12-2006

Channel-Access and Routing Protocols for Wireless Ad Hoc Networks with Directional Antennas

Arvind Swaminathan
Clemson University, arvind.phd@gmail.com

CHANNEL-ACCESS AND ROUTING PROTOCOLS FOR WIRELESS AD HOC NETWORKS WITH DIRECTIONAL ANTENNAS

A Dissertation
Presented to
the Graduate School of
Clemson University

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy
Electrical Engineering

---

by
Arvind Swaminathan
December 2006

Accepted by:
Dr. Daniel L. Noneaker, Committee Chair
Dr. Harlan B. Russell
Dr. Carl W. Baum
Dr. Jim Martin

ABSTRACT

Medium-access control (MAC) and multiple-hop routing protocols are presented that exploit the presence of directional antennas at nodes in a wireless ad hoc network. The protocols are designed for heterogeneous networks in which an arbitrary subset use directional antennas. It is shown that the new protocols improvement the network's performance substantially in a wide range of scenarios.

A new MAC protocol is presented that employs the RTS/CTS mechanism. It accounts for the constraints imposed by a directional antenna system, and it is designed to exploit the capabilities of a directional antenna. It is shown that the receiver blocking problem is especially detrimental to the performance if the network includes nodes with directional antennas, and a simple solution is presented. A further improvement to the MAC protocol is presented which results in more efficient spatial reuse of traffic channels in the heterogeneous network. The protocol includes a mechanism by which a negotiating node pair dynamically determines if a traffic channel that is in use in the local area can be used concurrently to support additional traffic. It is shown that the new protocol yields significantly better performance than two existing approaches to the reuse of traffic channels. It is also shown that the improvements are achieved over a wide range of network conditions, including different network densities and different spread-spectrum processing gains.

A new distributed routing protocol is also presented for use in heterogeneous wireless ad hoc networks. Two components of the routing protocol are jointly designed: a congestion-based link metric that identifies multiple routes with low levels of congestion, and a forwarding protocol that dynamically splits traffic among the multiple routes based on the relative capabilities of the routes. It is shown that the new routing protocol is able to exploit the decoupling of paths in the network resulting from the presence of nodes with directional antennas. Furthermore, it is shown that the protocol adapts effectively to the presence of advantaged nodes in the network. This approach to joint routing and forwarding is shown to result in a much better and more robust network performance than minimum-hop routing.

DEDICATION

To Appa and Amma.

To Mangalam Thathi.

To the *knowledge is power* atmosphere at Padma Seshadri and the Rajagopalan family household.

To all the teachers who have taught me over the years.

ACKNOWLEDGMENTS

This opportunity to thank the people who have influenced and supported over the years is definitely one of the sweetest rewards of completing my Ph.D. dissertation. I would like to begin by thanking my extended family. It is their influence that has led me to develop a passion for intellectual pursuits. I would like to thank my parents, Geetha and Swaminathan Eisenhower, who have sacrificed a lot in their lives so that my siblings and I could get the best education possible. I would also like to thank my sister Arthi who has been a constant source of encouragement and support during the last six years, and my brother Abhi who has had to endure all my “from experience” talks. My grandmother, Mrs. Mangalam Rajagopal, has also been a pillar of support, as have been all my uncles, aunties and cousins. Kumar mamma in particular has been a constant source of support and I only wish I had his energy and enthusiasm for life. Last but not the least, I would like to thank Pooja and her family. The ups and downs associated with Ph.D. research have been easier to ride this past year because of you, honey.

This brings me to the set of people who have helped me convert the enthusiasm for “knowledge acquisition” created by my family into tangible results, one of which is this dissertation. Once again an exhaustive list of all the teachers who have helped me understand the complicated and beautiful concepts in math, science and engineering would take an entire book by itself and hence I restrict myself to mentioning a few of them. I would like to thank my high-school science teacher in Palakkad, Warrier Mash, under whose tutelage that I first began the journey from learning things by rote to understanding the concepts and using it to solve problems. I would also like to express my thanks to all the teachers in Calicut REC who started my engineering education.

Without doubt, my years at Clemson have been the best in terms of accomplishing my lifetime goal of moving closer to becoming an expert in a field and doing original research. In large part, this is due to the encouragement, enthusiasm and the direction provided by my advisor Dr. Dan Noneaker. His professionalism, ability to shape my raw ideas, generous praise and his patient ear for my problems both in the research and non-research fronts have been extremely important. I also had the pleasure of working closely with Dr. Harlan Russell during my Ph.D.. I have fond memories of the many lengthy discussions about various issues in wireless networks that I have had with him and Dr. Noneaker. The insights on communication systems that these discussions have given me will definitely make me a better communications engineer in the future. I would also like to thank Dr. Carl Baum and Dr. Jim Martin for serving on my committee. Finally, I would like to thank all the faculty members in the Clemson Wireless program for developing such an excellent set of courses that helped me develop the fundamentals required to perform my research and which, I am certain, will serve me well in future.

It is impossible to survive the many years I have spent in Clemson without having great friends, and I've had the good fortune of having had great company both in the wireless lab and outside. My ex room-mates Varad and Deepesh endured me for many years and our times together is greatly missed. The times with my current room-mates Venki, Balajee and Karthik will also be missed and I wish all of them the best in the remainder of their Ph.D. programs. I have also enjoyed the fun times with Aman Behal, the AID gang, Harish and his room-mates, and the current gang consisting of Rohan, Raghavan, Rane, Brigu and Kunal.

I would like to end by thanking the Office of Naval Research, the Army Research Office and the other agencies that have generously supported my Ph.D. studies.

TABLE OF CONTENTS

Table of Contents (Continued)

Page

Table of Contents (Continued)

Page

LIST OF TABLES

List of Tables (Continued)

LIST OF FIGURES

List of Figures (Continued)

Figure Page

List of Figures (Continued)

Figure Page

List of Figures (Continued)

Figure Page

List of Figures (Continued)

Figure Page

CHAPTER 1

INTRODUCTION

Wireless communications has been one of the biggest technological success stories of the last decade. The convenience of accessing information from any place has lead to an explosive growth in the adoption of wireless technologies. In order to support the increasing data rates required from wireless communication systems, impairments like multipath-fading and multiple-access interference have to be overcome. Using directional antennas to eliminate interference in the spatial domain and to combine signals received along multiple paths is a promising technique to overcome these impairments.

As a result, the use of directional antennas to improve the performance of wireless communication systems has received considerable attention in the recent past. The majority of research in the topic [1]-[4] has focused on the use of directional antennas to improve system performance as measured by performance at the physical layer. Both a directional transmitter antenna and a directional receiver antenna increase the link gain, and a directional receiver antenna also provides spatial discrimination against sources of interference. Each of these factors contributes to an increased signal-to-interference-plus-noise ratio and thus an increased link capacity for a given communications channel. The directional-antenna gains also result in an extended range for a viable link. Furthermore, in many circumstances the use of a directional transmitter antenna reduces the number of neighboring nodes that are susceptible to interference from the transmitter's signal.

Much of the research on directional antennas for wireless communications has addressed systems using direct-sequence spread-spectrum modulation, and most of that research is focused on commercial code-division multiple-access cellular networks [3]. The use of a star topology centered on the cellular base station can be exploited in designing steered-beam or adaptive-array antenna subsystems, and the resulting performance gains at the physical layer result immediately in corresponding improvements in network-level performance.

Direct-sequence (DS) spread-spectrum modulation also provides significant advantages for multiple-channel mobile ad hoc packet radio networks used in the tactical military environment or disaster-relief operations due to its resistance to jamming and its low probability of intercept [5]. Directional antennas have the potential to greatly improve the performance of these networks as well. Yet the improved physical-layer performance provided by directional antennas can only be translated into significant gains in the end-to-end performance of these networks if the channel-access and network-layer protocols are designed to exploit the capabilities of directional antennas effectively. In this dissertation, we present the design of such protocols to translate the physical layer benefits of directional antenna systems into network-level improvements in performance. We will see that this requires the design of protocols that account for certain unique characteristics of the directional antenna system and that exploit certain additional degrees of freedom that result from the presence of nodes with directional antennas.

A number of researchers have proposed protocols for ad hoc networks with directional antennas, though most of the work is focused on single-channel, narrowband communications. Existing work on channel-access protocols with directional antennas is focused almost entirely on networks in which all of the nodes have directional antennas, and many of the protocols depend on the use of complicated directional antenna subsystems that require location information to achieve beam-pointing. In contrast, the work in this dissertation is focused on the design of simple distributed protocols that support an arbitrary mix of nodes with antennas of differing degrees of directionality and that do not require complex antenna-subsystem capabilities at the nodes. In keeping with this approach, our results are presented in the context of a simple, more-robust directional antenna system (described in Chapter 2), and we focus on realistic scenarios in which only a moderate fraction of the nodes in the network have directional antennas. The protocols we present in this dissertation are applicable to ad hoc networks employing multiple-channel communications, a mix of directional and omni-directional antennas, and either DS spread-spectrum modulation or narrowband modulation.

The design of channel-access protocols for ad hoc networks with directional antennas is illustrated by the work in [6]-[10]. The existing work demonstrates that the use of directional antennas improves the spatial utilization of network resources in comparison with a network in which omni-directional antennas are employed. It is also shown that if the channel-access protocols are designed to exploit this advantage, the link-layer performance can be improved.

The work in [10], [7], [8] generalizes the concept of a network allocation vector (NAV) [11] to a directional-NAV. The directional-NAV is similar to the basic channel-access mechanism we have examined for exploiting directional antennas. Some shortcomings of this approach are shown in Chapter 6 of this dissertation, however, and an enhancement to the protocol that mitigates the problem is presented in the same chapter.

The design of a channel-access protocol in a wireless ad hoc network is driven by two counter-acting objectives: efficient spatial reuse of traffic channels, and providing each on-going transmission with adequate protection against multiple-access interference. Clearly, the number of simultaneous packet exchanges permitted by the channel-access protocol is greater if the spatial reuse is more aggressive. However, if the spatial reuse is such that the resulting multiple-access interference at many of the receivers is too high, a large fraction of these simultaneous packet exchanges will fail, thus reducing the effective throughput. Hence, network throughput is maximized by striking a balance between the amount of spatial reuse and the resulting multiple-access interference. The best reuse policy for use by a node will depend in general on the packet transmission format, the network topology, and other conditions in the neighborhood of the node.

Conventional channel-access protocols exclusively reserve network resources within range of a transmitter-receiver pair using a Request-to-Send (RTS) and Clear-to-Send (CTS) exchange [12]. Thus the first of the two counter-acting objectives specified above is implicitly sacrificed in favor of the second. This exclusive-reservation is particularly wasteful in a system employing DS spread-spectrum modulation because of the multiple-access interference rejection capability of the modulation format. In Chapter 7, we present enhancements to the channel-access protocol that implement a more aggressive spatial reuse approach. We demonstrate that these protocols improve performance by striking a better balance between the two counter-acting objectives.

There has been less extensive prior research [13]-[15] on network-layer (routing) protocols than on channel-access protocols for ad hoc networks with directional antennas. Some of the research addresses ways of exploiting directional antennas to reduce the flooding of route requests in on-demand routing protocols [13],[14]. The work reported in [15] concerns the development of link-state congestion-avoidance routing protocols that exploit the availability of directional antennas. As with most of the prior research, the protocols described in [15] are for networks in which all the nodes have directional antennas.

In this dissertation, we present the design of a distance-vector routing protocol that exploits the presence of nodes with directional antennas in a heterogeneous network containing an arbitrary mix of nodes with directional antennas and nodes with omnidirectional antennas. Employing insights from simulations based on a simplified network model, we demonstrate that the presence of directional antenna nodes reduces the *mutual coupling* (interference) between the different routes in the network. This suggests that a routing protocol which attempts to reduce the level of congestion in the network by finding alternate paths for the different source-destination pairs with lower mutual coupling will improve the performance of a network with directional antenna nodes. In Chapter 8 we present the design of such a routing protocol that uses a modified path metric which accounts for the amount of congestion in a path. We demonstrate that this routing protocol exploits the additional degree of freedom made possible by the presence of directional antenna nodes, and it consequently results in dramatic improvements in the network's performance under many conditions.

CHAPTER 2

SYSTEM DESCRIPTION

A subset of the nodes in the network employ multiple directional antennas at the physical layer, and each antenna has a corresponding half-duplex radio transceiver. Each of the remaining nodes has an omnidirectional antenna and a single radio transceiver. In the following, the coverage area of each directional or omnidirectional antenna is referred to as a *sector*. Hence there is a one-to-one correspondence between transceivers and sectors at each node. The network uses two or more frequency channels, with one frequency channel serving as a MAC-layer control channel and the remaining frequency channels serving as MAC-layer data channels. Each transceiver for a directional antenna or an omnidirectional antenna can use any one frequency channel at a time. Hence a transceiver can at most receive on one channel or transmit on one channel at any instant of time, but not both.

Imperfect electromagnetic isolation among multiple transceivers at a node introduces the possibility of *co-site interference* among the transceivers if two or more of the transceivers concurrently employ the same frequency channel. In particular, concurrent transmission by one half-duplex transceiver and reception by another half-duplex transceiver at the same frequency can result in a poor signal-to-interference ratio in the received signal. In this dissertation it is assumed that a node with directional antennas is susceptible to co-site interference among the transceivers for its different sectors and that the MAC protocol must be designed to account for this possibility. Thus if a node transmits into one of its sectors on a given channel, the MAC protocol must be designed to preclude concurrent use of the same channel for reception from any sector at that node.

2.1 The Physical Layer and the Channel

Each packet is transmitted using direct-sequence spread-spectrum modulation with a quaternary spreading sequence, and the packet consists of an acquisition preamble followed by the packet's data portion. The chip rate of each transmission is 8 Mchips/s in the networks with two frequency channels, and it is 4 Mchips/s in the networks with four frequency channels. The length of the acquisition preamble is the same for all packets in a given network. A preamble length of 1000 chips is used if the chip rate is 8 Mchips/s, and the preamble length is 500 chips of the spreading sequence for the 4 Mchips/s networks. (The preamble has no data modulation.) The information in each packet is encoded with the NASA-standard, rate one-half, constraint-length seven convolutional encoder, and the encoder output constitutes the data portion of the packet. It is sent using binary antipodal modulation of the spreading signal with a processing gain that depends on the type of packet and the information rate of the transmission.

Each network that we consider uses a single packet format for all MAC-layer data packets and a single packet format for all MAC-layer control packets, though the formats differ for different networks. The number of information bits in a data packet depends on the instantaneous information rate of the packet format. For the networks using a chip rate of 8 Mchips/s, we consider five information rates for the data packets: 100 kbits/s, 200 kbits/s, 400 kbits/s, 800 kbits/s and 4 Mbits/s. The packets contain 500, 1000, 2000, 4000 and 20,000 bits of information, respectively, for the five information rates. The encoder outputs are spread with a processing gain of 40, 20, 10, 5 and one quaternary chips per binary channel symbol for the five respective information rates. For each network using a chip rate of 4 Mchips/s, the data packets have an information rate of 100 kbits/s. They contain 500 bits of information and are spread with 20 chips per channel symbol. The transmission time is thus 5.25 ms for all MAC-layer data packets in each network considered in the dissertation.

Each MAC-layer control packet is spread with the same processing gain as the data packets in the network. Each control packet contains 25 bits of information in the networks using a chip rate of 4 Mchips/s. The information content of the control packets differs for different networks using a chip rate of 8 Mchips/s. If the data packets in the latter type of network have an instantaneous information rate of 100 kbits/s, 200 kbits/s, or 400 kbits/s, the control packet each contain 50 bits of information. If the data packets have an information rate of 800 kbits/s, each control packet contains 100 bits of information; and if the data packets have an information rate of 4 Mbits/s, each control packet contains 500 bits of information. Thus the transmission time for a control packet ranges between 250 μs and 625 μs for the various networks. The details of the transmission formats used in the networks in each chapter are given in Table 2.1.

Some types of packets are transmitted using a spreading code that is common to all the nodes in the network (*common spreading*), and other types of packets are transmitted using a spreading code that is specific to the intended link destination (*receiver-directed spreading*). The packet format and MAC protocol are such that a link data rate of 71 packets/s is achieved by a single link-level source-destination pair using pacing (see Chapter 3) in the absence of noise, interference, and contention. All transmissions occur at the same power.

The channel is modeled as an additive white Gaussian noise (AWGN) channel with a path-loss coefficient of three and a power-gain constant of α. If the transmitter and receiver both use omnidirectional antennas, the received signal power P_r is thus given by

$$P_r = P_t \times \alpha \times \left(\frac{\lambda}{4\pi d}\right)^3 \tag{2.1}$$

where P_t is the transmit power, λ is the wavelength and d is the distance between the transmitter and receiver. (Only a power-gain factor of one is considered in most of the dissertation. Power-gain factors of two, four, eight, and forty are also considered in Chapter 7, however.)

The received signal is also subjected to multiple-access interference due to other concurrent transmissions in the network. The transmitted power and the receiver's noise power spectral density are such that in the absence of multiple-access interference, the signal-to-noise ratio for a bit of information is 11.5 dB if the transmitter and receiver both use an omnidirectional antenna, they are separated by a distance of 1350 m, and the packet format of Chapter 6 is used in a channel with a power-gain factor of one.

The receiver employs noncoherent serial, matched-filter acquisition based on the acquisition preamble. The acquisition stage is designed to acquire the preamble of a packet that uses either the common spreading code or the node's unique spreading code. The data is detected by coherent demodulation and hard-decision Viterbi decoding.

Chap.	Num. of chnls.	Chip Rate (Mchips/s)	PG (chips/ ch. sym.)	Preamble Length (chips)	Packet Type	Info. Rate (kbps)	Packet Info content (bits)	Packet Duration
6	2	8	40	1000	Data	100	500	5.125 ms
6	2	8	40	1000	Control	100	50	625 μs
7	2	8	40	1000	Data	100	500	5.125 ms
7	2	8	40	1000	Control	100	50	625 μs
7	2	8	20	1000	Data	200	1000	5.125 ms
7	2	8	20	1000	Control	200	50	375 μs
7	2	8	10	1000	Data	400	2000	5.125 ms
7	2	8	10	1000	Control	400	50	250 μs
7	2	8	5	1000	Data	800	4000	5.125 ms
7	2	8	5	1000	Control	800	100	250 μs
7	2	8	1	1000	Data	4000	20,000	5.125 ms
7	2	8	1	1000	Control	4000	500	250 μs
8	4	4	20	500	Data	100	500	5.125 ms
8	4	4	20	500	Control	100	25	375 μs

Table 2.1 Transmission parameters used in the dissertation.

CHAPTER 3

BASELINE MAC PROTOCOL

In this chapter we describe the MAC protocol that we designed to meet the following objectives: accounting for the co-site limitation of the directional antenna system and exploiting the presence of directional antenna nodes to increase the spatial reuse of network resources. The MAC protocol is a generalization of the MAC protocol described in [16]. The generalization is intended to support the co-site operation of transceivers. It is referred to here as the *baseline MAC protocol*.

For each transmission of a network-layer data packet, a Ready-to-Send (RTS)/Clear-to-Send (CTS) packet exchange is employed on the control channel. The RTS packet is sent by the link-level source node/sector, and it advertises the traffic channels that are available at the source node/sector. If one or more of the advertised traffic channels are available at the intended destination node/sector, the destination selects one of the traffic channels and responses with a CTS that specifies that traffic channel. The link-level source and destination node/sectors then exchange a MAC-layer data packet and acknowledgment packet on the selected traffic channel. The traffic-channel exchange utilizes receiver-directed spreading codes, which results in a unicast link-layer transmission and also mitigates the effect of multiple-access interference with any other transmissions in the network that are using the same traffic channel concurrently. Failed transmission attempts result in an exponential back-off with respect to retransmission of the packet. The source node/sector and destination node/sector undergo *pacing* [16] after a traffic-channel exchange to account for their lack of current information about the state of the traffic channels at the end of the transmission.

Each CTS packet is transmitted using the common spreading code so that it can be overheard by idle third party node/sectors. (Each node/sector listens on the control channel when it is idle.) Each overhearing node treats the traffic channel specified in the CTS packet as unavailable for the duration of the corresponding data transmission. Thus the RTS/CTS mechanism serves to determine the availability of the intended link-level destination, determine if at least one traffic channel is available at both the link-level source and destination, negotiate the selection of such a traffic channel, and reserve the channel for exclusive use in the neighborhood of the destination. Since the RTS packet must advertise multiple candidate traffic channels in general, third-party nodes cannot use it to infer reservation of the particular traffic channel that is ultimately selected. Thus the RTS packet is not used for that purpose. Consequently, it is sent using a receiver-directed spreading code that prevents unnecessary acquisition of the RTS by third-party nodes.

Two techniques have been incorporated to exploit the ability of nodes with directional antennas to restrict directions in which they cause interference. Firstly, when a directional antenna node overhears a CTS, the advertised traffic channel is blocked *only* in the sector in which the overheard CTS is received. Secondly, a node with directional antennas builds a table called the *Neighbor-Sector (NS) table*. This table specifies the sector to be used to reach a particular neighbor and is built by simply observing the control and data packets being exchanged in the network. Suppose node A wants to send a packet to node B, and suppose that both nodes have directional antennas. Then node A uses the NS table to send the RTS only on the sector that is in the direction of node B, and node B sends the CTS only on the sector that is in the direction of node A. This restricts the area over which the RTS and CTS packets cause interference and the node/sector's which are blocked by the CTS. In addition, the data packet from node A to node B, and the ACK packet from node B to node A are also sent from only one sector. This restricts the region over which these packets lead to interference.

For each transmission of a network-layer control packet, a RTS packet is sent on the control channel using the common spreading code. The packet advertises a single traffic channel that is available at the link-level source node/sector, and the MAC-layer data packet containing the network-layer control packet is subsequently transmitted on the specified traffic channel using the common spreading code. Thus the transmission results in an unacknowledged link-level broadcast within the sector.

The network layer at each node includes a scheduler that maintains separate queues for network-layer control packets and network-layer data packets. If the node has multiple sectors, the control queue includes one copy of a (broadcast) control packet for each sector. In contrast, the data queue contains one copy of each (unicast) data packet, and each data packet is associated with the appropriate sector for its transmission. A dispatcher allocates packets to the MAC layer for transmission in each sector based on the constraints of both available traffic channels within each sector and the constraint of avoiding co-site interference.

