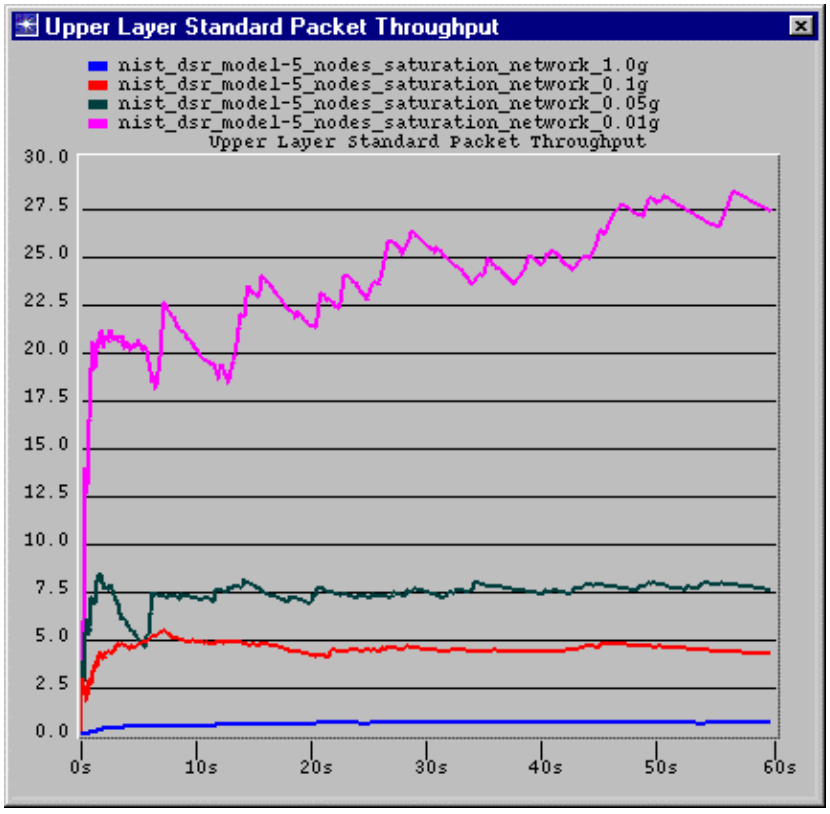

Figure 3.14. Upper layer standard packet throughput.

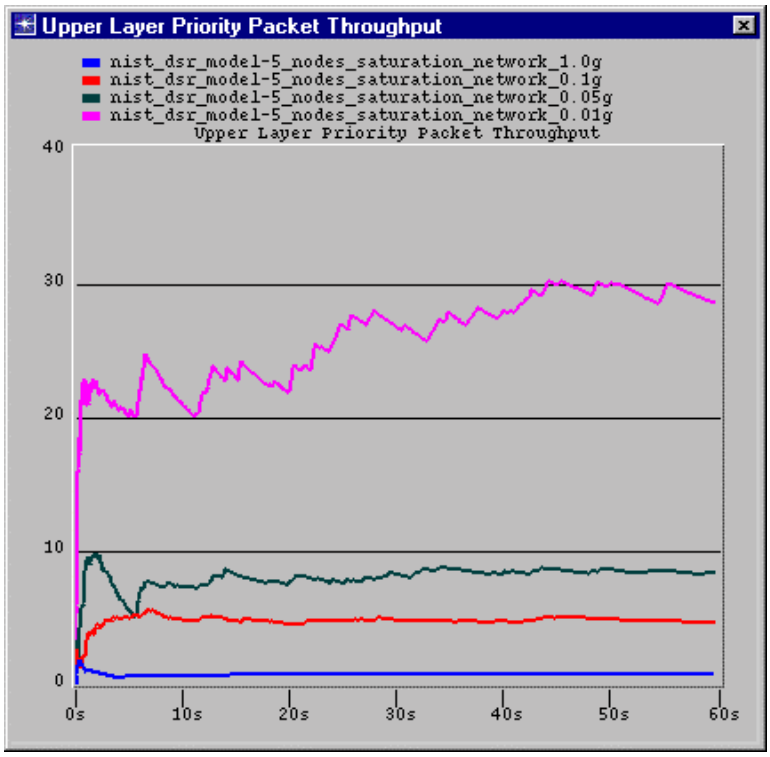

Figure 3.15. Upper layer priority packet throughput.