In order to account for the co-site limitation, the nodes with directional antennas employ a scheduler and a channel-management module. When a packet reception is expected on a particular channel at any sector in such a node, the channel-management module blocks the traffic channel for transmissions in the other sectors at the same node. In addition to the availability of a sector, the scheduler takes into account both the unavailability of traffic channels due to overheard CTSs and due to the actions of the channel-management module while scheduling packet transmissions. A detailed description of the baseline MAC protocol, including a description of the data structures employed, the scheduler and the protocol state-machine description is given in Appendix A.

CHAPTER 4

LEAST RESISTANCE ROUTING PROTOCOL

The network topologies that we consider require in many instances that a packet traverse multiple links to reach its destination. Thus the network must employ a multiple-hop routing protocol, and the design of the protocol is a key factor in determining the network's performance. In all of the results in Chapters 6 and 7 of this dissertation, we consider a routing protocol referred to as *least-resistance routing (LRR)* which is based on the distributed Bellman-Ford distance-vector algorithm [17]. A modification of the LRR protocol is also used in Chapter 8 as a baseline routing protocol against which we compare the performance of the new routing protocol introduced in that chapter.

The LRR protocol was developed originally for frequency-hop spread-spectrum packet radio networks, and the initial design is described in [18, 19]. Many of the details of the protocol are based on features of the DAPRA SURAN packet radio network [20]. The version of the LRR protocol considered in this dissertation is based on a version previously introduced for use in a direct-sequence spread-spectrum packet radio network [21]. It differs from the version in [21] only with respect to the link metrics that are employed and in the fact that our version accounts for the possibility of multiple sectors at a node. The link metric described below is used with the LRR protocol throughout this dissertation, though it is supplemented in Chapter 8 by a second type of link metric (which is described in that chapter).

A nominal value is used for the metric of each link, and it is greater than one only if recent packet transmissions have failed on the link. Thus in the absence of link transmission failures, the path metric equals the number of links (hops) in the path. The LRR protocol consequently results approximately in distance-vector, minimum-hop routing, and it is referred to as a *min-hop* routing protocol in subsequent chapters.

Each node maintains a forwarding table with two entries associated each destination: the primary outgoing link and the secondary outgoing link for the destination. The primary link at the node with respect to a given destination corresponds to the outgoing link of the lowest-metric path to the destination. The secondary link at the node with respect to a given destination corresponds to the outgoing link of the lowest-metric path to the destination among paths that don't include the primary link. (See [18] for details.)

Transmission attempts for a data packet occur on the primary link or secondary link for the packet's destination, based on a forwarding algorithm, and the metric on each link is updated based on the outcome of the transmission attempts for the data packet. Each packet is transmitted by the node up to a maximum allowable number of attempts if necessary. Transmission is attempted first on the primary link, up to a specified number, and the remaining allowable attempts occur on the secondary link. (In the examples considered in the dissertation, a maximum of six transmission attempts are allowed, of which the first four occur on the primary link and the remaining two occur on the secondary link.) The metric of a link (primary or secondary) is incremented by a constant for each failed transmission attempt on the link, and the metric is reset to one after a successful transmission attempt on the link. (An increment of 0.2 is used in the examples in the dissertation.)

The network-layer protocol uses PROP packets [22] that are broadcast periodically at the link layer by each node to disseminate information to its neighbors. A node uses the PROP packets received from its neighbors to build the forwarding table. Each PROP packet contains a table entry for each node in the network that is known to the node broadcasting the PROP packet. Among the information included in the packet about each node is the *sector-to-use* field. If the field has a value of zero in the entry for a particular node, that node is not a neighbor of the node issuing the PROP. Otherwise, the node sending the PROP uses the value in the sector-to-use field to specify its own sector in which it receives transmissions from the node corresponding to that table entry. This information aids the node receiving the PROP packet in forming a picture of the network that includes the spatial discrimination currently provided by the directional antennas of its neighboring nodes.

CHAPTER 5

SIMULATION OF NETWORK PERFORMANCE

An event-driven simulation is used to investigate the performance of the MAC and routing protocols in a distributed direct-sequence packet radio network. Each transmitting node in the network generates data packets for each of its destination nodes according to a Poisson process. The network simulation includes accurate models of physical-layer performance, including models of packet acquisition and detection performance similar to those used in [16]. The model of acquisition performance is discussed in [23]. An upper bound on the packet error probability for hard-decision decoding [24] is used in the simulation, employing convolutional-code parameters from [25]. Multiple-access interference is modeled as Gaussian noise, and the directionality and antenna gain of the directional antennas are taken into account.

For a node with n sectors, the directional antenna that implements each sector has a beam pattern that is approximated as an ideal pattern covering $360/n$ degrees in the azimuth and results in an n-fold in-beam power gain. For a given total transmission power, the use of the directional antenna for transmission thus results in an n-fold gain in the in-beam power density at a given distance compared with the use of an omnidirectional antenna. The use of the directional antenna for reception of an in-beam signal results in an n-fold gain in the received signal power for a given signal power density at the receiver. Thus, if n_t is the number of sectors at the transmitter and n_r is the number of sectors at the receiver, equation 2.1 generalizes to

$$P_r = n_t \times n_r \times P_t \times \alpha \times \left(\frac{\lambda}{4\pi d}\right)^3 . \qquad (5.1)$$

CHAPTER 6

THE RECEIVER BLOCKING PROBLEM IN WIRELESS AD HOC NETWORKS WITH NODES USING DIRECTIONAL ANTENNAS

In this chapter we examine the performance of the baseline MAC protocol introduced in Chapter 3 and demonstrate that conventional RTS-CTS based protocols suffer from the *receiver blocking problem*. Using examples and geometric arguments, we show that while the problem also occurs when all the nodes in the network have omnidirectional antennas, it occurs with a higher probability when the nodes have directional antennas. We then present a simple solution to mitigate the problem and demonstrate the resulting improvements.

6.1 The Receiver Blocking Problem and Its Effect on Performance

In this section we illustrate the receiver blocking problem and demonstrate its impact on performance if the network includes directional antennas. Two eight-node networks are considered. Network I consists only of nodes with omnidirectional antennas. In network II, nodes 0-3 each have multiple directional antennas that form three sectors and nodes 4-7 have omnidirectional antennas. The two networks have the same node locations, which are shown in Fig. 6.1 along with the sector orientations of nodes 0-3 of network II. Either network employs only two frequency channels: the control channel and one traffic channel.

The distance between nodes 0 and 1 is 300 m while the distance between nodes 0 and 2 is 643.35 m. Nodes 4 and 5 are separated by a distance of 900 m, while nodes 4 and 6 are separated by a distance of 1243.5 m. By the symmetry of the network topology, the distances between all the other nodes can be calculated from these distances. Traffic is generated at node 0 at a rate of G packets/s, and it is all destined for node 1. Node 2 generates traffic for node 3 at the same rate. Nodes 4-7 each generate traffic at a rate of H/2 packets/s to each of their two nearest neighbors along the outer square, so that the net generation rate at each of the nodes 4-7 is H. (For instance, node 4 generates packets to nodes 5 and 6.) No other traffic is generated in the network. Routing is turned off in this example, and only link throughput is considered.

6.1.1 An Illustration of The Receiver Blocking Problem

The receiver blocking problem is illustrated by considering the throughput for traffic originating at node 0, which is referred to as simply the *throughput*. (By symmetry, the throughput for traffic originating at node 2 is the same.) The throughput is shown in Fig. 6.2 as a function of the generation rate G of nodes 0 and 2. Four values of the generation rate H at the interfering nodes 4-7 are considered. Examination of the network topology in Fig. 6.1 suggests that transmissions between node 0 and node 1 are decoupled from transmissions between node 2 and node 3 in network II. This in turn suggests that the throughput in network II should be approximately twice the throughput in network I if the generation rate G is large, since only one traffic channel is available. An approximate doubling is in fact observed under that condition if the generation rate of the interfering traffic is low. But as the level of interference increases, the throughput in network II degrades more severely than the throughput in network I. Indeed, if the rate of interference is sufficiently high, the throughput in network II is poorer than the throughput in network I. This is illustrated by considering the performance at H = 20 packets/s in Fig. 6.2.

This result appears counter-intuitive since one might expect that the improved interference rejection directional antennas give to nodes 0-3 in network II should provide them with greater immunity to transmissions by nodes 4-7. But it is understood by considering the information available at nodes 0 and 1 about interfering transmissions in network II. Every RTS-CTS exchange among nodes 4-7 results in a CTS that is overheard by either node 0 or node 1 in network II, but not both, with this topology. This results in a mismatch of information between node 0 and node 1 about the availability of the sole traffic channel.

For example, if an RTS is sent to node 4 and the sole traffic channel is available at that node, it responds with a CTS that reserves the traffic channel. The CTS is overheard in sector II at node 1 which marks the traffic channel as unavailable in that sector. But the CTS is not overheard in sector III at node 0. Thus if node 0 has a packet to send to node 1, it considers the traffic channel to be available and transmits a RTS into its sector III. But node 1 receives the RTS in its sector II and ignores it, since it has no available traffic channels in that sector. Thus node 0 initiates a (potentially lengthy) back-off that would not have occurred if node 0 had overheard the CTS in its sector III and refrained from transmitting the RTS to node 1. The back-off results in a loss in efficiency. By symmetry, the same phenomenon occurs with transmissions between nodes 2 and 3.

As the traffic rate at the interfering nodes increases, this phenomenon occurs with increasing frequency and the throughput is decreased accordingly. It is an instance of the receiver blocking problem, and in this example it arises as a result of the unique characteristics of a network with directional antennas, since no corresponding information mismatch occurs in network I. Note that the receiver blocking problem is distinct from the more familiar hidden-terminal problem (as well as the exposed-terminal problem) which can arise in random-access radio networks [26]. Indeed the receiver blocking problem is one consequence of the mechanism used to mitigate the hidden-terminal problem (e.g., the use of the CTS packet in this network).

6.1.2 Directional Antennas and the Receiver Blocking Problem

The presence of directional antennas increases the network's vulnerability to the receiver blocking problem, as illustrated by a simple example. Consider the link between a pair of nodes, A and B, and the occurrence of the receiver blocking problem for an intended transmission from node A to node B. The receiver blocking problem occurs if the receiver is blocked by an overheard CTS from a node that lies within the receiver's range but not within the transmitter's range. The severity of the receiver blocking problem is characterized by the probability that the problem occurs with an interferer that is placed randomly within the region of coverage of the relevant sector of node B. (The concept of "coverage areas" used in this subsection is for purposes of illustration only. All of the simulation results in the dissertation employ the accurate link model described in chapter 2.)

The situation if both nodes have omnidirectional antennas is shown in Fig. 6.3. The shaded area indicates the region in which an interferer that generates a CTS can be overheard by node B but not by node A - thus resulting in the receiver blocking problem at node A. The probability of occurrence of the receiver blocking problem is almost zero if node A and B are close to each other. The probability increases to 0.6089 as node B is moved towards the outer edge of node A's coverage range and the ratio of the inter-node distance to coverage range increases.

Suppose instead that node A and node B employ directional antennas and that each node lies along the boresight of a directional antenna at the other node. (I.e, the antennas' boresights are aligned.) This is shown in Fig. 6.4 for antenna beamwidths of 120^o. Again if an interferer in the shaded region generates a CTS, it can be overheard in the relevant sector at node B but not in the relevant sector at node A. The probability that the receiver blocking problem occurs is 0.17292 (with 120^o beamwidths) if nodes A and B are at the limits of each other's coverage range and the probability approaches one as node A approaches node B.

If instead nodes A and B employ directional antennas that face each other but are not perfectly aligned, as illustrated in Fig. 6.5, the probability that the receiver blocking problem occurs depends on the misalignment of the antenna patterns. For a given distance between A and B, the probability that the receiver blocking problem occurs increases as the boresight misalignment increases, and it approaches one as the boresight misalignment approaches the beamwidth. Thus if the pair of nodes with directional antennas are close or they have substantially misaligned boresights, the link between them is highly vulnerable to the receiver blocking problem. The problem is exacerbated if the beamwidths of the transmitting and receiving antennas are narrow. As the beamwidths employed by the two nodes are decreased towards zero, the probability that the receiver blocking problem occurs increases towards 0.5 even if the boresights are aligned and the inter-node distance equals the coverage range. (For example with a beamwidth of 60^o the probability that the receiver blocking problem occurs is 0.4485 when the inter-node distance equals the coverage range and increases towards one as the inter-node distance decreases towards zero.)

6.2 Mitigation of the Receiver Blocking Problem

6.2.1 Some Insight into Mitigation of the Problem

An idealized modification of the baseline MAC protocol suggests how the protocol might be modified in practice to reduce the frequency of occurrence of the receiver blocking problem. Specifically, consider an enhancement of the protocol by the inclusion of a "genie" which provides side information to a node/sector that has transmitted a RTS as the prelude to the transmission of a data packet. If the destination node/sector has no available traffic channels and one or more of the channels are blocked due to an overheard CTS, the genie informs the source node/sector of the time at which each of those traffic channels at the destination will become available. The source node/sector then employs the genie-provided information as if it had itself obtained the information by overhearing the corresponding CTS packets. If the baseline MAC protocol is used with the eight-node network II considered in the example above, for instance, the introduction of the genie substantially enhances link throughput. In particular, the throughput for data originating at node 0 is several-fold greater with the genie than without it under conditions of heavy traffic.

6.2.2 An Alternative Protocol: The NCTS MAC Protocol

The second MAC protocol we introduce in this chapter is a practical enhancement of the baseline MAC protocol that approximates the behavior of the genie described in the previous subsection. The fact that a node/sector has had all its traffic channels blocked as a result of overhearing CTS packets does not preclude it from transmitting on the control channel, and the enhanced protocol exploits that fact. If the node/sector receives a RTS on the control channel while all the requested traffic channels are blocked in this manner, in the enhanced protocol the node/sector responds with a *negative CTS (NCTS)* packet addressed to the sender of the RTS. (Hence we refer to the enhanced baseline MAC protocol as the *NCTS MAC protocol*.) (The term NCTS has also been used in [27], [28] and [29], but the information contained in the NCTS packet in these works differs from its use in our protocol).

The NCTS is transmitted on the control channel using the common spreading code, and thus it can be overheard by any listening node within range. Multiple NCTS packets may be transmitted by a node/sector during a given channel-blockage interval as a result of receiving a RTS packet from more than one neighbor. Each NCTS packet echoes the traffic list in the corresponding RTS, and it contains the time at which each traffic channel in the list will be free (unblocked). For each designated channel, the packet also includes the number of times the responding node/sector has transmitted a NCTS packet listing the channel since the last time the channel was free at the node/sector. (Thus in our protocol, the NCTS packet contains information that is not provided in the NCTS mechanism considered in [27] and [28].)

The contents of both the NCTS packets addressed to a node and the NCTS packets that are overheard by a node are used by a node to schedule subsequent transmission attempts to a blocked node/sector so that contention at the end the of the channel-blockage interval is limited. (Thus, the information in the NCTS packets is used in a different manner to schedule retransmissions compared to the protocol in [29]. In [29], the effect of contention at the end of the channel-blockage interval is not taken into account while scheduling retransmissions based on the information in the NCTS.) Note that the NCTS mechanism is feasible because the network employs a distinct frequency channel solely for MAC-layer control packets. A detailed description of the NCTS MAC protocol is given in Appendix B.

6.3 Comparison of Link-level Performance with the Baseline and NCTS MAC Protocols

The effect of introducing the NCTS mechanism into the MAC protocol is illustrated by examining the performance of the eight-node network II considered above. As in the example above, routing is turned off and performance is characterized in terms of link throughput only. In Fig. 6.6, the aggregate link throughput of network II is shown as a function of the generation rate of nodes 0 and 2 with both the baseline and NCTS MAC protocols. The total generation rate at each of nodes 4-7 equals the generation rate at nodes 0 and 2. The throughput with the two MAC protocols is approximately the same if the traffic generation rate is low. Under conditions of heavy load, however, the aggregate link throughput of the network is 165 packets/s with the NCTS protocol. That is an increase of 36.3% over the heavy-load aggregate link throughput of 121 packets/s with the baseline protocol. Clearly, the introduction of the negative NCTS control packet and the corresponding protocol results in a significant improvement in overall performance in this instance.

The NCTS protocol provides a greater aggregate link throughput due to its superior utilization of the capabilities of the nodes with directional antennas as illustrated in Fig. 6.7. Specifically, a heavy-load link throughput of 50 packets/s is achieved at each of nodes 0 and 2 with the NCTS protocol, whereas the throughput at either node is only 21 packets/s with the baseline protocol. The improvement of 138% in the link throughput at nodes 0 and 2 is obtained at the cost of a 17.7% decrease in the overall throughput at the remaining source nodes, however. The average of the link throughputs at nodes 4-7 is 19.75 packets/s with the baseline MAC protocol but only 16.25 packets/s with the NCTS MAC protocol.

In contrast, there is only a minimal difference in the performance with the baseline and NCTS MAC protocols if each of the eight nodes has an omni-directional antenna (i.e., if network I is considered). For network I, the introduction of the NCTS protocol increases the link throughput at nodes 0 and 2 by only 6% under a heavy load, and it has negligible impact on the link throughput at nodes 4-7. The heavy-load aggregate link throughput of the network is increased from 119 packets/s to 123 packet/s due to the NCTS mechanism - a difference of only 3%. Note, moreover, that the introduction of directional antennas at nodes 0-3 of the eight-node topology results in an increase of only 6% in the aggregate link throughput if the baseline MAC protocol is used, whereas the corresponding change improves the throughput by 34 % with the NCTS protocol. The example thus motivates investigation of the impact of the NCTS mechanism on both the aggregate performance and the per-data-flow performance for more interesting (i.e., larger) networks with a specific focus on networks that include nodes with directional antennas.

The link-level performance with the two MAC protocols is illustrated for a larger network by considering a network of thirty nodes with a randomly generated topology. Twenty of the nodes employ an omnidirectional antenna, and each of the remaining ten nodes uses directional antennas that form three sectors. The nodes are located in a region of dimension 4.5 km by 4.5 km, and the average network throughput is evaluated by simulation for a random placement of the nodes within the region. Each node generates packets for transmission to a single destination node, where the destination node is chosen at random among the nodes that lie at a distance between 500 m and 1.5 km from the source node. Once again, routing is turned off and performance is characterized in terms of link throughput only.

The aggregate link throughput of the network is shown in Fig. 6.8 for the two MAC protocols, where the throughput is averaged over six randomly generated topologies for the thirty-node network. In each of the six network topologies, each of the nodes has at least one neighbor within the required range of distances so that each node generates traffic. The throughput is 386 packets/s under a heavy load if the baseline MAC protocol is used, and it is 440 packets/s if the NCTS MAC protocol is used. Thus the NCTS MAC protocol results in a 14.4% improvement in aggregate link throughput in comparison with the baseline MAC protocol. Moreover, the NCTS protocol results in a higher aggregate throughput for each of the six network topologies, and the percentage improvement over the baseline MAC protocol ranges between 10.4% and 19.5% for the six topologies. As with the eight-node network, the introduction of the NCTS mechanism is of particular benefit to the nodes that experience a severe receiver blocking problem with the baseline protocol. Among the links for which at least 20% of the RTS packets are transmitted to a node/sector that has no traffic channels available, the use of the NCTS protocol results in an average increase in throughput of 38%, as illustrated in Fig. 6.9. In contrast, the link throughput averaged over the remaining links is essentially the same with either MAC protocol. If the ten nodes with directional antennas are replaced by nodes with omni-directional antennas, the NCTS protocol results in an increase in aggregate throughput of only 7.4% over the baseline protocol.

6.4 Comparison of Network-Level Performance with the Baseline and NCTS MAC Protocols

In this section we consider the performance with the baseline and NCTS MAC protocols for a 30-node network in which multiple-hop routing is enabled. The network dimensions and the mix of node capabilities are the same as in the last example of the previous section. The performance measures of interest are the aggregate end-to-end throughput of the network, the distribution of end-to-end throughputs achieved for the individual network-layer source-destination pairs, and the percentage of generated packets in the network that successfully arrive at the destination (i.e., the *completion percentage* for the network). The completion percentage is an important network performance characteristic for many applications, and it can have a substantial impact on the performance of a transport-layer connection. Thus the throughput under a constraint on the completion percentage serves as one fair measure of the network performance under different conditions. Each performance result in this section is an average over the same six randomly generated topologies considered in the last example of the previous section.

Three traffic scenarios are considered in the examples. In each scenario, the generation rate is the same for the Poisson source of each active source-destination pair in the network. In *traffic scenario one*, each node generates packets for one destination node that is chosen randomly among the nodes that lie between 500 m and 1.5 km from the source node. (Note that a given node may be the destination of more than one source node.) The source-destination pairs correspond to those considered in the last example of the previous section, and at any given time the routing algorithm is likely to have selected single-hop routes for most pairs. Thus scenario one results in a network in which most of the traffic is local, a circumstance that arises in many applications of ad hoc networks.

The performance of the network in traffic scenario one is shown in Fig. 6.10 with both the baseline MAC protocol and the NCTS MAC protocol. The aggregate end-to-end network throughput and the completion percentage of the network are shown as functions of the aggregate generation rate for the network. The baseline MAC protocol achieves a completion percentage of 90% with a throughput of 288 packets/s. But the NCTS protocol achieves a 90% completion percentage with a throughput of 326 packets/s, which is a 13% increase in throughput compared with the baseline protocol. If the constraint on completion percentage is tightened to 95%, the throughput with the NCTS protocol is 275 packets/s, which is 16% greater than the throughput of 237 packets/s with the baseline MAC protocol.

Suppose instead that each node generates packets for one destination node that is chosen at random from among all the other nodes in the network. This reflects a circumstance in which traffic is less localized than in traffic scenario one, and it is denoted *traffic scenario two*. At any given time, a significant fraction of the source-destination pairs in this scenario are likely to employ multiple-hop routes. The performance of the network in traffic scenario two is shown in Fig. 6.11 with both MAC protocols. The NCTS protocol results in an aggregate network throughput of 194 pkts/s at a completion rate of 90%, whereas the corresponding throughput is only 179 pkts/s with the baseline protocol. If the completion is 95%, the throughput for the NCTS and baseline MAC protocols are 176 packet/s and 163 packet/s, respectively. Thus the NCTS protocol results in a throughput that is 9% greater and 8% greater than with the baseline protocol for the two respective completion percentages.