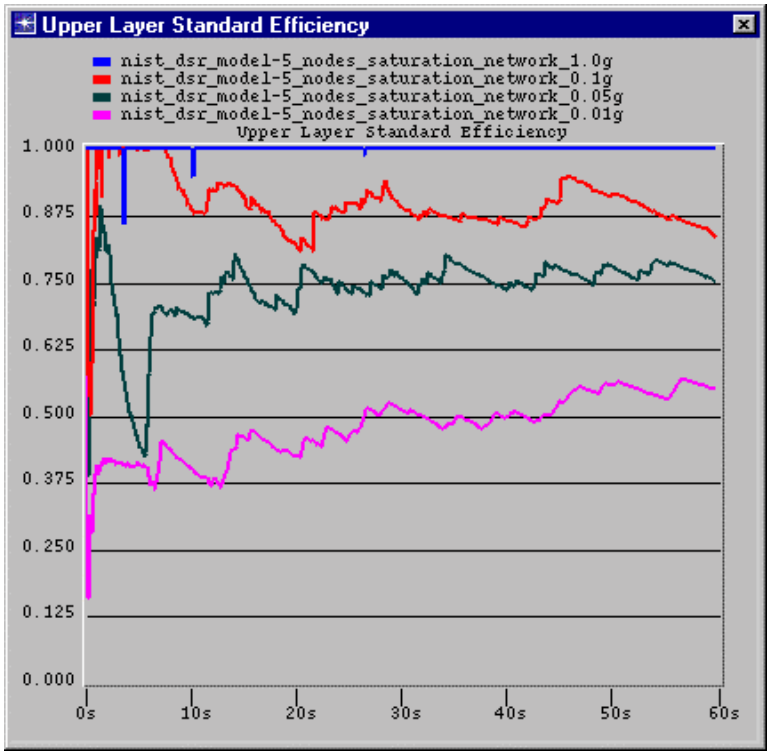

Figure 3.16. Upper layer standard packet efficiency.

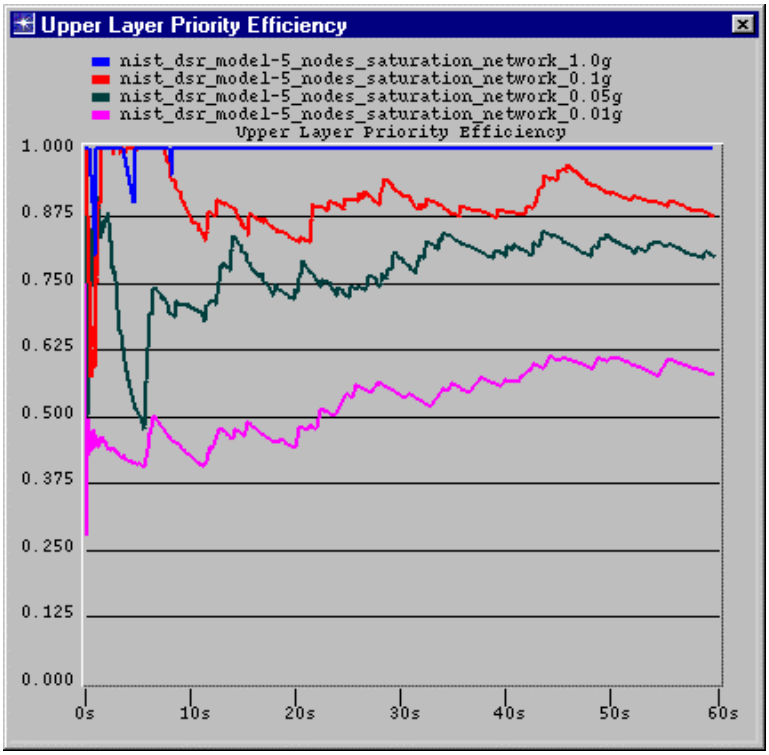

Figure 3.17. Upper layer priority packet efficiency.

For the source data rate of 100 packets/sec, which we know is above the capacity of the network, the throughput for either type packet begins around 20 packets/sec (40% efficiency) and rises toward 30 packets/sec (60% efficiency) during the simulation. The explanation for this behavior is apparently that, once the queues become backlogged, the traffic is served not in bursts as it is generated but at a more even pace that permits a more efficient and "orderly" access to the channel.

3.3.5 Comparison of MAC at Center and Outer Nodes

Since the network simulation scenario of Figure 3.3 is set up to create congestion at the center node, it is interesting to compare MAC layer statistics at the center node with those at the uncongested outer nodes. In Figures 3.18 and 3.19, the MAC transmission queue size for standard packets is shown as a function of time for node packet average interarrival times of 1.0, 0.1, 0.05, and 0.01 seconds (source rates: 1, 10, 20, and 100 packets per second) at an outer node and at the center node, respectively. The same quantity for priority packets is shown in Figure 3.20 for an outer node and in Figure 3.21 for the center node.

Comparing outer node to center node, the difference in congestion is quite obvious in its effect on the MAC queue size. There is roughly four times as much traffic being offered to the MAC layer of the center node, and the peak queue sizes for the center node are about four times those for the outer node when the source packet rate is 100 packets/sec.

Comparing standard packet processing to priority packet processing, there are two effects that can be observed in Figures 3.18 to 3.21. First, the peak values of priority MAC queue sizes

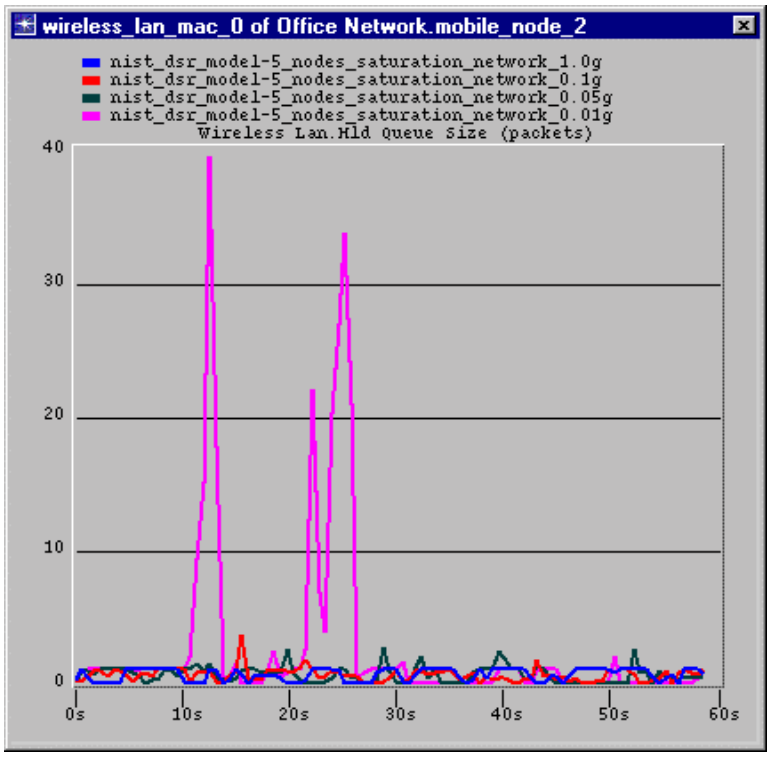

Figure 3.18. MAC layer queue size for standard packets at outer node.