The fact that the performance improvement provided by the NCTS protocol is greater in traffic scenario one than in traffic scenario two is understood by considering a node/sector that is susceptible to the receiver blocking problem on the links to some, but not all, of its neighbors. If the baseline MAC protocol is employed and the node/sector undergoes a back-off interval with respect to one of its links due to the receiver blocking problem, the node/sector can still utilize the link to another neighbor during the interval if it has data to send to that neighbor. Thus as the number of data flows in which a typical node participates is increased, under-utilization of the local channel capacity due to the receiver blocking problem thus becomes less significant factor in limiting network performance. The potential for performance improvement with the NCTS protocol is correspondingly smaller, as occurs in the progression from traffic scenario one to traffic scenario two.

This point is illustrated further by considering *traffic scenario three* in which each node generates traffic at the same rate for every other node in the network. Traffic scenario three represents an extreme of randomized traffic flows, and most node/sectors participate in a larger number of data flows. The performance of the network in traffic scenario three is shown in Fig. 6.12 with both MAC protocols. We see that the baseline protocol and the NCTS protocol achieve a completion percentage of 90% with respective throughputs of 204 packets/s and 209 packets/s in this traffic scenario - a difference of about 2%. If the completion percentage is 95%, the corresponding throughputs are 185 packets/s and 186 packets/s for both - a difference of less than 1%.

Not only does the NCTS protocol result in better aggregate network performance than the baseline protocol, but the improved performance is widely shared among the network-level source-destination pairs in the network. This is illustrated by considering two performance measures that are derived from the cumulative distribution function (c.d.f.) of the end-to-end throughputs achieved by individual source-destination pairs. The empirical c.d.f. with either MAC protocol is shown in Figures 6.13, 6.14, and 6.15 for traffic scenarios one, two, and three, respectively, and a network completion percentage of 90%. Each result is an average over thirty randomly generated topologies for the 30-node network, rather than the six topologies considered in the discussion above, resulting in 900 samples for each empirical c.d.f. (Since the generation rate is the same for each active source-destination pair in the network, the distribution of completion rates for individual source-destination pairs can be determined immediately from the network completion rate and the distribution of end-to-end throughputs.)

One measure of the per-pair performance is the throughput at the n-th percentile for a specified value of n. It is apparent from Figures 6.13 and 6.14 that the NCTS protocol results in better performance than the baseline protocol with respect to this performance measure, regardless of the value of n. For example, 90% of the source-destination pairs achieve a throughput of at least 7.4 packet/s with the baseline protocol in traffic scenario one. With the NCTS protocol, however, 90% of the pairs achieve a throughput of at least 8.55 packets/s. Thus the NCTS protocol results in 16% better performance than the baseline protocol with respect to this measure. The most-favored 95% of the source-destination pairs achieve a minimum throughput of only 6.4 packets/s with the baseline protocol, but they achieve a minimum throughput of only 7.7 packets/s with the NCTS protocol (a difference of 22%). In traffic scenario two, the NCTS protocol results in 6% and 13% performance improvements over the baseline protocol under the same two respective criteria. The difference in the performance of the two protocols in traffic scenario three is negligible under either of the two criteria.

Another measure of per-pair performance is based on a specified level of service in terms of the end-to-end throughput. For example, suppose that a throughput of seven packets/s is required to provide acceptable service for a particular application. In traffic scenario one, 6.4% of the source-destination pairs fail to receive acceptable service with the baseline MAC protocol but only 4% of them fail to receive acceptable service with the NCTS protocol. In traffic scenario two, 83% do not receive acceptable service with the baseline protocol, but percentage is reduced to 63% with the NCTS protocol. Suppose instead that a throughput of eight packets/s is required. Then the baseline and NCTS protocols do not result in acceptable service for 14% and 6.4% of the source-destination pairs, respectively, in traffic scenario one. In traffic scenario two, only 4% of the pairs receive acceptable service with the baseline protocol, whereas 10% are still able to receive acceptable service with the NCTS protocol. Once again, the difference in the performance of the two protocols in traffic scenario three is negligible under either of the two criteria.

The improvement in overall network performance provided by the NCTS protocol is thus enjoyed comparably by nodes that are in an advantageous circumstance and those that are in a disadvantageous circumstance with respect to factors other than their susceptibility to the receiver blocking problem. While this does not mean that each data flow necessarily achieves greater throughput with the NCTS protocol than with the baseline protocol, it demonstrates that the NCTS protocol yields better performance in terms per-data-flow statistical criteria. In this respect, the NCTS MAC protocol results in universally superior network-wide performance to that obtained with the baseline MAC protocol.

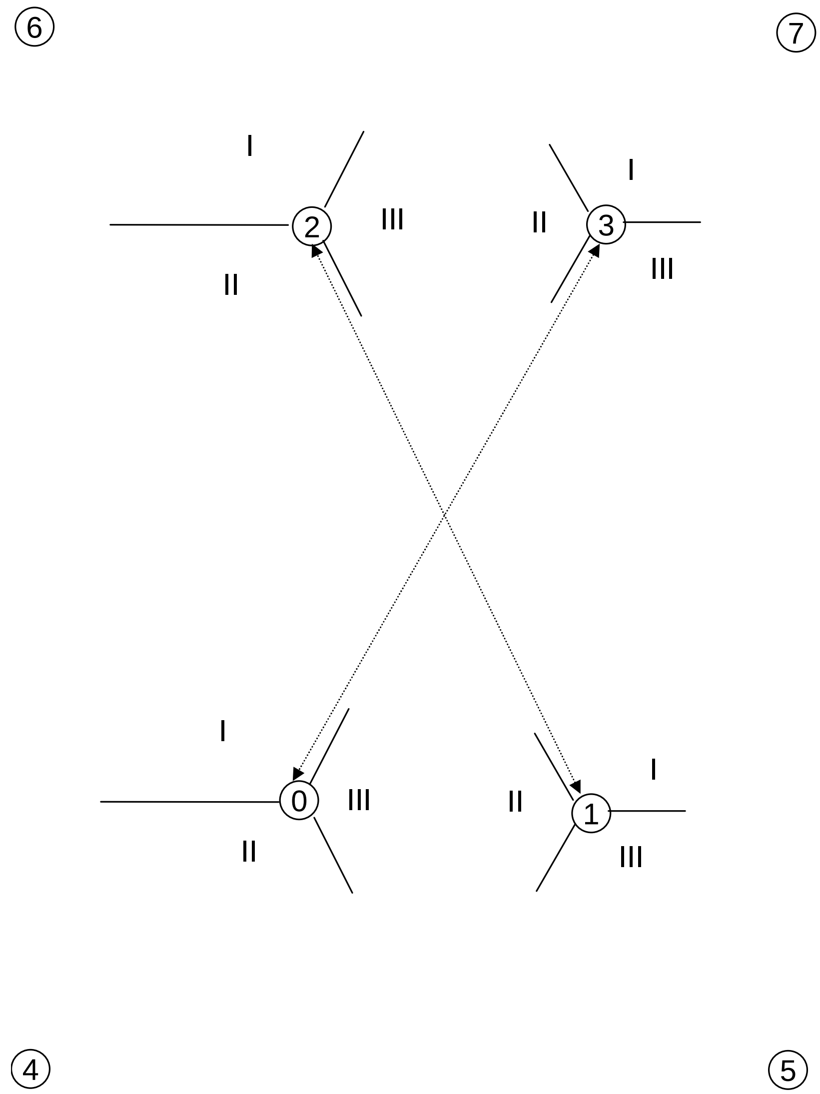

Figure 6.1 Network topology.

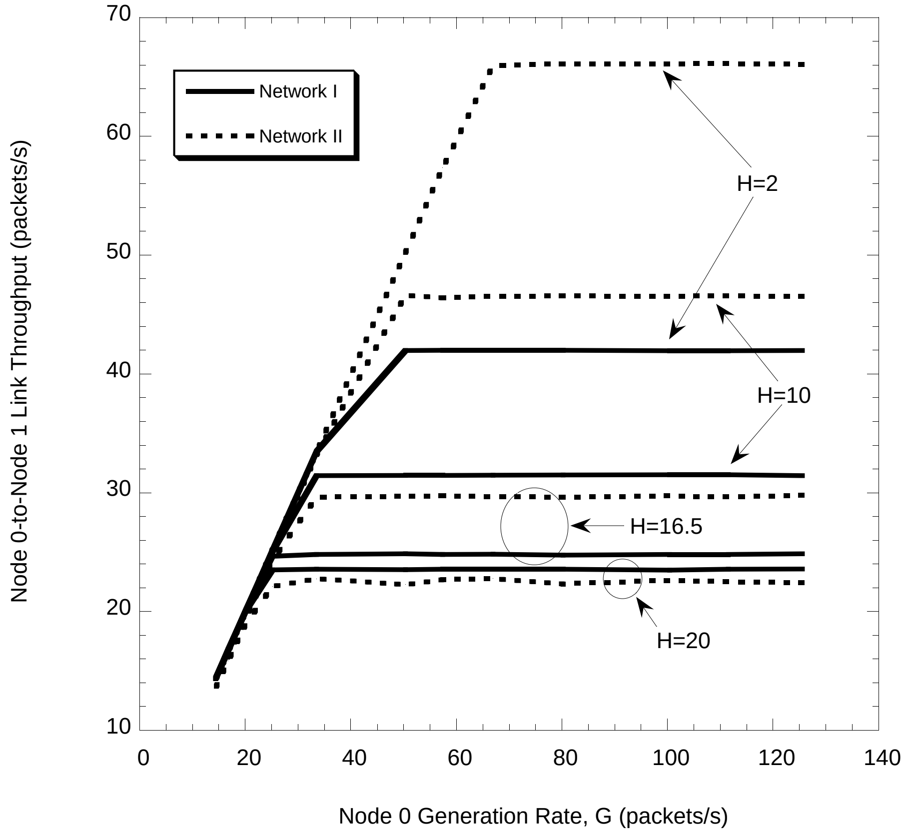

Figure 6.2 Throughput of node 0 for networks I and II with various generation rates H at the interfering nodes.

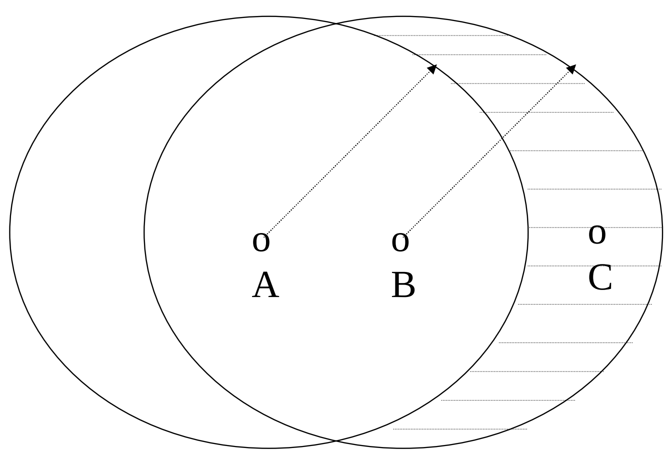

Figure 6.3 Illustration of the receiver blocking problem if each node has omnidirectional antennas.

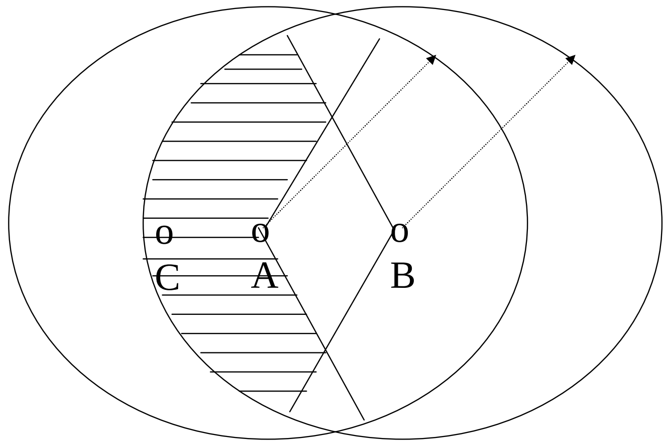

Figure 6.4 Illustration of the receiver blocking problem if transmitter A and receiver B have directional antennas.

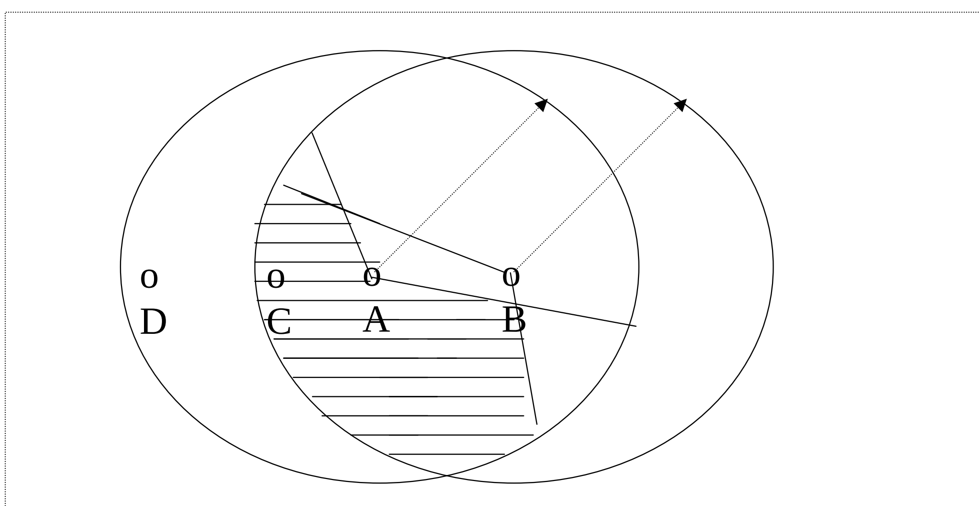

Figure 6.5 Illustration of the receiver blocking problem if both transmitter and receiver have directional antennas with imperfect overlap between sectors.

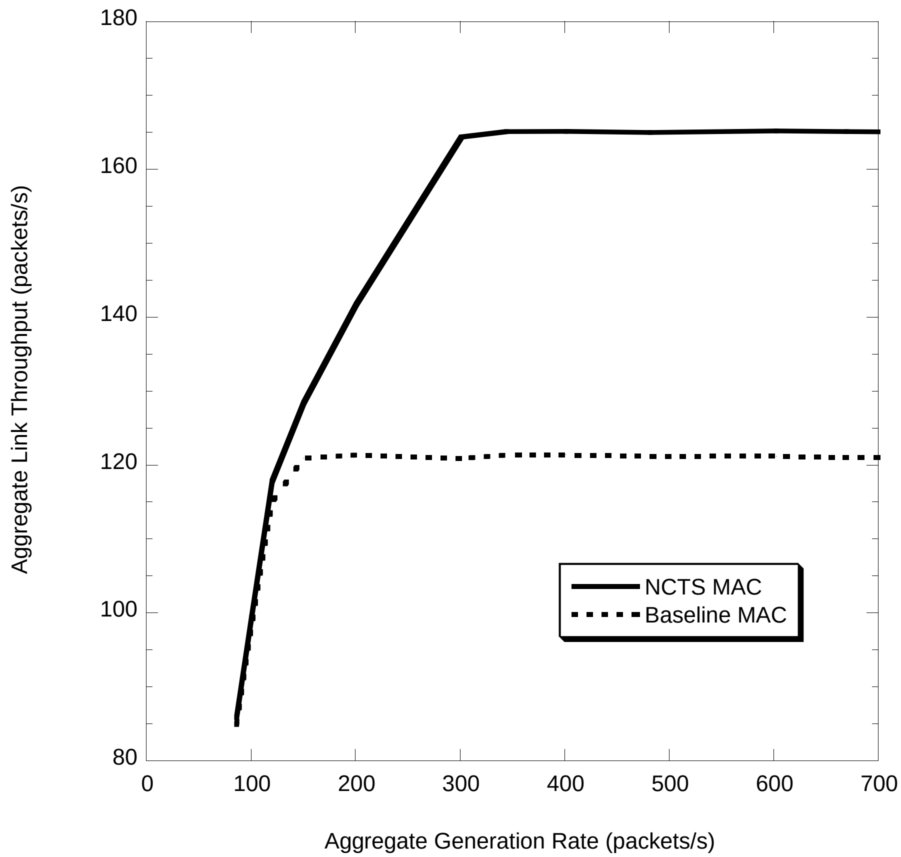

Figure 6.6 Aggregate link throughput of network II with both MAC protocols.

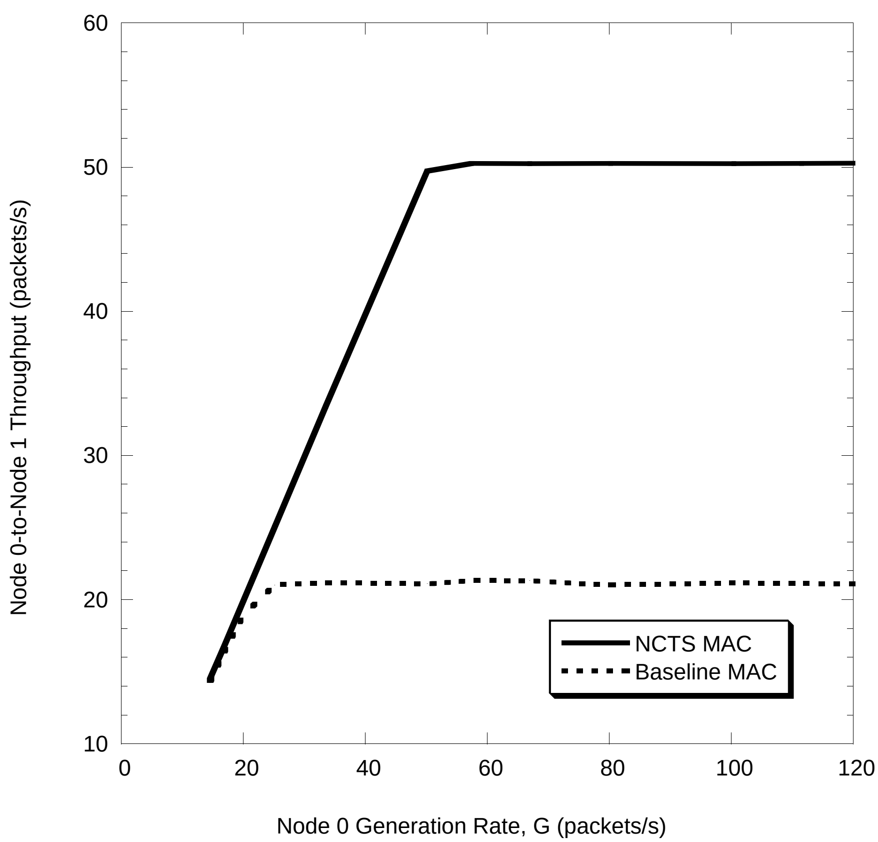

Figure 6.7 Throughput of node 0 of network II with both MAC protocols.

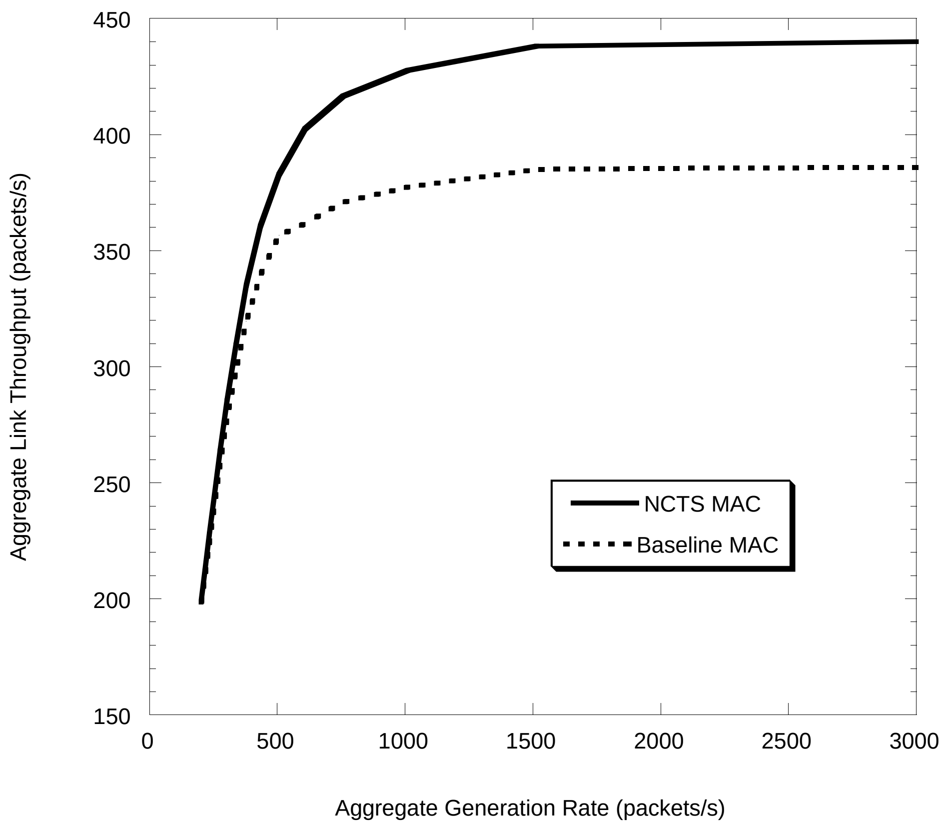

Figure 6.8 Aggregate link throughput for the 30-node network with both MAC protocols.