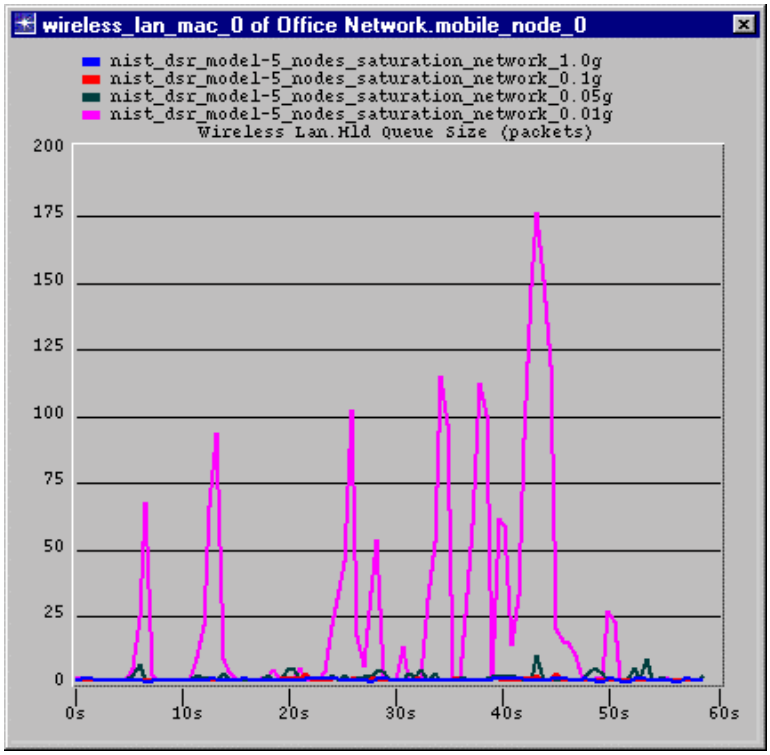

Figure 3.19. MAC layer queue size for standard packets at center node.

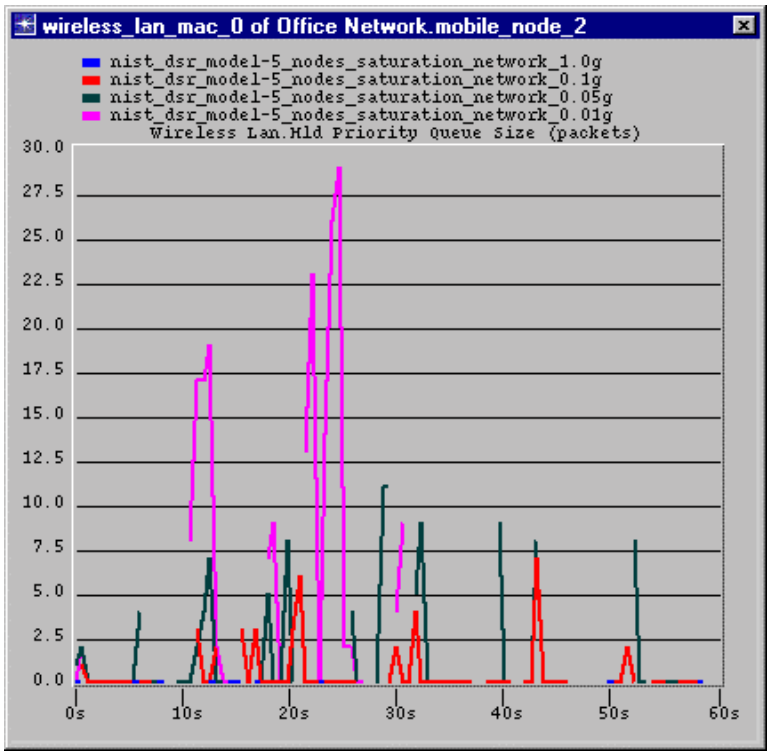

Figure 3.20. MAC layer queue size for priority packets at outer node.

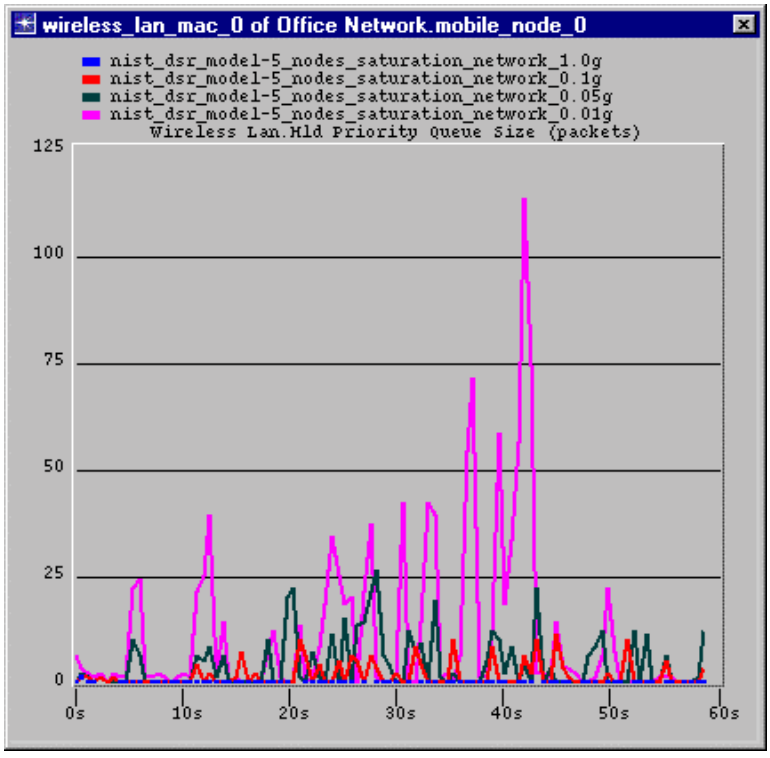

Figure 3.21. MAC layer queue size for priority packets at center node.

are much reduced relative to those of standard MAC queue sizes for 100 packets/sec. Second, the peak MAC queue sizes for priority traffic for nonblocking packet rates less than 100 packets/sec are generally larger than the MAC queue sizes for standard traffic for the same packet rates. The explanation for this behavior is not certain; it could be that the preference given to priority packets at the DSR layer results in bursts of priority packets being sent to the MAC layer, or it could be that the short backoff times of priority packets results in more collisions during contention access.

3.4 Conclusion and Objectives for Further Work

We conclude that the simulation results obtained so far indicate that the new IEEE 802.11 MAC layer priority mechanisms that have been proposed can be effective. The feasibility of these mechanisms having been demonstrated, further work is needed in order to perform a thorough parametric evaluation of the mechanisms. Objectives for this further work include the following:

- Perform further diagnostic simulations using the scenario of Figure 3.3 to clarify the role of the different parameters controlling the priority mechanisms. In particular, perform simulations using data packet sizes sufficient to trigger the IEEE 802.11 RTS/CTS collision avoidance mechanism in order to observe the working of the priority mechanisms for different numbers of collisions.
- Develop "typical" WLAN and MANET scenarios and evaluate the effectiveness of the priority mechanisms. Adjust the parameters of the mechanisms to determine the sensitivity of the results to the values of the parameters.
- Extend the two-level priority mechanisms to three- and four-level mechanisms and determine the appropriate values of the parameters distinguishing the levels of priority traffic and controlling the mechanisms.
- Investigate "fairness" concepts and develop them into constraints on the parameters of the priority mechanisms.

REFERENCES

[1] Z. J. Haas, "Guest Editorial" in special issue on wireless ad hoc networks, *IEEE J. on Selected Areas in Communications*, vol. 17, no. 8 (August 1999).

[2] J. Macker and M. S. Corson, "Mobile Ad Hoc Networking and the IETF," *Mobile Computing and Communications Review*, vol. 2, pp. 9-14 (January 1998).

[3] ——, http://www.ietf.org/html.charters/manet-charter.html, home page of MANET working group.

[4] D. B. Johnson and D. A. Maltz, "Dynamic Source Routing in Ad-Hoc Wireless Networks," *Mobile Computing*, T. Imielinski and H. Korth, eds., Kluwer, 1996, pp. 153-81.

[5] J. Broch, D. B. Johnson, and D. A. Maltz, "The Dynamic Source Routing Protocol for Mobile Ad Hoc Networks," IETF Internet Draft (work in progress).

[6] X. Pallot, N. Roux, and J.-S. Pegon, "NIST OPNET Model for DSR," available online at http://w3.antd.nist.gov/wctg/DSRreadme.doc, December 2000.

[6] ——, http://www.manta.ieee.org/groups/802/.

[7] ——, http://www.manta.ieee.org/groups/802/11/main.html.

[8] ——, IEEE Standard 802.11–1997, "Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications," LAN MAN Standards Committee of the IEEE Computer Society, 26 June 1997.

[9] L. Kleinrock and F. A. Tobagi, "Packet Switching in Radio Channels: Part I—Carrier Sense Multiple-Access Modes and Their Throughput-Delay Characteristics" and "…Part II—The Hidden Terminal Problem in Carrier Sense Multiple-Access and the Busy-Tone Solution," *IEEE Trans. on Communications*, vol. COM-23, pp. 1400-1433 (December 1975).