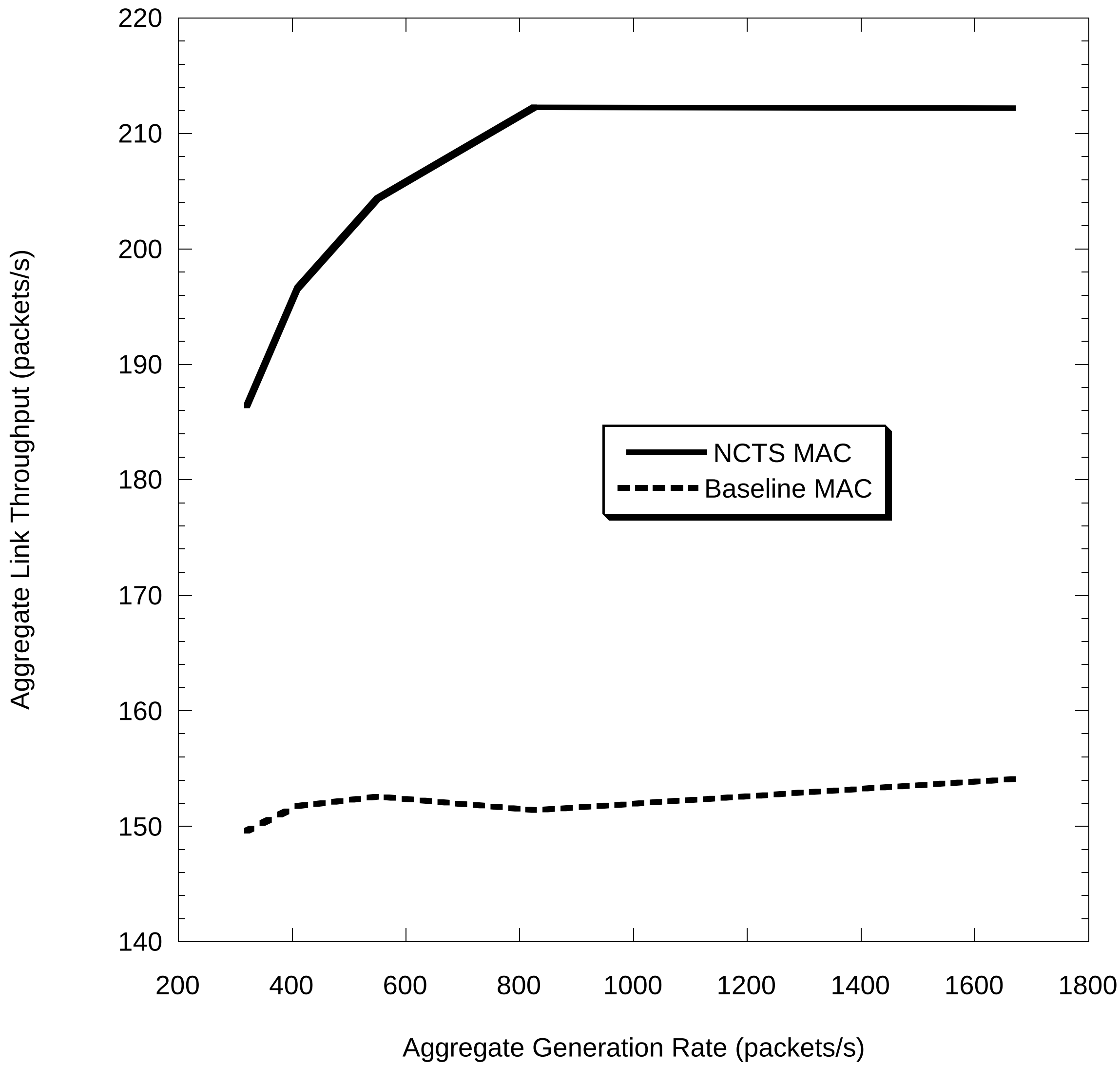

Figure 6.9 Aggregate throughput of the links that are affected by the receiver blocking problem at least 20% of the time for the 30-node network with both MAC protocols.

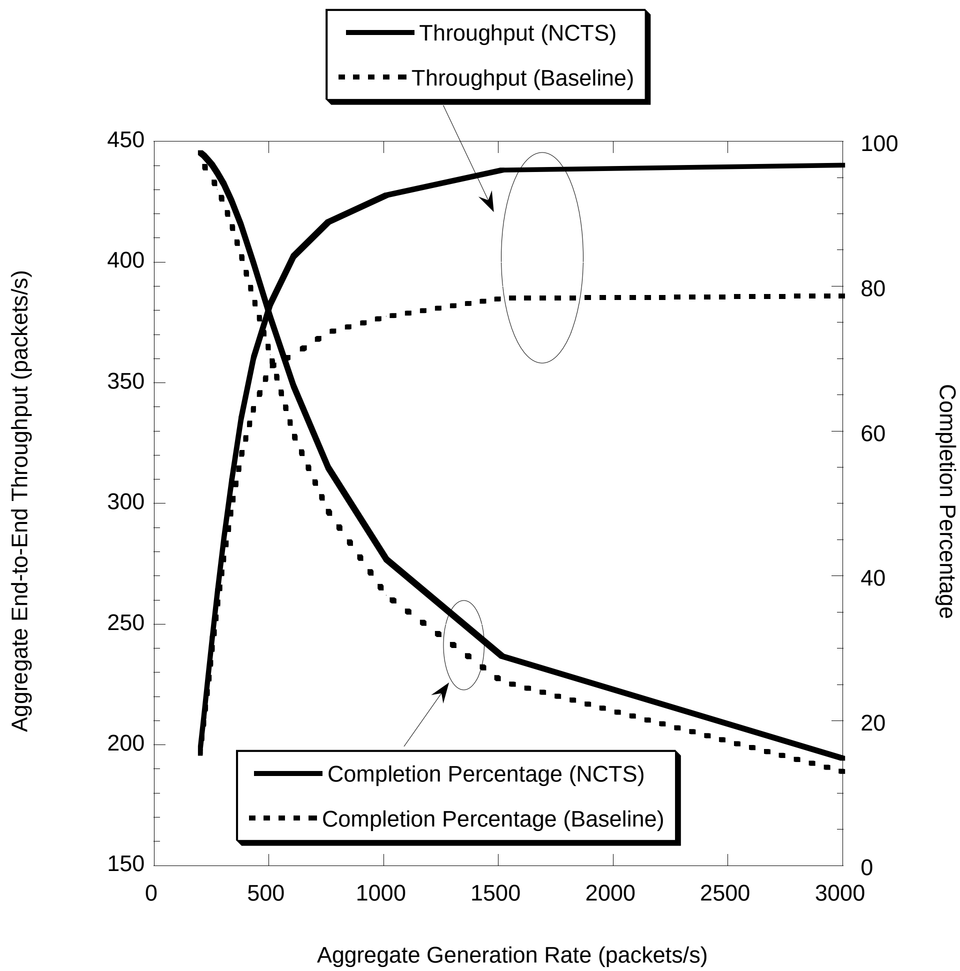

Figure 6.10 Completion percentage and aggregate end-to-end throughput for the 30-node network with both MAC protocols for traffic scenario one.

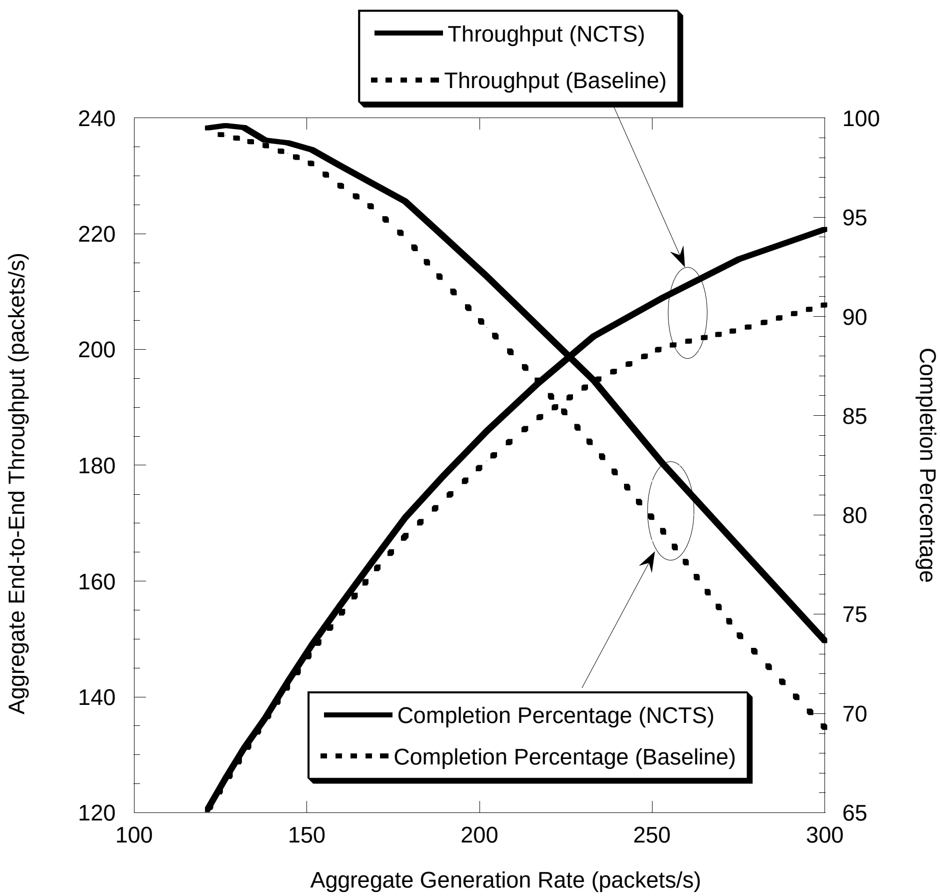

Figure 6.11 Completion percentage and aggregate end-to-end throughput for the 30-node network with both MAC protocols for traffic scenario two.

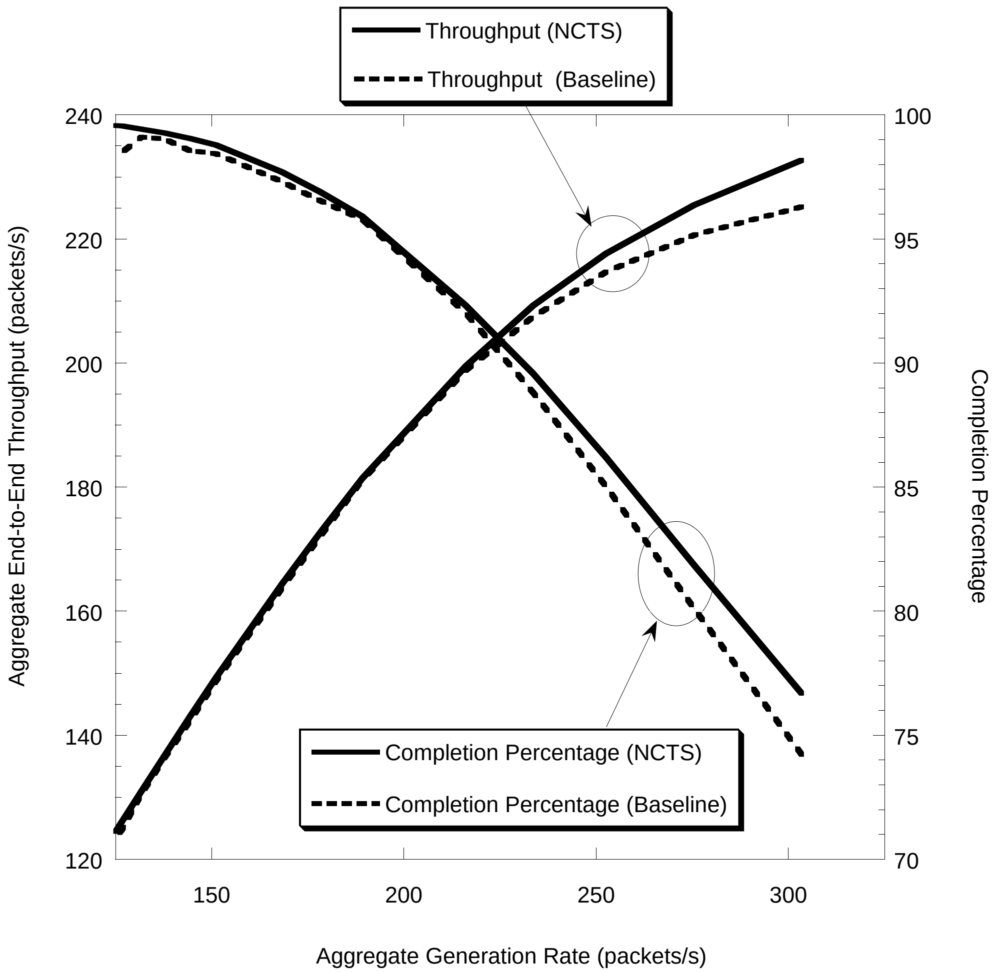

Figure 6.12 Completion percentage and aggregate end-to-end throughput for the 30-node network with both MAC protocols for traffic scenario three.

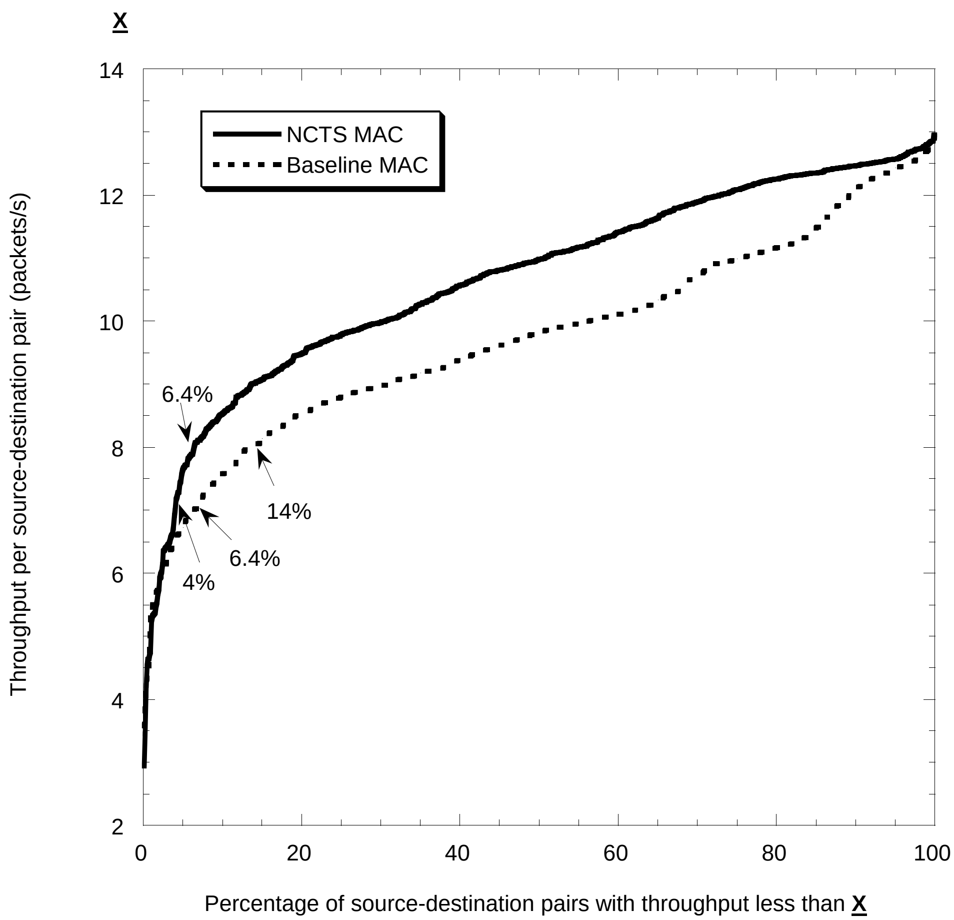

Figure 6.13 Empirical c.d.f of per-flow throughputs for thirty 30-node networks with traffic scenario one.

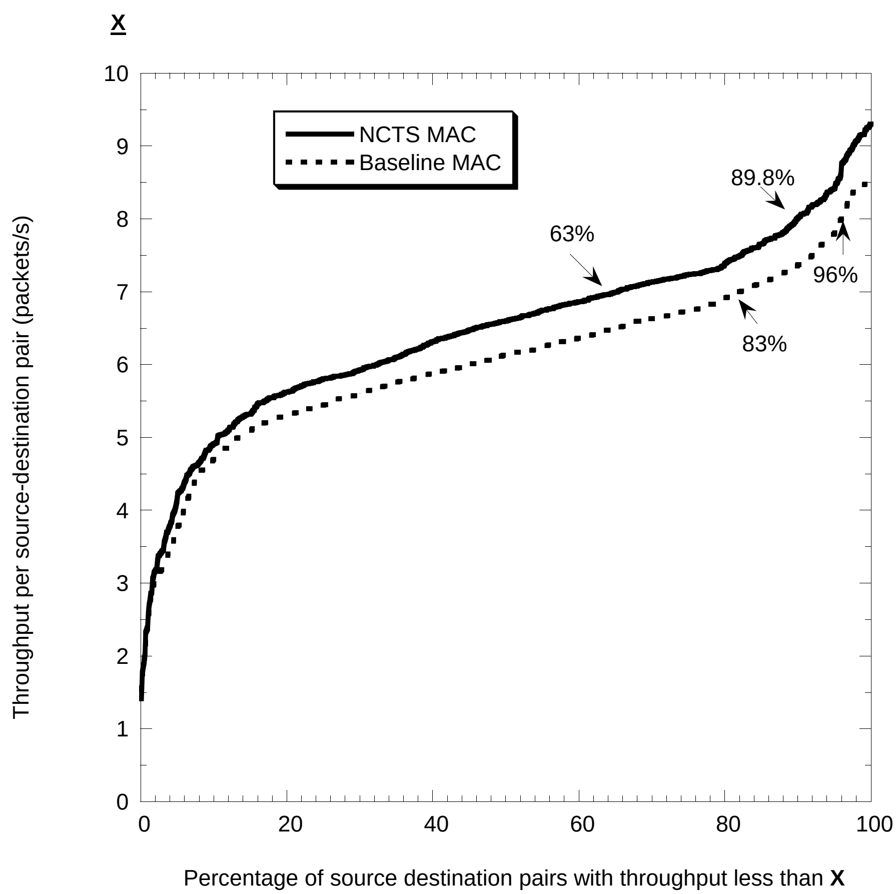

Figure 6.14 Empirical c.d.f of per-flow throughputs for thirty 30-node networks with traffic scenario two.

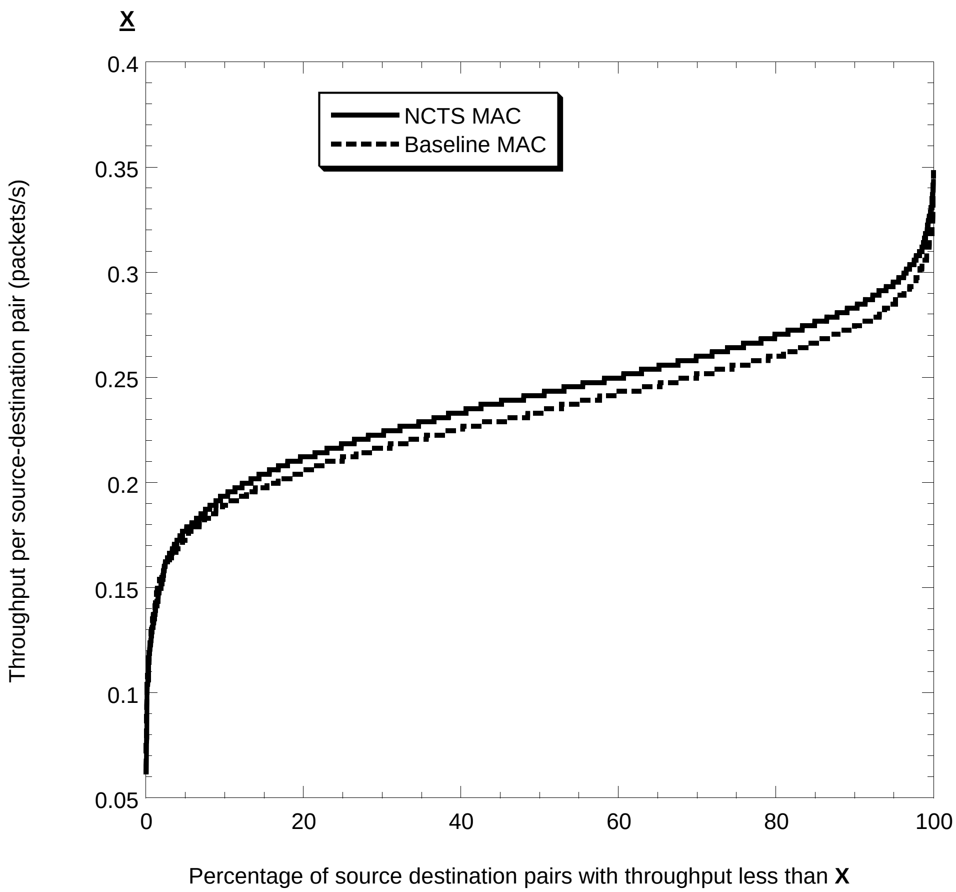

Figure 6.15 Empirical c.d.f of per-flow throughputs for thirty 30-node networks with traffic scenario three.

CHAPTER 7

TECHNIQUES TO IMPROVE SPATIAL REUSE OF TRAFFIC CHANNELS IN WIRELESS AD HOC NETWORKS

Any partially connected packet radio network is susceptible to the hidden-terminal problem, and much of the development of channel-access protocols for ad hoc packet radio networks has focused on protocols that are designed to mitigate the hidden-terminal problem through short-term reservation of a traffic channel within the local area of a transmitter or receiver [30]. This approach is typified by the use of Ready-to-Send (RTS) and Clear-to-Send (CTS) control packets as a mechanism for local channel reservation [12].

Exclusive local reservation of a communication channel in an ad hoc network can often result in overly conservative utilization of the channel, however. This phenomenon is especially pronounced in a network using DS spread-spectrum modulation, since the modulation format provides each receiver with the opportunity to achieve a processing gain against sources of multiple-access interference. Yet the modulation format does not eliminate the susceptibility of a receiver to the near-far interference problem [5]. Thus the design of the channel-access protocol for a DS spread-spectrum packet radio network can be viewed as driven by two counter-acting objectives: efficient spatial reuse of each traffic channel, and adequate protection against multiple-access interference for each transmission. The first of these objectives is sacrificed implicitly in favor of the second in the design of most reservation-based protocols.

Several approaches have been proposed recently for achieving efficient spatial reuse of traffic channels in an ad hoc network through adaptive control of transmission power [31]. Among the approaches are the use of a dedicated control ("busy tone") channel in which each receiver advertises the excess link margin it currently enjoys [32] and the advertisement of similar information in CTS control packets [33]. In either approach, nearby potential interfering transmitters adjust their power for subsequent transmissions in an attempt to ensure that acceptable link quality is preserved for the advertising receivers. Other approaches to power control in ad hoc networks require a degree of centralized control [34]. While each provides some useful insight into the design of ad hoc networks for efficient spatial reuse of traffic channels, most proposed designs fail to account for one or more important practical limitations of tactical packet radio networks, such as restriction to half-duplex radio operation, the vulnerability of link-level acknowledgement packets to reception error, and the need to avoid centralized control in order to ensure network robustness.

In this chapter, we describe several variants of a channel-access protocol for DS spread-spectrum packet radio networks that is designed to exploit available opportunities for spatial reuse of traffic channels while limiting concurrent transmissions that might create excessive multiple-access interference at any receiving radio. The protocol is designed to work effectively in the environment and within the limitations of practical packet radio networks. In particular, it is designed for a network containing an arbitrary mix of nodes with half-duplex radios, some employing omnidirectional antennas and other employing directional antennas. In addition, each node obtains information about the state of the network by using only quality metrics it derives from detection of packets and information contained in the payload of control packets exchanged as part of the protocol. The protocol is an extension of the baseline MAC protocol described in Chapter 3.

It is shown that the new protocol, designed with the twin objectives of spatial reuse efficiency and limitation of multiple-access interference, achieves significantly better performance than the more conservative approach of using the RTS/CTS exchange to provide exclusive local reservation of a channel. It is also shown that the new protocol results in better performance than a protocol that permits unrestricted concurrent use of a channel.

7.1 A Motivating Illustration

The tension between the design objectives of efficient spatial reuse and limitation of multiple-access interference is illustrated by considering the simple four-node network shown in Figure 7.1. Suppose the network employs a channel-access protocol with two frequency channels: a control channel dedicated to the exchange of RTS and CTS control packets, and a traffic channel dedicated to the exchange of data and acknowledgement (ACK) packets. An RTS/CTS exchange serves the purpose of determining and securing the availability of the intended link-level destination for an upcoming data packet transmission. Moreover, if the protocol is designed to reserve the traffic channel, the RTS/CTS exchange serves to prevent other listening neighbors from using the traffic channel during the upcoming data transmission and acknowledgement interval.

Now consider the circumstance in which node 1 has data queued for transmission to node 2 and node 4 has data queued for transmission to node 3. (The distance d between either communicating node pair is assumed to be small enough to allow communication in the absence of multiple-access interference.) Suppose at first that the distance X between nodes 2 and 3 is small. Reception at node 2 of a data transmission from node 1 is somewhat vulnerable to multiple-access interference from a data transmission by node 4 and even more vulnerable to interference from an acknowledgement transmission by node 3. A complementary vulnerability exists for reception by node 3 of a data transmission from node 4. Thus it is desirable in this circumstance that the RTS/CTS exchange reserves the traffic channel for exclusive use by one of the two node pairs at a time in order to limit the effects of multiple-access interference.

Now suppose that the distance X is increased. Beyond a certain distance, node 3 will no longer be able to detect (overhear) a CTS transmitted by node 2 to node 1 and the channel-reservation property of the protocol is irrelevant. Moreover, beyond another certain (likely different) distance, the level of multiple-access interference that arises at node 2 due to transmissions from node 3 or node 4 is sufficiently small that it has negligible impact on the probability of detection error. In a network employ DS spread-spectrum modulation with a significant processing gain, the latter distance is likely to be much smaller than the former distance. Thus there is a range of conditions in which node 3 can overhear a CTS transmitted by node 2 to node 1 (and node 2 can overhear a CTS from node 3 to node 4) but in which the two node pairs can engage in concurrent data transmissions without suffering from severe multiple-access interference. It is desirable in this circumstance that the RTS/CTS exchange between either node pair does *not* reserve exclusive use of the traffic channel, thus permitting more efficient spatial reuse.

This example illustrates the limitations of a channel-access protocol that is designed to limit multiple-access interference by securing exclusive channel reservations in all conditions. Likewise, it illustrates the limitations of a protocol that is designed to achieve efficient spatial reuse by allow concurrent channel use under all conditions. Thus the example motivates the desirability of a channel-access protocol that adaptively regulates local concurrent use of a traffic channel to maximize overall network performance.

7.2 Signal-Quality Metrics Derived from Received Packets

For the protocols we introduce in this chapter, signal-quality measurements are derived as part of the detection of each PROP packet (see Chapter 4) at a receiver. These measurements may be employed along with measurements from the reception of other types of packets in order to generate link metrics for use in the distributed routing protocol. These signal-quality measurements form the basis of the adaptive channel-access protocols described in this chapter. Any of several practical techniques for determining signal-quality in a DS spread-spectrum packet radio network can be employed [35]. In this dissertation, we use the power in the desired component (the signal of interest) of the received signal as the measure of signal quality, and we employ the approximating assumption that this quantity is measured with complete accuracy by the receiver for each detected PROP packet.

7.3 Additional Data Structures Maintained at each Node

Each node in the network maintains two data structures for each of its sectors that are used with the channel-access protocols described in the next section. The *overheard CTS table (OCTS table)* at each node/sector contains one entry for each CTS transmission that has been overheard recently by the node. (Entries in the table are removed as they expire.) Each CTS packet contains the identifiers of the node/sector transmitting the CTS (the CTS node) and the node/sector to which the CTS is responding (the RTS node), the CTS node's current measure of signal quality for transmissions it receives from the RTS node, and the traffic channel selected for use for the subsequent data/acknowledgement exchange. All of this information in the overheard CTS is stored as an entry in the node's OCTS table along with the time stamp designating when the data/acknowledgement exchange resulting from that CTS will be completed.

Each node/sector also maintains a table called the *signal-quality table*. Each entry in the signal-quality table consists of measurements of signal quality that have been obtained during reception of PROP packets from the node's neighbors. There is one entry for each neighbor/sector and each traffic channel. Each entry may be expunged due to age using the same criterion used for entries in the routing table at the node. (As discussed above, in this dissertation we consider measurements of received signal power as the signal-quality measurements stored in the OCTS table and the signal-quality table.)

7.4 Reference Channel-Access Protocols and the Proposed Channel-Access Protocol

In this section we describe two channel-access protocol that employ a static channel-reuse scheme. Following this, we describe three variations of a channel-access protocol that employs an adaptive channel-reuse scheme designed to improve the spatial reuse of traffic channels in a DS spread-spectrum ad hoc packet radio network. Each channel-access protocol that is considered is used in conjunction with the local packet scheduler at each node described in Chapter 3. The packet scheduler governs the order of transmission attempts for outgoing packets at the node and schedules each transmission for the appropriate sector in the nodes that employ multiple directional antennas.

7.4.1 Exclusive-Use Channel-Access Protocol

The baseline MAC protocol described in detail in Chapter 3 and Appendix A is referred to as *the exclusive-use channel-access protocol* in this chapter. This terminology is used to emphasize that in the baseline MAC protocol, unlike the other protocols described in this chapter, a node overhearing a CTS always blocks the traffic channel advertised in the CTS in the sector in which the CTS was overheard. In addition, if a node/sector receives an RTS packet when all the traffic channels advertised in the RTS are blocked at the receiving node/sector, it rejects (ignores) the RTS.

7.4.2 Unrestricted-Reuse Channel-Access Protocol

The *unrestricted-reuse channel-access protocol* uses the RTS-CTS control-channel dialog only to reserve the attention of the node/sector for the subsequent data transmission and to determine the traffic channel on which the data/acknowledgement exchange will occur. It is not used to reserve the traffic channel. Thus nodes ignore overheard CTS packets, and it responds to each RTS packet it receives by selecting one of the advertised traffic channels and sending a CTS packet that designates the selected channel.

7.4.3 Selective-Reuse I Channel-Access Protocol

One variant of the new protocol is referred to as the *Selective-Reuse I (SR I)* channel-access protocol. The SR I protocol is identical to the exclusive-use protocol in the situation in which a node/sector receives an RTS advertising one or more traffic channels that are not blocked at the receiving node/sector. If instead all the advertised traffic channels are blocked, the SR I protocol uses the signal-quality information contained in the OCTS and signal-quality tables described in the previous section. The use of this information is illustrated by considering the following scenario.

Suppose that some traffic channels are currently blocked by one or more overheard CTS packets at node A/sector S and that the OCTS table for that node and sector has the entries shown in Table I. (The received signal power measured at node y due to a transmission from node x is denoted P_{xy}.) Furthermore, suppose an RTS then arrives at node A/sector S from node B/sector S_1 and that all the channels advertised in the RTS are blocked at node A/sector S. Rather than automatically rejecting (ignoring) the RTS, as would occur with the exclusive-use protocol, node A employs the following procedure to determine if any of the blocked traffic channels can be used in spite of the presence of one or more interfering transmissions. For each traffic channel, the maximum total interference power that the intended data transmission will suffer at node A/sector S is estimated, and it is divided by the estimated power with which the intended data transmission will be received. This ratio is compared against a threshold for each traffic channel. If the ratio is exceeded for all traffic channels, the RTS is ignored. But if not, one of the traffic channels passing the test is selected for use and a modified version of a CTS packet (referred to as a *Reuse CTS (RCTS)* packet) is transmitted to node B/sector S_1 on the control channel, designating the selected traffic channel. (A node/sector overhearing an RCTS packet treats it in the same manner as any other overheard CTS packet with respect to traffic-channel blocking and entries in its OCTS table.)

The total interference power for each traffic channel is estimated by node A from the corresponding entries in the OCTS table for its sector S. For example, there is one entry associated with traffic channel one in the OCTS table shown in Table I. Based on this entry, node A extracts the quantity P_{DA} from the entry for node

D/sector 2 in node A's signal-quality table for its sector S. Similarly, the signal-quality table is searched for an entry for node C/sector 1. If one exists, the quantity P_{CA} is extracted, and if not, P_{CA} is set to zero. The maximum interference at node A/sector S due to the data transmission reflected in the first entry of the OCTS table is given by the maximum of P_{CA} and P_{DA}, since the data and acknowledgement transmissions are not concurrent. This maximum is determined for each entry in the OCTS table associated with traffic channel one, and the sum of these maxima yields node A's estimate of the maximum total interference power P_I at its receiver for sector S. The quantity P_{BA} is extracted from the signal-quality table, and the ratio P_I/P_{BA} is compared with the threshold $MaxTolerableInt$. (If the signal-quality table lacks an entry for either P_{DA} or P_{BA} for the traffic channel, the test is treated as a failure for that traffic channel since node A lacks timely, relevant signal-quality information for a link that is known to exist.) The same calculations and comparison are performed for each candidate traffic channel to determine which one, if any, can be used. (In our current implementation, if two or more candidate traffic channels pass the test, the one with the smallest maximum total interference power is selected.)

If the result of these tests is the selection of a traffic channel by node A, the RCTS packet it transmits includes the information associated with the selected traffic channel contained in the OCTS table for node A/sector S. For the example, this corresponds to the first row shown in Table I if traffic channel one is selected. (In our implementation of the protocol, a traffic channel is not selected if there are more than three corresponding entries in the OCTS table. This serves to limit the size of the RCTS packet.) Node B determines if it can transmit on the selected traffic channel without causing excessive interference to any of the ongoing transmissions corresponding to the entries in the RCTS packet it has received, and its decision is made using its signal-quality table for sector S_1 and the entries in the RCTS.

If node B receives a "standard" CTS from node A, it responds as in the exclusive-use protocol. But if node B receives an RCTS from node A, it determines the prob-

RTS node	CTS node	Power	Time valid	Channel
node C/sector 1	node D/sector 2	P_{CD}	T_1	1
node E/sector 1	node F/sector 2	P_{EF}	T_2	2

Table 7.1 Example of OCTS table at node A/Sector S.

ability with which it will transmit a data packet on the specified traffic channel by determining the impact its transmission would have on each of the ongoing transmissions. This is illustrated by considering its procedure for an RCTS packet entry corresponding to the first OCTS table entry shown in Table I. The quantities P_{CB} and P_{DB} are extracted by node B from the signal-quality table for sector S_1, and the quantity P_{CD} is taken from the received RCTS packet. (If there is no entry for P_{CB} or P_{DB}, the quantity is set to zero.) Node B then determines the probability $p_{CD} = \min\{f(P_{CB}/P_{CD}), f(P_{DB}/P_{CD})\}$. In our implementation of the protocol, the function $f(x)$ is given by

$$f(x) = \begin{cases} 1, & x \leq MinRCTS \\ \frac{(MaxRCTS-x)}{(MaxRCTS-MinRCTS)} & MinRCTS < x < MaxRCTS \\ 0, & x \geq MaxRCTS \end{cases} \quad (7.1)$$

where *MinRCTS* and *MaxRCTS* ($MinRCTS \leq MaxRCTS$) are constants. This is designed to limit interference with either the data transmission from node C to node D or the acknowledgement transmission from node D to node C by providing protection in some instances in which node B has previously failed to overhear the CTS from node D to node C even though one or both is within range. (Note that this implementation of the function implicitly treats each link as exhibiting symmetric signal attenuation for a given traffic channel. Asymmetric attenuation can be accounted for more accurately by modifying the protocol so that each node/sector includes its signal-quality table in its PROP packets.) The probability that node B transmits a data packet is the minimum of the probabilities determined for the entries in the received RCTS packet.

There are two circumstances in which the SR I protocol may result in traffic-channel reuse that would have been precluded with the exclusive-use protocol. The first occurs if the node/sector initiating a transmission has previously failed to overhear a CTS from a neighbor due to either a detection error or busy-ness of the node/sector. The second occurs if the receiver-blocking problem [36] arises with respect to the desired transmission. In the illustration shown in Fig. 1, for example, suppose a transmission from node 1 to node 2 is ongoing on a given traffic channel. A "reuse decision" concerning a possible concurrent transmission from node 4 to node 3 on the same channel arises only if node 3 has received the recent CTS from node 2 but node 4 has not. It is shown in [36] that the receiver blocking problem is exacerbated by the presence of nodes in the network that have directional antennas, and thus we would expect the SR I protocol to show the greatest benefit for such a network. The results shown later bear out this expectation.

7.4.4 Selective-Reuse II Channel-Access Protocol

We also consider a second variant of the new channel-access protocol, referred to as the *selective-reuse II (SR II)* protocol, that follows a more aggressive approach to traffic-channel reuse than the SR I protocol. This is achieved by allowing a node/sector with outgoing data to initiate a data transmission under some conditions in which all the traffic channels would have been blocked at that node/sector by overheard CTS packets using the exclusive-use or SR I protocol. Under the SR II protocol if a node/sector overhears a CTS packet, it determines whether or not it should treat the traffic channel designated in the overheard CTS as blocked for transmission based on a probability distribution. Using the signal-quality information in the overheard CTS and in the local node/sector signal-quality table, the probability is generated using the function given in (7.1) except that the constants *MinOCTS* and *MaxOCTS* are used in place of *MinRCTS* and *MaxRCTS*, respectively. With this probability, the traffic channel designated in the overheard CTS is not blocked from inclusion in the list of channels advertised in a subsequent RTS transmission by the node. Even for this outcome, however, the node retains the information in the overheard CTS, and it follows the same procedure as for the SR I protocol if it is the recipient of an RTS or the addressed recipient of a "standard" CTS or an RCTS.

7.4.5 Selective-Reuse III Channel-Access Protocol

We also consider a third variant of the new channel-access protocol, referred to as the *selective-reuse III (SR III)* protocol. In this variant, we adopt the approach of the SR II protocol in that a node with data to send is allowed to transmit an RTS packet that advertises traffic channels that are blocked at the node under the conditions specified in the previous subsection. However, in the SR III protocol the recipient of the RTS packet follows the approach of the exclusive-use channel-access protocol. That is, it ignores the received RTS packet if all the advertised traffic channels are blocked at the receiving node. Similarly, the node that has sent the RTS packet follows the approach of the exclusive protocol when it receives the responding CTS packet. The approach used in the SR III protocol allows reuse of traffic channels under some conditions without requiring the increased control-packet size that is required for the RCTS packets in the SR I and SR II protocols.

7.5 Performance Comparison of the Channel-Access Protocols

The performance of each of the protocols described in the previous section is evaluated by determining the total throughput of all links of a network (referred to simply as the *throughput*), except where otherwise noted in Section 7.5.3. Each network considered in the examples uses one control channel and one traffic channel. The results in Sections (7.5.1)-(7.5.4) address link-level performance only (with no multiple-hop routing). Heavy traffic conditions are considered so that every active link has an infinite backlog of data queued at the corresponding source node. Multiple-hop routing is considered in the results of Section 7.5.5 for sources with a finite generation rate.

Protocol	Heterogeneous Network	All-Omnis Network
Exclusive Use	469.69	333.3
SR-I	550.68	390.69
SR-II	735.9	522.7
SR-III	691.5	479.97

Table 7.2 Throughput (kbits/s) for example in Section 7.5.1.

7.5.1 An Illustrative Example

In this subsection and the next subsection, the performance of each protocol is evaluated for six specific network topologies that were constructed using randomly generated node placement. The results are given in terms of the average throughput over the six topologies. Each of the six network topologies consists of thirty nodes. Two scenarios are considered. In scenario one, twenty of the nodes have an omni-directional antenna, while the other ten nodes each contain three-sector directional antennas. In scenario two, all the thirty nodes have omni-directional antennas. The nodes are placed in a 4500m x 4500m square region. Each node is a source of data for at least two and at most five link destinations. A particular node's destinations are chosen at random from among all its neighbors in the range 500m to 1500m, and the same set of active links is used in all examples for a given network topology.

The throughput is shown in the first column of Table 7.2 for the first scenario (with 10 directional antenna nodes) and in the second column for scenario two. The processing gain employed is 40 and the performance of each of the four channel-access protocols: the exclusive-use, SR I, SR II and SR III protocols is given. The three selective-reuse protocols and their corresponding parameter values are denoted by SR I (u, v, w), SR II (u, v, w, x, y) and SR III (x, y), where u=MinRCTS, v=MaxRCTS, w=MaxTolerableInt, x=MinOCTS, and y=MaxOCTS. In this example, SR I (1,2,5), SR I (1,2,5,1,2) and SR III (1,2) are considered. These represent a reasonable choice of values, though they are not necessarily optimal with respect to any particular operating conditions.

Noticeably better performance results for the heterogeneous network with the judicious traffic-channel reuse of the latter three protocols than with the exclusive-use protocol. The throughput with the SR I protocol is 550.7 packets/s, which is an 17.2% improvement over the throughput of 469.7 packet/s obtained with the exclusive-use protocol. The SR II protocol results in a throughput of 736 packets/s, which represents a 56.7% increase in throughput over the exclusive-use protocol. The SR III protocol results in a throughput of 691.5 packets/s.

The second column of Table 7.2 gives the throughput if all 30 nodes in each network employ omnidirectional antennas. With each of the four protocols the throughput is significantly lower if none of the nodes have directional antennas, but the benefits of the three selective-reuse protocols over the exclusive-use protocol are preserved. For instance, the SR II and SR III protocols provide respective gains of 57% and 44% over the exclusive-use protocol. Though the form of traffic-channel reuse employed in the SR I protocol appears to be of modest benefit, it is clear from this example that judicious traffic-channel reuse at the initiative of the source node (as used in the SR II and SR III protocols) is of greater benefit. Thus subsequent discussion of the reuse protocols is limited to the SR II and SR III protocols.

7.5.2 Comparison for Different Values of Processing Gain

The best choice of the processing gain in the design of a DS spread-spectrum packet radio network depends on numerous factors, and different networks may use greatly differing processing gains. Moreover, many future networks are likely to employ an adaptive transmission protocol in which transmission parameters including the processing gain are adapted in response to perceived link conditions [35]. In this subsection we will compare the performance of the different protocols over a wide range of processing gains in order to investigate the robustness of the selective reuse protocols with respect to that parameter of the DS transmission format.

For each network topology that is examined, we consider the operation of the network under five different channel propagation conditions which are characterized by different values of the power-gain constant α in the path-loss equation (5.1). Values of α of one, two, four, eight, and 40 are considered. The five values of α represent five different propagation environments ranging from the one exhibiting the most severe signal attenuation to the one exhibiting the least severe attenuation, respectively, with a corresponding impact on the quality of the link between a given pair of nodes in the network.

The performance of the network in each propagation environment is considered for a packet format using a processing gain (and a corresponding information rate) that is determined by the power-gain constant of the channels in the network. The processing gains are 40, 20, 10, 5, and one, and the corresponding information rates are 100 kbits/s, 200 kbits/s, 400 kbits/s, 800 kbits/s, and 4 Mbits/s, respectively. If the network is operating in an environment in which the channels have a power-gain constant of one, the highest processing gain of 40 (and the lowest information rate) is used. The processing gains of 20, 10, 5, and one are used in environments in which the network's channels exhibit power-gain constants of two, four, eight, and 40, respectively. Consequently, the signal-to-noise ratio for a bit of information is the same for the link between a given pair of nodes regardless of the value of the path-gain constant.

The value of considering the network performance under these different conditions is two-fold. First, it permits an evaluation of the relative performance of the various MAC protocols in DS networks with differing processing gains. Second, it is a first step in examining the performance that would result with the various MAC protocols in a network employing an adaptive processing gain to compensate for changes in link quality. Of course in a realistic environment, different conditions will exist on different links at any given time so that an adaptive-transmission protocol will not necessarily select the same transmission parameters for use on all links at that time. In contrast, in each circumstance we consider, all links in the network employ the same processing gain at a given time. But the examination can nonetheless provide some useful insights concerning the more general situation.