[10] H. Takagi and L. Kleinrock, "Optimal Transmission Ranges for Randomly Distributed Packet Radio Terminals," *IEEE Trans. on Communications*, vol. COM-32, pp. 246-257 (March 1984).

[11] N. Abramson, "The Throughput of Packet Broadcasting Channels," *IEEE Trans. on Communications*, vol. COM-25, pp. 117-128 (January 1977).

[12] D. Scott, "Priority Access Service (PAS)," *Proc. 14th Federal Wireless Users' Forum Workshop*, Dec. 12-14, 2000, Paradise Valley, AZ.

[13] S. Boztas and F. J. W. Symons, "Robust multi-priority topology-independent transmission schedules for packet radio networks," *Proc. 1994 IEEE Information Theory Symposium*, p. 411.

[14] J. Shor and T. G. Robertazzi, "Traffic Sensitive Algorithms and Performance Measures for the Generation of Self-Organizing Radio Network Schedules," *IEEE Trans. on Communications*, vol. 41, pp. 16-21 (January 1993).

[15] B. Hajek and G. Sasaki, "Link Scheduling in Polynomial Time," *IEEE Trans. on Information Theory*, vol. 34, pp. 910-917 (September 1988).

[16] A. Ephremides and T. V. Truong, "Scheduling Broadcasts in Multihop Radio Networks," *IEEE Trans. on Communications*, vol. 38, pp. 456-460 (April 1990).

[17] K. Enomoto, S. Shiokawa, and I. Sasase, "Performance of Inhibit Sense Multiple Access for Prioritized Traffic," *Proc. 1996 IEEE Internatl. Conf. on Universal Personal Communications*, pp. 22-26.

[18] P. Papantoni-Kazakos et al., "Transmission Algorithms for a Multi-Channel Packet Radio System with Priority Users," *Proc. GlobeCom '92*, pp. 89-93.

[19] A. M. Glass and R. L. Brewster, "Reservation with Contention-Based Traffic Demand Assignment Protocol for Land Mobile Radio Communications," *Proc. 1989 International Conf. On Mobile Radio and Personal Communications*, pp. 6-9.

[20] M. Ogawa, T. Sueoka, and T. Hattori, "Priority Based Wireless Packet Communication with Admission and Throughput Control," *Proc. IEEE 2000 Spring Vehicular Technology Conference*, pp. 370-374.

[21] A. Petrick, J. Zyren, and J. Figueroa, "Delivering Voice over IEEE 802.11 WLAN Networks," Harris Corp. white paper. http://www.intersil.com/prism/papers/

[22] T. Papantoni-Kazakos, N. B. Likhanov, and B. S. Tsybakov, "A Protocol for Random Multiple Access of Packets with Mixed Priorities in Wireless Networks," *IEEE J. on Selected Areas in Communications*, vol. 13, pp. 1324-1331 (September 1995).

[23] L. Georgiadis and M. Paterakis, "Performance Analysis of Window Type Random-Access," *Proc. IEEE INFOCOM '89*, pp. 505-511 and "Bounds on the Delay Distribution of Window Random-Access Algorithms," *IEEE Trans. on Communications*, vol. 41, pp. 683-693 (May 1993).

[24] J. Sau and C. Scholefield, "Scheduling and Quality of Service in the General Packet Radio Service," *Proc. 1998 IEEE Internat'l Conf. On Universal Personal Communications*, pp. 1067-1071.

[25] Q. Pang et al., "Service Scheduling for General Packet Radio Service Classes," *Proc. 1999 IEEE Wireless Commun. and Networking Conf.*, pp. 1229–1233.

[26] "M. Chuah et al., "Access Priority Schemes in UMTS MAC," *Proc. 1999 IEEE Wireless Commun. and Networking Conf.*, pp. 781–786.

[27] J. J. Zavgren, "The Moment-of-Silence Channel-Access Algorithm," *Proc. MILCOM '89*, pp. 388-394.

[28] X. Pallot, N. Roux, and J.-S. Pegon, "DSR Model Reference Appendix," available as the file DSRappx.pdf at http://w3.antd.nist.gov/wctg/prd_dsrfiles.html

[29] M. W. Subbarao, "Ad Hoc Networking Critical Features and Performance Metrics," white paper, Wireless Communications Technology Group, NIST, 15 September 1999.