A large processing gain results in a reduced susceptibility to multiple-access interference compared with a small processing gain for concurrent transmissions in the same frequency channel. Thus one would expect the degree of aggressiveness in traffic-channel reuse that is most appropriate for the channel-access protocol to differ depending on the processing gain. This expectation is borne out in the result illustrated in Table 7.3 in which the aggregate link throughput of the network is shown in kbits/s as a function of the processing gain for several channel-access protocols which include the exclusive-use protocol and the unrestricted-reuse protocol. The throughput is averaged over the six randomly generated, 30-node network topologies with 10 directional antenna nodes each as described in Section 7.5.1. (Results for the same topology are also shown in Fig. 7.2 for processing gains between 5 and 40.) Transmissions are highly vulnerable to multiple-access interference if the processing gain is small, which dictates a conservative approach to traffic-channel reuse. This is seen with a processing gain of 5, for which a higher throughput of 3632 kbits/s results with the exclusive-use protocol than with the unrestricted-reuse protocol which results in a throughput of 3210 kbits/s. We can also see that the vulnerability of the unrestricted reuse protocol to multiple-access interference is most apparent in a narrow-band system (i.e., with a processing gain of one), in which the throughput with the exclusive-use channel-access protocol is 14,024 kbits/s while that with the unrestricted reuse protocol is only 7031 kbits/s.

Transmissions are much less vulnerable to multiple-access interference if the processing gain is large, which dictates a more aggressive approach to traffic-channel reuse. This is illustrated by considering the performance with the exclusive-use and unrestricted-reuse protocols as the processing gain is increased. If the processing gain is 10 or greater, the unrestricted-reuse protocol results in better performance. For instance at a processing gain of 40, the exclusive-use protocol results in 470 kbits/s

Processing Gain	Exclusive Use	Unrestricted Reuse	SR-III (1,2)	SR-III (2, 4)	SR-III (0.2, 0.5)	SR-III (0.7, 1.5)
40	469.69	700.30	691.95	723.58	612.11	677.68
20	967.27	1271.7	1456.8	1506.8	1297.2	1429.8
10	1930.7	2106.6	2850.0	2850.9	2594.2	2814.5
5	3632.4	3210.4	4861.8	4607.4	4731.8	4904.8
1	14,024.0	7,031.2	12,010.0	10,961.0	14,288.0	12,414.0

Table 7.3 Aggregate link throughput (kbits/s) with the three protocols averaged over six 30-node networks at different processing gains.

while the unrestricted reuse protocol results in a throughput of 700 kbits/s. Moreover, the performance of the unrestricted-reuse protocol relative to the performance of the exclusive-use protocol improves consistently as the processing gain is increased.

Table 7.3 also shows the performance of the SR III protocol for four different combinations of parameter values. In contrast with the static traffic-channel reuse strategies of the exclusive-use and unrestricted-reuse protocols, the SR III protocol makes dynamic reuse decisions based on local conditions at the time of each transmission. Thus it is able to exploit reuse opportunities when they are present without creating excessive multiple-access interference. This is illustrated in the figure, where it is seen that for each combination of parameters, the SR III protocol results in high throughput regardless of the processing gain. In particular, at a processing gain of 5, the SR III protocol with parameters $MinOCTS = 1$ and $MaxOCTS = 2$, results in a 34% improvement over the exclusive-use protocol and a 51.5% improvement over the unrestricted-reuse protocol. The SR III protocol with the same parameters results in a 47.7% improvement over the exclusive-use protocol and a 36% improvement over the unrestricted reuse protocol at a processing gain of 10. At a high processing gain of 40, SR III(1, 2) performs the same as the unrestricted reuse protocol and 49% better than the exclusive-use protocol. Even in the narrow-band system (PG=1), the throughput obtained with the SR-III (0.2, 0.5) protocol is greater than that with the exclusive-use protocol.

Thus, there is at least one combination of parameter values (SR III (2,4)) for which the SR III protocol results in throughput that is greater than the throughput of the unrestricted-reuse protocol over the full range of processing gains. Similarly, there is at least one combination of parameter values (SR III (0.2, 0.5)) for which the SR III protocol results in a throughput that is greater than the throughput of the exclusive-use protocol over the full range of processing gains. (We can also see that the performance with the SR III (2,4) protocol is better than that with either the exclusive-use or the unrestricted-reuse protocol over the range of processing gains 40-5.)

A similar comparison is illustrated in Fig. 7.3 with the SR II protocol used in place of the SR III protocol. Three combinations of parameter values are considered for the SR II protocol. The SR II protocol exhibits performance gains relative to the exclusive-use and unrestricted-reuse protocols that are similar to those exhibited by the SR III protocol. Indeed for all of the combinations of parameter values, the performance of the SR II protocol is strictly superior to that of the exclusive-use and unrestricted-reuse protocols over the full range of processing gains. In particular, the SR III (1.5,3,1.5,3,5,1,2) results in 28% greater throughput than the exclusive-use protocol at the lowest processing gain and 8% than the unrestricted-reuse protocol at the highest processing gain. The SR II (0.5, 1, 0.5, 1, 5, 0.5, 1) protocol results in performance that is slightly better than or equal to any of the three variants of the SR III protocol over the full range of processing gains. The performance gains that result with the SR II protocol are not much higher when compared to that obtained with the SR III protocol, however. As a result, since the SR II protocol is more complicated than the SR III protocol due to the additional tests and the use of the RCTS packets, we restrict our attention to the simpler SR III protocol in the remainder of the chapter.

Finally, the results in Table 7.4 depict the performance with the exclusive-use, unrestricted reuse and SR III protocols with the same six 30-node network topologies

Processing Gain	Exclusive-Use	Unrestricted-Reuse	SR-III (1, 2)	SR-III (2, 4)
40	333.3	465.6	479.97	523.3
20	671.63	702.27	999.7	1067.6
10	1332.3	939.45	1909.4	1895.7
5	2476.6	1166.8	3021.3	2692

Table 7.4 Aggregate link throughput (kbits/s) with the three protocols averaged over six all-omni 30-node networks.

when all the 30 nodes have omni-directional antennas. We see that the general trends are similar to the results for the heterogeneous network. However, the absolute values of the performance improvements that result with the SR III protocol are greater in the all omni-directional node case. In addition, the unrestricted reuse protocol performs very poorly (200% worse than the best performing SR III and 100% worse than the exclusive-use protocol at PG=5) at the lower processing gains.

From these results it is clear that either selective-reuse protocol provides greater robustness with respect to the processing gain than either the exclusive-use or unrestricted-reuse protocol. Thus dynamic traffic-channel reuse can prove especially beneficial if the processing gain in the network is held constant during operation but is treated as a system parameter that can be specified at the time of network configuration. The results also suggest that dynamic traffic-channel reuse would be highly beneficial in a network that employs an adaptive transmission protocol in which the processing gain differs for different transmissions.

7.5.3 Comparison for Different Network Densities

Some tactical communication scenarios are likely to result in a network with a low density of nodes, whereas other scenarios may result in a high network density. Thus an important measure of the robustness of a channel-access protocol is its ability to support good performance over a range of network densities. In this subsection, the performance of the exclusive-use protocol, the unrestricted-reuse protocol, and the SR III protocol are examined in two settings in which the density of nodes in the network is varied. For each of the selective-reuse protocols, parameter values are chosen that result in robust performance.

For each setting, the area containing the network is fixed. The initial network topology for a setting is obtained by randomly generated placement of ten nodes (the *nodes of interest*) of which three have three directional antennas and the rest have omnidirectional antennas. Each of the nodes of interest is a source of data for at least two and at most four link destinations, all of which are also among the nodes of interest. In addition, a specified number of additional nodes (the *interfering nodes*) are placed randomly. Each of these also has between two and four link destinations, all of which are other interfering nodes. If the number of interfering nodes is 5, 10, 15 or 20, the number with three directional antennas is 2, 3, 5 and 7, respectively. For these examples, the throughput is defined as the sum of the link throughputs for links involving the ten nodes of interest. The throughput with each protocol is considered as the number of interfering nodes in the network (and hence, the network density) is varied. The processing gain is 20 in each instance.

The performance of the four channel-access protocols in shown in Fig. 7.4 for a setting with dimensions of 1250m by 1250m. Each source-destination pair is separated by a distance of between 350m and 1200m. The throughput with each protocol is shown as a function of the number of interfering nodes in the network. In this setting, the unrestricted-reuse protocol results in the poorest performance regardless of the network density. The SR III protocol results in a greater throughput than either the unrestricted-reuse or exclusive-use protocol.

Note that in this setting, the performance of the exclusive-use protocol degrades more markedly than the performance of the unrestricted-reuse protocol with increasing network density. As the network density is increased, the unrestricted-reuse protocol is affected primarily due to an increase in the level of multiple-access interference at the receiving nodes of interest. In contrast, the exclusive-use protocol is affected primarily due to an increase in the contention for reservation of the traffic channel, which reduces the availability of the channel to the nodes of interest. It appears that though multiple-access interference is the more deleterious of the two phenomena in this setting, the relative impact of contention increases as the density increases. Thus the SR III protocol is effective both in limiting occurrences of excessive multiple-access interference and in exploiting available traffic-channel reuse opportunities even as demand for the channel increases.

The performance of the four channel-access protocols in shown in Fig. 7.5 for a setting with dimensions of 2500m by 2500m. Each source-destination pair is separated by a distance of between 300m and 1400m. In this second setting, the exclusive-use protocol results in the poorest performance regardless of the network density. Once again, the SR III protocol results in a greater throughput than either the unrestricted-reuse or exclusive-use protocol.

It is seen from these two examples that even for a fixed processing gain, the better choice of a static reuse strategy (represented by the exclusive-use and unrestricted-reuse protocols) can differ depending upon the operational setting of the network. Yet a channel-access protocol that implements a well-designed dynamic reuse strategy can provide an added degree of robustness with respect to the setting. The result is an improvement in network performance.

7.5.4 Comparison for a Clustered Network Topology

The nodes in each of the previous examples have been placed randomly, but many tactical communication scenarios lead to a network consisting of several clusters of closely spaced nodes. In this subsection, the performance of the exclusive-use, unrestricted-reuse, SR II and SR III channel-access protocols are considered for a network with a clustered topology. The network consists of three clusters of ten nodes each. Three of the nodes in each cluster have three directional antennas and the rest have omnidirectional antennas. Each cluster lies within a separate area of

Processing Gain	Exclusive-Use	Unrestricted-Reuse	SR III (1, 2)	SR III (2, 4)
40	355.83	568.32	554.89	581.36
20	710.48	927.75	1137.6	1178.1
10	1396.8	1419.1	2188.2	2202.5
5	2624.0	2095.6	3836.5	3593.3

Table 7.5 Aggregate link throughput (kbits/s) with a clustered topology.

dimensions 1200m by 1200m, and the centers of the three areas lie on the vertices of an equilateral triangle with edge length 1800m. The nodes of the three clusters are placed at random within their respective areas. Each node generates traffic for four other randomly chosen nodes within the same cluster.

The performance of the three channel-access protocols is shown in Table 7.5 as a function of the processing gain. The technique used to vary the processing gain is the same as that in Section 7.5.2. Once again the general trends are similar to that seen in that section. Comparing the performance obtained with the exclusive-use and unrestricted reuse protocols shows that the former outperforms the latter at PG=5, but the latter begins to perform progressively better at the higher processing gains. The performance with the SR III (1, 2) protocol results in a 56%, 60.2%, 56.7% and 46.2% over the exclusive protocol at PG=40, 20, 10 and 5 respectively. These improvements are higher than those obtained in the non-clustered topology in section 7.5.2. At PG=40, the unrestricted reuse protocol performs 2% than the SR III (1,2) protocol, while the SR III (1,2) protocol outperforms the unrestricted reuse protocol by 27%, 55.2% and 71.4% at PG=20, 10 and 5 respectively. The SR III (2, 4) protocol outperforms the unrestricted reuse protocol over the entire range of processing gains.

7.5.5 Performance with Multiple-Hop Routing

In this section we consider the performance with the three protocols for a 30-node network in which the destination nodes are possibly multiple hops away. The performance will be averaged over six instances of the 30-node network, with each instance containing a random placement of the nodes. In addition, in each instance, ten of the 30 nodes have 3-sector directional antennas. Packet generation at each node is modeled as an independent Poisson process with mean rate of R packets/s, and each packet is assigned a destination chosen at random among all the other nodes in the network. The routing protocol employed is least-resistance routing (LRR).

The performance measures of interest are the *end-to-end completion percentage* and the *average delay* with each of the three protocols. The end-to-end completion percentage is the percentage of packets generated in the network that successfully reach their final destination. The delay experienced by a successfully received packet is the difference between the time at which the packet reaches its final destination and the time at which the packet is generated, and the average delay is the delay averaged over all the successfully received packets.

The end-to-end completion percentage with each of the three protocols is shown as a function of the total packet generation rate in the network in Fig. 7.6. The results are shown for a processing gain of 20. If the target end-to-end completion percentage is 90%, the maximum generation rate at which this is achieved with the exclusive use protocol is 197 packets/s with the exclusive-use protocol, 164.4 packets/s with the unrestricted reuse protocol and 238 packets/s with the SR III (1, 2) protocol. Thus we obtain a 21% improvement over the exclusive-use protocol and a 45% improvement over the unrestricted reuse protocol. At a processing gain of 40 the maximum generation rate at which 90% completion percentage is achieved is 197.9 packets/s with exclusive-use, 241.9 packets/s with unrestricted reuse and 254 packets/s with the SR III (1, 2) protocol.

The average delay experienced with each of the three protocols is shown in Fig. 7.7 as a function of the total packet generation rate in the network at a processing gain of 20. At any particular generation rate, we can see that the SR III (1,2) outperforms both of the static reuse approaches in terms of delay performance.

At a high generation rate the unrestricted reuse protocol has much smaller delay than the exclusive-use protocol. The reverse is true at lower generation rates, however. At the generation rate at which 90% completion is achieved with each of the protocols, the average delay experienced by the successfully received packets is 150ms with exclusive-use, 110ms with unrestricted-reuse and 160ms with SR-III (1,2). These results demonstrate that the MAC layer performance benefits that result from the use of the selective reuse protocol translate well into network-wide improvements in performance.

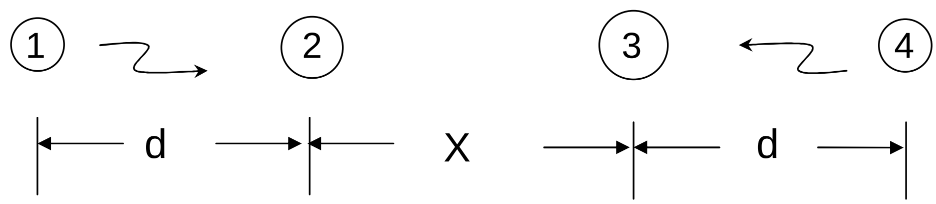

Figure 7.1 Illustration of factors affecting spatial-reuse opportunities.

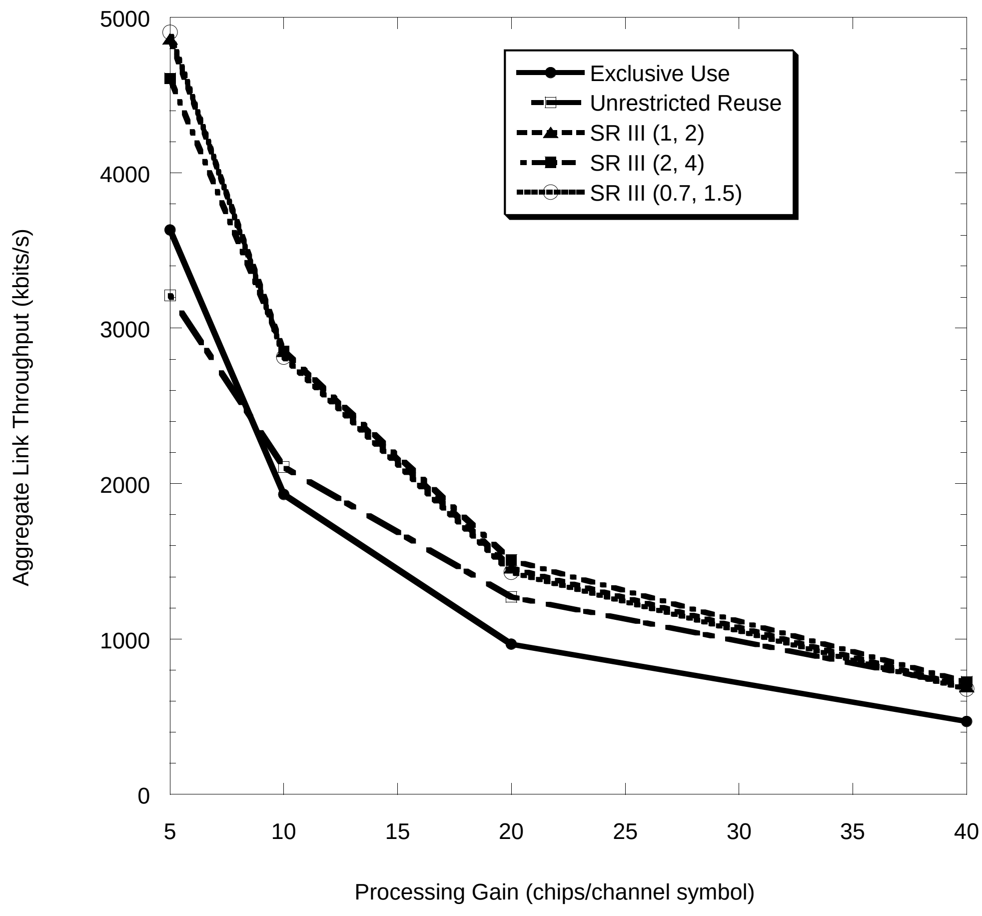

Figure 7.2 Performance of three protocols including SR III for different processing gains in a random 30-node network.

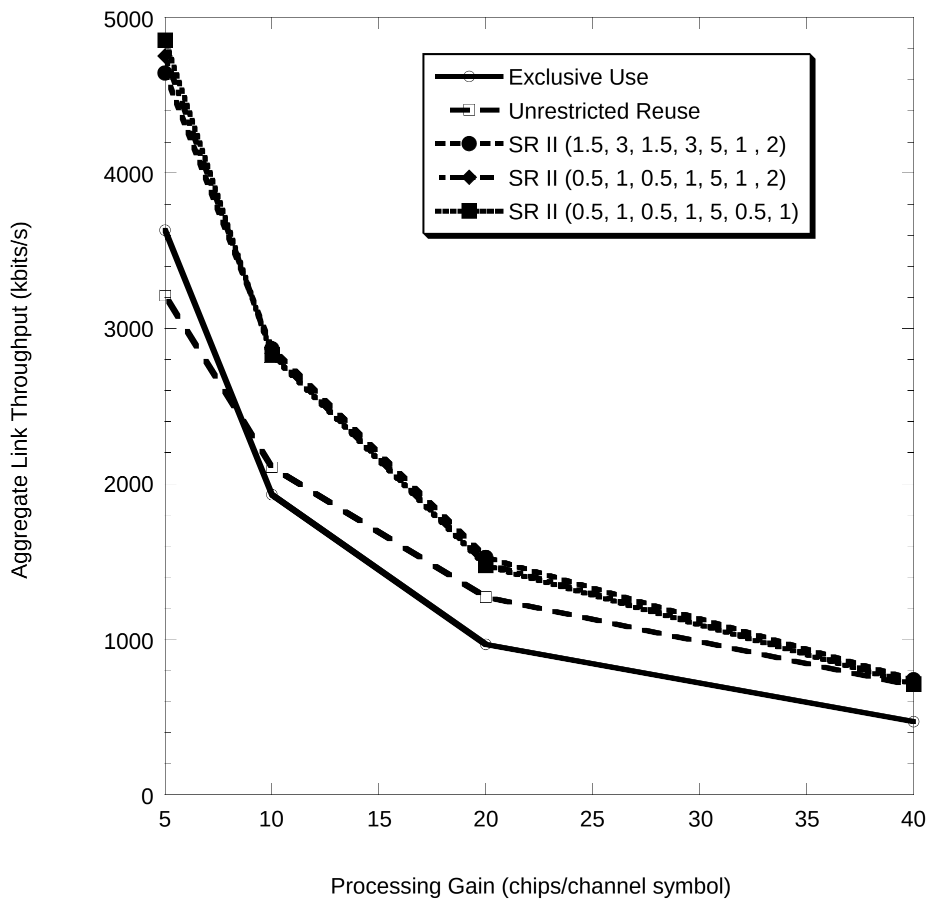

Figure 7.3 Performance of three protocols including SR II for different processing gains in a random 30-node network.

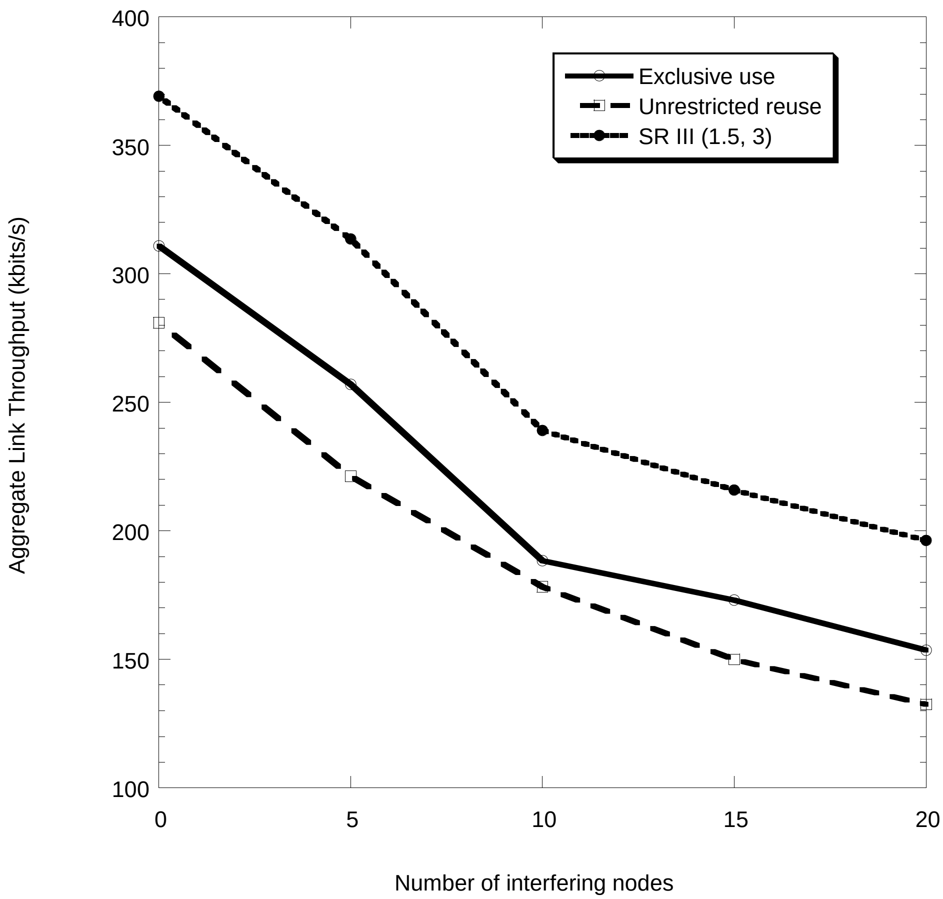

Figure 7.4 Performance of four protocols for different network densities in setting one.

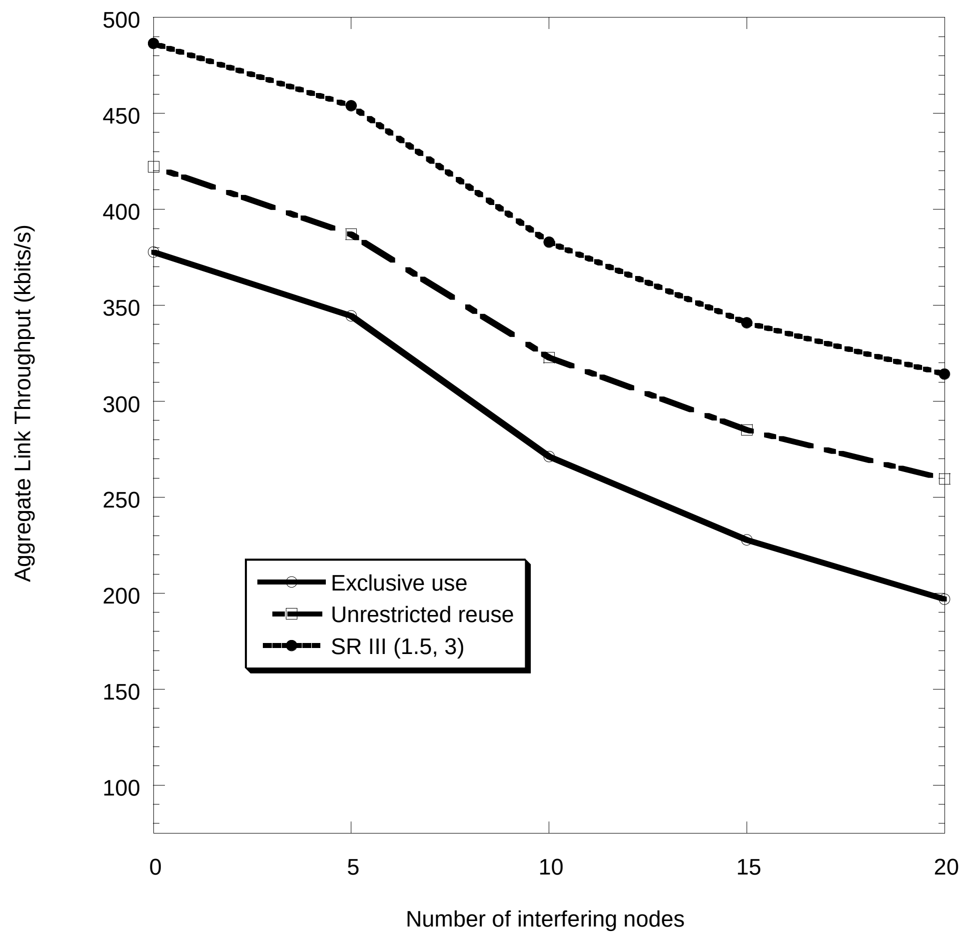

Figure 7.5 Performance of four protocols for different network densities in setting two.

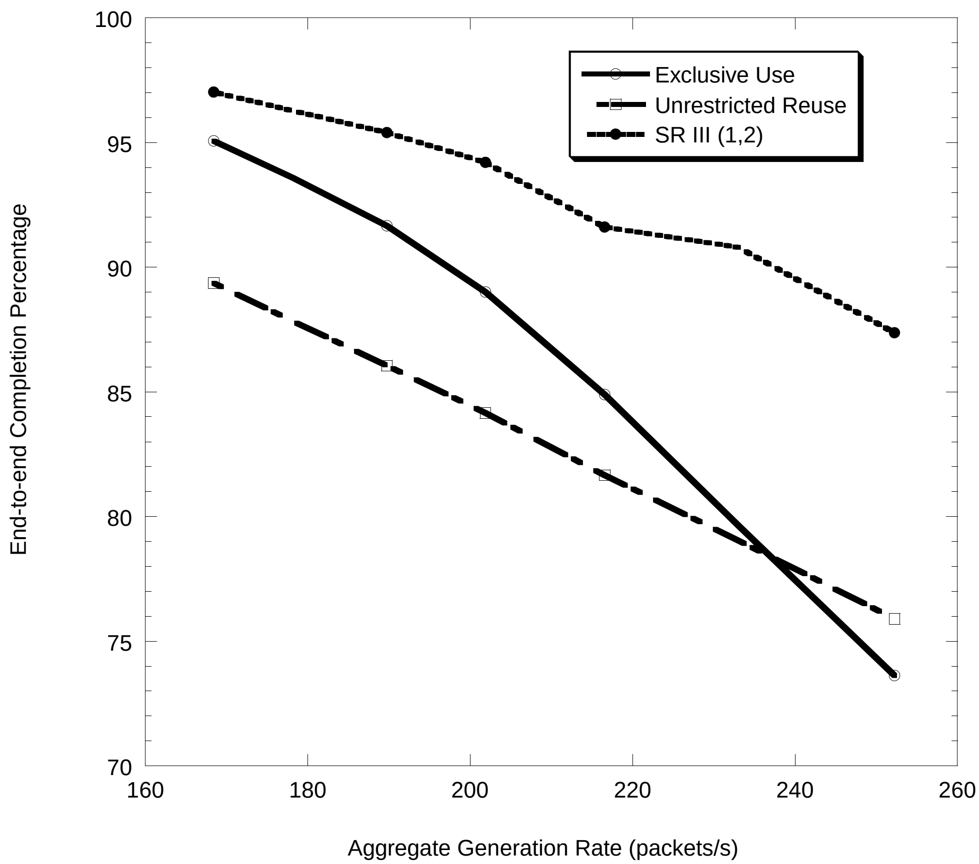

Figure 7.6 End-to-end completion percentage with three protocols in random 30-node networks.

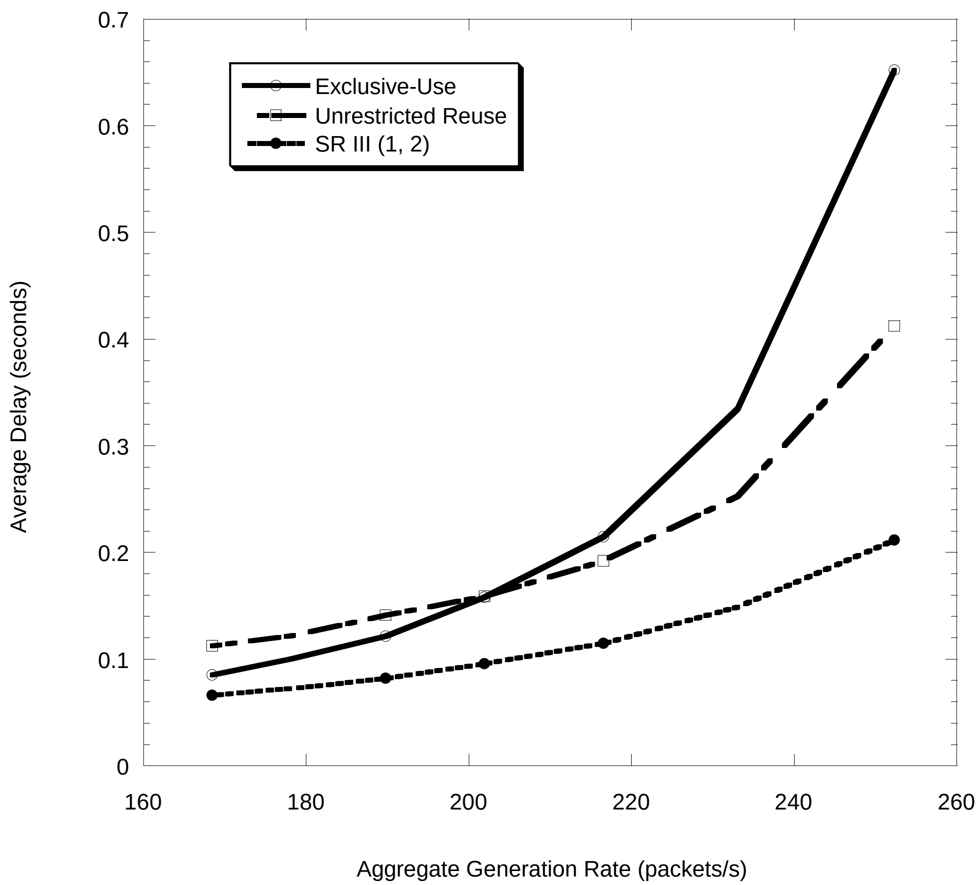

Figure 7.7 Average delay with the three protocols in random 30-node networks.

CHAPTER 8

DESIGN OF A ROUTING PROTOCOL THAT EXPLOITS THE AVAILABILITY OF NODES WITH DIRECTIONAL ANTENNAS

The performance for a data flow following a particular path through the network is affected by the presence of interfering transmissions along the route, contention for the attention of radio receivers along the route, and the backlog of traffic to be relayed by each node in the route. These forms of interference and contention for resources among data flows following various paths (collectively referred to here as *mutual coupling* among the paths) determine the level of congestion encountered by each data flow, and thus incorporation of any of these factors into the path metric used for routing can potentially mitigate congestion. This is the approach employed in several *congestion-avoidance (CA) routing protocols* that have been introduced previously, including protocols using link-state routing, others using distance-vector routing, and others using on-demand routing. Measures of local congestion used in the protocols include the number of active links near a given link [15, 37] or near a given node [38], the level of traffic on nearby links [39], the queue occupancy of a node [40], and the busyness of a node [41]. The presence of nodes with directional antennas can provide quite different mechanisms for limiting or responding to congestion than are available in a network with only omnidirectional antennas. Yet there is only limited prior consideration of congestion-avoidance routing protocols in a network with directional antennas [15].

It has been shown for a wired network that optimal routing employs multiple paths between each source and destination pair together with splitting of the corresponding data flow among the multiple paths [42]. The optimal split of a data flow among the paths depends on the congestion in the paths, the capacity of each link, and the rate of traffic that must be supported between each source-destination pair. Yet there has been only limited prior consideration of traffic splitting as a mechanism for congestion avoidance in a wireless network [37, 43], perhaps due to the perceived problem of high mutual coupling among paths in a network of nodes with omnidirectional antennas.

In this chapter, we demonstrate that the presence of nodes with directional antennas reduces the level of mutual coupling among active paths in the wireless network, thus providing an increased opportunity to exploit traffic splitting. A congestion-avoidance, distance-vector routing protocol is introduced that employs traffic splitting to limit the level of mutual coupling. It is designed for networks that include an arbitrary mix of nodes with directional antennas and nodes with omnidirectional antennas. The performance of the congestion-avoidance protocol is shown to result in significantly better performance than a minimum-hop (*min-hop*) routing protocol. It is also shown that while congestion-avoidance routing improves the performance in a network of all omnidirectional antennas, the greatest improvements are achieved in the presence of nodes with directional antennas.

8.1 Related Prior Work

Congestion-avoidance routing for ad hoc networks without traffic splitting is considered in several previous papers [15], [38], [39], [40], [41]. To the best of our knowledge, only [15] addresses ad hoc networks in which the nodes have directional antennas. (Unlike in our work, only networks in which *all* the nodes have directional antennas are considered in [15].) Based on a simple example network, it is claimed in [15] that congestion-avoidance routing is beneficial only if *zone-disjoint* routes are available between the different source-destination pairs and that the probability of finding such routes is high only when all the nodes have directional antennas. In contrast, we demonstrate that congestion-avoidance routing results in noticeable benefits even if a small fraction or even if none of the nodes have directional antennas.

The routing protocol in [15] is a distributed link-state protocol in which each link is assigned a cost based on a metric referred to as the *correlation factor*, a measure of the number of active routes within range of the transmitter's beam. The level of traffic in each of the active routes is not taken into account, however (in contrast with the protocol we introduce in this chapter). Furthermore, the protocol in [15] does not consider route-splitting. Finally, since a link-state approach is employed in [15] it requires network-wide dissemination of state information that may be rapidly time-varying.

None of the other work addressing congestion-avoidance routing for ad hoc wireless networks considers an ad hoc network in which the nodes have directional antennas. The protocols in [38],[39] and [40] are used to improve the performance of *on-demand routing protocols* by incorporating a metric that accounts for congestion during the route-discovery phase. Other than the metric used to calculate path cost, the route-discovery phase is similar to conventional on-demand routing protocols.

In [38], each node is assigned a *nodal activity* that is the sum of the number of active paths in the node's neighborhood, and the cost of a route is the sum of the nodal activities of its constituent nodes. In [39], the nodes exchange *hello* messages to estimate their *available bandwidth* as the residual bandwidth within their two-hop neighborhood. The available bandwidth along a route is then calculated as the minimum of the available bandwidths of its constituent nodes. Finally, in [40] the authors propose three different path metrics that are all a function of the number of packets queued in the nodes in the route. In the first metric, which is shown to yield the best performance, the route cost is just the sum of the number of packets queued in each of the individual nodes in the route.

In [38],[39], the initial route request serves as a probe that identifies the least congested path. Once this is done there is no adaptation of the route based on subsequent traffic that arrives into the network. The protocol in [40] does provide some adaptation, however. If the cost of the route at a certain point in time exceeds the initial route cost by a certain threshold, the destination initiates a route request towards the source, and a new low-cost route is established between the source and destination using the route requests received at the source.

Routing metrics that reflect the congestion in a network are also considered in [41] for frequency-hop packet radio networks. Each node in [41] determines an *activity metric* that measures the fraction of time the node is busy. If the measured activity metric exceeds a threshold, the node switches from a routing protocol based only on energy-efficiency to one that uses a hybrid metric that also includes the effect of hop-count and link-quality. The rationale behind the approach is that a routing protocol based on energy-efficiency alone usually selects routes with a large number of short length hops, leading to poor performance at high traffic loads.

Theoretical results for wired networks suggests that the performance of routing protocols can be improved by incorporating the capability to split traffic for a single source-destination pair along multiple paths. In spite of this, it appears there has been only limited work on routing protocols for wireless ad hoc networks that exploit traffic splitting [37],[43].

The protocol described in [37] is an augmented on-demand routing protocol with the ability to use multiple disjoint paths for each source-destination pair. In the route-discovery phase, the destination sends a route reply along the first received route-request's path. From among the alternate paths advertised in the route requests received after the first one within a certain period of time, the destination selects a second route reply along the route which is maximally disjoint from the path used by the first arriving route request. (The definition of maximally disjoint used in [37] is not clear, but it appears that node-disjointedness is used.)

It is shown in [43] that the potential advantage of multiple-path routing is much less pronounced in wireless ad hoc networks than in wired networks. (It is critical to note that all the nodes in [43] employ omnidirectional antennas, however). If a single traffic channel is employed, the increase in aggregate throughput due to multiple-path routing is shown to be only on the order of 2%. If multiple traffic channels are available, performance gains on the order of 20% are possible since the alternate routes are coupled only if they share the same node. (The number of traffic channels is implicitly assumed to be infinite for the multiple-channel case in [43].)

8.2 Motivating Illustration

The introduction of directional antennas at some or all nodes in an ad hoc network can result in a reduction in the mutual coupling among routes in the network, and this reduction is illustrated by considering the network topology shown in Figs. 8.1 and 8.2. In this idealized network, the distances are such that nodes 2, 4 and 5 are not within range of each other, and the processing gain is such that a data reception at node x is unaffected by transmissions from nodes y and z for all possible combinations of (x, y, z) from the set $\{2, 4, 5\}$. However, the distances and processing gains are such that if any node in the set $\{1, 2, 3\}$ is receiving packets, then the probability that the packet reception is successful is very low when any of the other nodes in the same set are transmitting. The network employs four frequency channels, including one control channel and three data channels, and the data rate and packet sizes are such that if only a single transmitter uses a given channel, it can transmit at most X packets/s.

Suppose nodes 1 and 2 each generate packets at a rate of R packets/s destined for nodes 4 and 3, respectively, and that no other source data is generated in the network. If min-hop routing is employed, the path 2-3 will be used to carry the packets from node 2 to node 3, and the path 1-3-4 will be used to carry the packets from node 1 to node 4. Consequently, stable operation of the network can be achieved only if nodes 1, 2 and 3 are each able to transmit at a rate of R packets/s. It is clear that the offered traffic that can be supported is limited by the constraint that no two of the following events can take place concurrently: node 1 transmits to node 3, node 2 transmits to node 3, and node 3 transmits to node 4.

If nodes 1, 2, and 3 use transmission opportunities fairly and channel-access arbitration time is negligible, the maximum packet generation rate per source, R, that can be supported is $R = X/3$ packets/s. This results in a total network throughput of $2X/3$ packets/s. In this instance it is the availability of the half-duplex nodes rather than the availability of traffic channels that limits throughput, and in particular, node 3 is the bottleneck.

Suppose instead that nodes 1 and 3 of Fig. 8.1 employ three-sector directional antennas, resulting in the network depicted in Fig. 8.2. If min-hop routing is employed, the same routes (2-3 and 1-3-4) result as in the network of Fig. 8.1. But in this instance the offered traffic that can be supported is limited only by the less stringent constraint that transmissions from node 1 to node 3 and node 2 to node 3 cannot occur concurrently. As a result, if nodes 1 and 2 use transmission opportunities fairly and channel-access arbitration time is negligible, the maximum offered load that can be supported is $R = X/2$ packets/s. In this instance it is the availability of sector II of node 3 that creates a bottleneck.

Consider the alternative path 1-5-3-4 from node 1 to node 4 in the network of Fig. 8.2. Due to the presence of directional antennas, the path 1-5-3-4 does not have any mutual coupling with the path 2-3 (under the approximation that channel-access arbitration time is negligible). Furthermore, the mutual coupling between paths 1-3-4 and 1-5-3-4 is less than in the network with only omni-directional antennas. In the network of Fig. 8.1, the two paths share node 1, node 3 as both a link destination and a link source, and node 4. But in the network of Fig. 8.2, the two paths share only node 3 as a link source and node 4 (and hence the link 3-4).

The performance of the network in Fig. 8.2 can be improved by the following routing protocol. The routing protocol ensures that all traffic from node 2 to node 3 is routed on the path 2-3. Further suppose that the traffic from node 1 to node 4 is routed on the path 1-3-4 as long as the aggregate traffic into sector II of node 3 is no more than X packets/s. If the aggregate traffic into sector II of node 3 reaches X packets/s, any excess traffic generated at node 1 is routed on the alternate path 1-5-3-4. This approach permits stable network operation for any aggregate generation rate of $3X/2$ packets/s as long as neither source generates traffic at a rate of more than X packets/s. For instance, stable operation is achieved if the respective generation rates at nodes 1 and 2 are $X/2$ packets/s and X packets/s, and stable operation is also achieved if the generation rate at each source node is $3X/4$ packets/s.

This example illustrates one form of congestion-avoidance routing. It requires a routing protocol that is able to detect the occurrence of excessive congestion on the path 1-3-4 and to determine that rerouting traffic from node 1 to node 4 onto the alternate path 1-5-3-4 will mitigate the congestion and improve network performance.

Thus the presence of directional antennas decreases the mutual coupling among paths in two important respects. Mutual coupling is decreased among alternative paths between the same source-destination pair (such as between the paths 1-3-4 and 1-5-3-4 in our example). It is also decreased among paths between different source-destination pairs (such as between the path 1-5-3-4 and 2-3 in our example). A properly designed congestion-avoidance routing protocol can exploit the reduced mutual coupling to improve network performance. Moreover, the lower path coupling achieved with directional antennas can be critical to the effectiveness of congestion-avoidance routing. In the example considered here, for instance, the congestion-avoidance protocols result in improved performance *only* if the directional antennas are present. (We later show that in large networks with randomly placed nodes, noticeable benefits are also obtainable when a small fraction or even none of the nodes have directional antennas.)

8.3 A Simplified Model of Congestion-Avoidance Routing

A simplified network model is used to gain some insight into the value of congestion-avoidance routing and the impact of directional antennas on the effectiveness of congestion-avoidance routing. The network consists of nodes distributed within a rectangle of X m by Y m. Viable radio links are determined by the disc-graph model with a disc radius of R m. The same radius determines the range for both reception of desired transmissions and susceptibility to interfering transmissions.

Parameters of the network include the number of nodes, the percentage of nodes with directional antennas, and the number of active source-destination pairs. The network topology is determined by random placement of the nodes within the rectangle and random selection of the subset that has directional antennas. Active source-destination pairs are selected at random under the constraint that no source-destination pair has a direct radio link. (Thus all traffic requires multiple-hop routing.) A single route is used for the traffic of a given source-destination pair, and the route is chosen based on the routing protocol under consideration. (For both routing protocols considered using the simplified network model, routes are determined using the centralized Bellman-Ford algorithm.) Each randomly selected source-destination pair determines a unidirectional traffic flow, so a given pair of nodes exhibit bidirectional end-to-end traffic only if the two directions are generated as separate random draws.

Network performance is characterized in terms of a congestion-based route metric that measures the level of busy-ness of each node in a route and the level of contention in each link in the route. For a given set of network parameters, a large number of network topologies are generated randomly. For each topology, the effectiveness of the routing protocol in mitigating congestion is measured by the sum of the route metrics for all the active routes. The performance for a given set of network parameters is determined by averaging the measure of congestion over all the randomly generated topologies.

Idealized forms of two routing protocols are evaluated using this network model: standard min-hop routing, and (non-splitting) congestion-avoidance min-hop routing. (The standard min-hop routing protocol is subsequently referred to simply as the *min-hop routing protocol*, and the congestion-avoidance min-hop routing protocol is referred to simply as the *congestion-avoidance routing protocol*.) In the min-hop routing protocol, a single route between a given source-destination pair is chosen randomly among all the shortest paths between the source and the destination, and that route is used as the sole route for the corresponding traffic. In contrast, the congestion-avoidance min-hop routing protocol determines multiple shortest paths (up to a specified maximum number) as candidate routes between a given source and destination. It selects a single path as the route for each source-destination pair, and the paths for all source-destination pairs are chosen to optimize a congestion metric defined below. In either protocol, independent routes are selected for the two directional flows if a node pair exhibits bidirectional end-to-end traffic.

The congestion-avoidance protocol uses iterative applications of the Bellman-Ford algorithm to determine the set of candidate routes for a given source-destination pair. In the first iteration, one shortest path is identified and selected at random among the shortest paths and used as the initial entry in the list of candidate routes for the source-destination pair. The connectivity graph of the network is then modified to exclude the links in the first candidate route. In the second iteration, another path with the same hop count as the first candidate route is identified in the modified graph (if any such paths exist), and it is added to the list of candidate routes. In each subsequent iteration the links in previously selected candidate routes are excluded from the connectivity graph. One additional path of the same hop count as the candidate routes is sought for addition to the list of candidate routes. The iterations continue until the specified maximum number of candidate routes have been selected or until no additional paths of the minimum hop count exist in the modified graph. The iterative procedure is performed for each source-destination pair, and the network graph is reinitialized to its original connectivity at the start of the procedure for each pair.

8.3.1 Contention-Based Metrics

The aggregate contention experienced in the assigned routes in the network is captured by a single network-wide metric that is determined from contention metrics at the node, link, and route levels. The network-wide metric is used as a basis of comparison between the two routing protocols for different mixes of directional and omni-directional antennas. The definition of the metrics is motivated by the behavior of a network in which an RTS/CTS channel-access protocol is employed with common spreading sequences for both RTS and CTS packets and in which reception of either an overheard RTS or CTS packet serves to block the subsequently assigned traffic channel. (This conceptual model differs from the actual channel-access protocols considered in the previous sections. Those protocols use receiver-directed spreading for RTS packets, and only overheard CTS packets are used to block traffic channels. Moreover, the conceptual model is idealized, since an RTS packet in an actual multiple-channel network may advertise multiple traffic channels - thus limiting its usefulness as a basis for channel blocking.)

The delay encountered by a packet traveling from a source node to a destination node is the sum of the delays encountered in the individual links (and the corresponding nodes) in the path. The delay encountered in a link between two nodes depends on the contention for the use of the traffic channels and the availability of the nodes. This is illustrated by considering the three networks shown in Fig. 8.3. In the first network, node 1 transmits packets to node 3. Congestion at node 2 results from the requirement that it support incoming traffic from node 1 and outgoing traffic to node 3, and it is the sum of these two quantities that limits the availability of node 2 to support other traffic. The second network includes an additional intermediate node (node 4) that is located between nodes 2 and 3. In this network, node 2 must support incoming traffic from node 1 and outgoing traffic to node 4. In addition, if only a single traffic channel is available, node 2 must remain idle whenever a 4-3 transmission occurs. In contrast, if two or more traffic channels are available, a 4-3 transmission and a 1-2 transmission can take place concurrently. The third network in Fig. 8.3 also includes a second route, 5-6-2-7. In this case, node 2 must support transmissions for four links: 1-2, 2-4, 6-2 and 2-7. Furthermore, it may have to remain idle during transmissions on links 5-6 or 4-3, depending on the number of available traffic channels. Clearly, the ability of a node to support additional traffic is determined by two quantities: the level of traffic that the node is currently supporting, and the level of traffic between other nodes in the neighborhood. Consequently, the availability of a given link is determined by the effect of these factors on the pair of nodes that form the link. These observations generalize to links in which one or both nodes use directional antennas.

Based on these observations, we employ a metric referred to as the *total node contention (TNC)* for each node/sector in the network. The TNC of a node/sector is the sum of two metrics intended to reflect the two factors affecting the availability of the node: the *node-overlap contention (NOC)* metric, and the *nodes-within-range contention (NWRC)* metric. The rationale for the definition of the metrics is based on the approximation that the traffic at each node/sector in the network is proportional to the number of routes passing through the node/sector.

The NOC metric for a node/sector is defined as the number of traffic flows on routes that pass through the sector. The metric does not account for the effect of co-site interference at the node. The NRWC metric for a node/sector is given by the sum of the number of traffic flows on routes passing through neighboring nodes/sectors that are within range, excluding those flows that pass through the node/sector of interest. Note that the same traffic flow is counted twice if both nodes/sectors forming the corresponding link are within range of the node/sector of interest.

The TNC metric for each node/sector is defined as the sum of its NOC metric and the larger of zero and $NWRC - numChans + 1$, where $numChans$ is the number of traffic channels in the network. The *total link contention (TLC)* of each link in the network is defined as the larger of the TNC metrics for the two nodes/sectors that form the link. The *route congestion (RC)* metric for each route in the network is defined as the sum of the TLC metrics for the links in the route, and the *total congestion (TC)* metric for the network is defined as the sum of the RC metrics for all the routes in the network. It is shown in Section 8.3.2 that the TC metric can serve as an accurate predictor of the relative aggregate throughput of a network under heavy load for different mixes of directional and omni-directional antennas.

The calculation of the TC metric for a network using either the min-hop protocol or the congestion-avoidance protocol is illustrated by considering a network in which there are two source-destination pairs, P_1 and P_2, and there are a maximum of three min-hop paths allowed in the list of candidate routes for each pair. The candidate routes are denoted as R_{ij}, for all $i \in \{1, 2\}$ and $j \in \{1, 2, 3\}$. For each randomly generated topology and combination of source-destination pairs, one candidate route is selected at random for each pair and the corresponding TC metric is determined. The metric is taken as the TC metric for that topology and selection of source-destination pairs using the min-hop protocol. For the congestion-avoidance protocol, the TC metric is determined instead for each of the (up to) nine combinations of candidate routes for the two pairs. The smallest of the TC metrics is taken as the TC metric for that topology and selection of source-destination pairs using the congestion-avoidance protocol. For either routing protocol, the average of the TC metrics over all the randomly generated topologies is the measure of average congestion for the given set of network parameters.

8.3.2 Validation of Simplified Model and Total-Congestion Metric

In this subsection, we evaluate the simplified network model with the TC metric as a tool for comparing network performance under different conditions. An example network with fixed node placements and source-destination pairs is considered for two different sets of static routes. The relative performance of the network using the first set of routes and the network using the second set of routes is determined using the simplified model, and the relative performance using the two sets of routes is also determined using a more realistic network model. It is shown that the simplified model is an accurate predictor of the relative performance for the two scenarios under heavy traffic conditions.

The results from the simplified model are expressed in terms of the TC metric, and they are compared with the results of an Opnet-based simulation in which the results are expressed in terms of aggregate network throughput. The Opnet simulation employs source routing [44], and thus the same routes can be prescribed that are considered with the simplified model. In contrast with the simplified model, however, the Opnet simulation includes the physical layer and the baseline MAC protocol described in Chapters 2 and 3 of the report. Thus the simplified model differs from the Opnet simulation in several respects. The simplified model does not account for the effects of random noise. It uses a deterministic rather than a probabilistic model of the effects of multiple-access interference (MAI), and it uses a much different model of the distance-dependent effects of MAI than does the Opnet simulation. Since the baseline MAC protocol does not use RTS packets for channel blocking, the Opnet simulation accounts for the possibility of concurrent use of a traffic channel in a neighborhood (which may be either successful or unsuccessful). The simplified model does not account for this possibility. Finally, the effect of co-site interference is modeled in the Opnet simulation, whereas the simplified network model does not account for the effect of co-site interference.

A network of thirty nodes using omni-directional antennas is generated by random placement in a square of 6000 m by 6000 m, and each node is assigned an identifier between 1 and 30. The number of traffic channels is three. Ten source-destination pairs are chosen at random. Two sets of route assignments are considered. The first set of routes is shown in Table 8.1, and the TC metric for the network using this route set is determined to be 434 using the procedure described earlier. The second set of routes is shown in Table 8.2, and the TC metric for the network using this route set is 914. Hence the network's TC metric with the routes in set two is approximately 2.1 times the network's TC metric with the routes in set one.

The results of the Opnet simulations using source routing are shown in Fig. 8.4. The peak aggregate network throughput obtained with the routes in set one is ap-

Pair	Nodes in Route
1	2-25-15-7-13
2	18-20-9-1
3	28-18-14-21-4-13
4	10-4-13
5	30-17-9-20
6	24-17-29-18
7	27-11-19-6-8
8	6-9-20
9	22-25-15-30
10	1-6-19-11

Table 8.1 First set of routes used.

Pair	Nodes in Route
1	2-25-21-4-13
2	18-29-10-1
3	28-20-9-17-24-13
4	10-4-13
5	30-17-9-20
6	24-17-29-18
7	27-4-10-9-8
8	6-9-20
9	22-29-17-30
10	1-10-4-11

Table 8.2 Second set of routes used.

proximately 125 packets/s, while the corresponding value obtained with the routes in set two is approximately 64 packets/s. Thus the maximum heavy-load throughput using route set one is approximately 1.95 times the maximum heavy-load throughput using route set two. This ratio is close to the inverse of the ratio of the respective TC metrics. A similar comparison has been performed for other topologies and various sets of source-destination pairs. For a given topology and set of source-destination pairs, the inverse of the TC metrics for different route sets is a consistently good predictor of the relative maximum heavy-load throughput for the route sets.

8.3.3 Performance Comparisons using the Simplified Network Model

In this subsection the simplified network model and the TC metric are used to compare the performance of congestion-avoidance routing with min-hop routing. A key issue that is considered is the extent to which the relative performance of the two protocols depends on the presence of nodes in the network with directional antennas. Though the congestion-avoidance protocol that is considered does not include route splitting, insights provided by the comparison are suggestive of the value of the route-splitting congestion-avoidance protocols.

Several network scenarios are considered which are parameterized by the number of nodes, the area covered by the network, the number of source-destination pairs, the number of traffic channels, and the maximum number of candidate routes per source-destination pair used in the congestion-avoidance protocol. The range of each node in the disk-graph model is 2000 m, and each node with directional antennas employs three sectors. For each scenario, 1000 trials are employed with random placement of nodes and random selection of source-destination pairs for each trial. The network's TC metric is determined for each trial for both the min-hop protocol and the congestion-avoidance protocol, and the values for the respective protocols are averaged over the trials. The average TC metrics are determined as a function of the percentage of nodes with directional (sectored) antennas as the percentage is varied from zero to 100.

The average TC metric is shown in Fig. 8.5 for a 30-node network that employs one traffic channel and covers a 7000 m by 7000 m square. The performance of the congestion-avoidance protocol is significantly better than the performance of the min-hop protocol. The improvement in performance is approximately 20% if all nodes have omni-directional antennas, 25% if 30% of the nodes have directional antennas and 39% if all of the nodes have directional antennas. The improvement provided with the congestion-avoidance protocol is even more pronounced if the same network employs three traffic channels. This is shown in Fig. 8.6. The performance gains range from the smallest gain of 23% if all the nodes have omni-directional antennas to the largest gain of 63% if all the nodes have directional antennas.

The congestion-avoidance protocol provides significantly better performance than the min-hop protocol even if a larger number of traffic sources are active. Figs. 8.7 and 8.8 show the performance of the same network with ten source-destination pairs. The performance with one traffic channel is show in Fig. 8.7, and the performance with three traffic channels is shown in Fig. 8.8. The performance improvement provided by the congestion-avoidance protocol ranges from 20% with all omni-directional antennas to 37% with all directional antennas. If there are three traffic channels, the corresponding range is 22% to 52%.

Finally, Figs. 8.9 and 8.10 show the average TC metric for a 40-node network employing one traffic channel and three traffic channels respectively. The number of source-destination pairs is 10 in either case. With one traffic channel case, congestion avoidance provides a performance improvement of between 21% and 38%. If there are three traffic channels the range of performance improvement is between 21% and 53%.

Thus the use of the congestion-avoidance routing protocol consistently provides much better network performance than the min-hop routing protocol as measured by the TC metric. The percentage improvement provided by congestion-avoidance routing increases as the fraction of nodes with directional antennas is increased. Yet contrary to some claims from prior work of others (including [15]), congestion-avoidance routing is beneficial even if all the nodes in the network have omni-directional antennas. Moreover, congestion-avoidance routing exploits multiple traffic channels more effectively than min-hop routing, and it provides its largest gains if the network has multiple traffic channels rather than only one traffic channel.

Note that our results are consistent with prior results in demonstrating that the use of directional antennas improves network performance regardless of the routing protocol. Indeed with the congestion-avoidance protocol, the TC metric decreases between 80% and 90% for different networks as the percentage of nodes with directional antennas is increased from zero to 100. Finally, note that these results for an "idealized" network do not incorporate other possible protocol improvements such as split-traffic routing for source-destination pairs or the use the directional antenna's gain to achieve longer hops.

8.4 Congestion-Avoidance Routing Protocol

In this section, we describe a practical proactive routing protocol that is effective in exploiting the availability of nodes with directional antennas. The protocol employs a metric that reflects the level of congestion encountered in a path through the network and that serves as the basis for a traffic-splitting approach to adaptive congestion-avoidance routing. The objective of the protocol is to achieve an allocation of traffic flows among paths that results in a small average traffic-weighted mutual coupling among the paths while employing links of sufficient reliability to achieve an acceptable probability of successful end-to-end packet transmission in each path. Specifically, the metric $R(P)$ calculated for path P at the source node is given by the sum of two components, a *traffic-contention metric* denoted by $TC(P)$ and a *path-quality metric* denoted by $Q(P)$, which are described below. Both metrics are calculated using the distributed Bellman-Ford distance-vector algorithm.

8.4.1 Traffic-Contention Metric

The traffic-contention metric for a path is the sum of traffic-contention metrics for the links in the path, and the link traffic-contention metrics are derived in turn from metrics associated with the node/sector pair that forms the link. Each node in the network determines current values of two *node-contention metrics* for each sector of the node and each link to a neighbor in that sector. These are updated at intervals of $TimeBwUpdate$ seconds. We define $TimeBwUpdate$ as $numAveragingSlots$ times the duration of a packet interval, where a packet interval is the time to complete an RTS, CTS, data packet, and ACK exchange measured in seconds. They are determined by four measurements based on the state of the channel-access sublayer that are made during each update interval: *TimeBusy*, *TimeCtrlBlk*, *TimeTCBlk[i]*, and *TimePacing*.

The quantity *TimeBusy* is the total time during the interval that the node/sector is busy either transmitting or receiving a packet. The quantity *TimeCtrlBlk* is the total time during the interval that the node/sector is idle but the control channel is unavailable for use (due to the co-site-interference constraint on a node with directional antennas). The quantity *TimeTCBlk[i]* is the total time during the interval that the node/sector is idle and the control channel is available but exactly i traffic channels are unavailable. The quantity *TimePacing* is the portion of the remaining time during the interval that the node/sector undergoes pacing.

One value of each of the quantities *TimeBusy*, *TimeCtrlBlk*, and *TimePacing* is determined at the node/sector at any point in time, and the three values are used in calculating the node-contention metrics for the links to all neighbors. In contrast, a separate value of *TimeTCBlk[i]* is determined for the link to each neighbor and is used in calculating the node-contention metric to that neighbor. The values of *TimeTCBlk[i]* for different links differ only in that each accounts for the time spent in exponential backoff with respect to that neighbor. (For instance, if a node/sector begins an exponential backoff of length T seconds with respect to a neighbor, $TimeTCBlk[NumChans-1]$ for that neighbor is incremented by T while the $TimeTCBlk[NumChans-1]$ for the remaining neighbors is left unchanged.) For convenience, we simplify the notation for the node-contention metrics in what follows by not showing the dependence on the neighbor.

The two node-contention metrics for the node/sector at the end of the nth update interval are given by

$$NC_{Rx}(n) = \alpha * NC_{Rx}(n-1) + (1-\alpha) * NC_{Rx,curr} \tag{8.1}$$

and

$$NC_{Tx}(n) = \alpha * NC_{Tx}(n-1) + (1-\alpha) * NC_{Tx,curr} \tag{8.2}$$

where

$$NC_{Rx,curr} = \frac{TimeBusy + TimeCtrlBlk + \frac{\sum_{i=1}^{NumChans-1} i * TimeTCBlk[i]}{NumChans-1}}{TimeBwUpdate} \tag{8.3}$$

and

$$NC_{Tx,curr} = NC_{Rx,curr} + \frac{TimePacing}{TimeBwUpdate} \tag{8.4}$$

and the values in the numerators of (8.3) and (8.4) are the measurements from the current (nth) update interval. The parameter α determines the degree of smoothing applied to sequential measurements, and the parameter $NumChans$ indicates the total number of channels. Note that $0 \leq NC_{Rx}(n) \leq NC_{Tx}(n) \leq 1$. The value of $NC_{Rx}(n)$ is intended to measure the level of busyness of the node/sector with respect its function as a receiver, and the value of $NC_{Tx}(n)$ provides the comparable measure with respect to the node/sector's function as a transmitter. The difference in the quantities reflects the fact that a node/sector can receive a packet while it is the pacing state, but it cannot initiate a transmission.

The current value of NC_{Rx} for a node/sector is included in channel-access-sublayer acknowledgment packets and the periodic PROP packets that are broadcast to the node/sector's one-hop neighbors. Each node/sector uses its local node-contention metrics and those received from its neighbors to determine the *link-contention metric* for each link that involves the node/sector. The link-contention metric at node i/sector s_i for the link from that node/sector to node j/sector s_j is given by

$$LC(i, s_i, j, s_j) = \max\{NC_{Tx,i,s_i}, NC_{Rx,j,s_j}\} \tag{8.5}$$

where NC_{Tx,i,s_i} denotes the current value of the corresponding node-contention metric, NC_{Rx,j,s_j} denotes the most recently received value of the corresponding node-contention metric.

The traffic-contention metric formed at node i/sector s_i for the link from that node/sector to node j/sector $s_j$is obtained from the corresponding link-contention metric and a monotonically increasing function f_{TC} by

$$TC(i, s_i, j, s_j) = f_{TC}(LC(i, s_i, j, s_j)). \tag{8.6}$$

The traffic-contention metric of the links serves as the additive metric which forms the traffic-contention metric for each path. Thus the traffic-contention metric of the path P consisting of node/sector $(n_1, s_{n_1}), (n_2, s_{n_2}), \cdots, (n_m, s_{m_1})$ is given by

$$TC(P) = \sum_{i=1}^{m-1} TC(n_i, s_{n_i}, n_{i+1}, s_{n_{i+1}}). \tag{8.7}$$

The path metrics are distributed in the PROP packets.

In the examples in this dissertation, the function f_{TC} that maps each link-contention metric into a link traffic-contention metric is the non-linear function shown in Fig. 8.11. Different choices of the parameters shown in Fig. 8.11 define different functions. This form for the function f_{TC} is motivated by the fact that the delay experienced by a packet while traversing a link is an exponential function of the traffic load in some simple queuing models of the link. The function f_{TC} provides a piecewise-linear approximation to an exponential function.

8.4.2 Path-Quality Metric

The path-quality metric for a path is the sum of link-quality metrics for the links in the path, and the link-quality metric for transmissions in a particular direction on a link is determined by signal-quality measurements at the receiver. Any of various measures of link quality can be employed, ranging from simple ones such as the measured probability of packet error on the link to more sophisticated ones such as post-detection signal-quality (PDSQ) statistics [35]. In the protocol considered in this dissertation, however, the link-quality metric is a hard-limited linear function of the received power in the most recent transmission on the link. The function is shown in Fig. 8.12.

The link-quality metric is zero if the received power on the link is greater than an upper threshold $p_{\max}$, and the metric increases linearly to a maximum value of $maxHopsQuality$ as the received power is decreased from $p_{\max}$ to a lower threshold $p_{\min}$. In the examples in this chapter, the two thresholds are determined by the power required to achieve a specified probability of data packet error in the presence of thermal noise alone. The upper received-power threshold $p_{\max}$ results in probability of packet error of 0.05, and the lower received-power threshold $p_{\min}$ results in a probability of packet error of 0.4.

The path-quality metric is intended to reflect whether or not the links in the path have a sufficiently high signal-to-noise ratio to support acceptable link performance. The inclusion of the path-quality metric in the overall path metric $R(P)$ is motivated by the observation that in many instances, the use of the total-contention metric alone results in poor performance for some of the source-destination pairs in the network. The problem arises if the path that is used includes one or more links that are at the limits of viable connectivity. The probability of packet error is high on such links, and consequently the path may represent a poor choice even if it is not congested.

8.4.3 Forwarding with Traffic Splitting

The approach to traffic splitting considered here allows for splitting of the traffic for a source-destination pair among an arbitrary number of paths. In the description and numerical results that follow, however, we restrict the splitting so that a node chooses between at most two next-hop links to a given destination at any given time. Each node maintains a routing table that designates its two best next-hop links to each destination based on the minimum-metric paths to the destination that are determined by the local instance of the Bellman-Ford algorithm.

Traffic for a given destination that originates at or is forwarded by a node is split between its two next-hop links designated in the node's routing table based on an independent weighted coin toss for each packet. (An ordered schedule could be employed instead.) The link that is used with the greater probability is referred to as the *primary link*, and the link used with the lesser probability is referred to as the *secondary link*. The usage probability for the secondary link (referred to as the *splitting probability*) and the designation of the primary and secondary links at the node evolve over time based on changes in the paths metrics calculated at the node.

The designation of the forwarding links and splitting probabilities at each node affects the level of traffic in each individual path in the network. This in turn affects the values of the path metrics calculated at each node, which in turn affects the traffic splitting at the nodes. The resulting negative feedback can be beneficial in balancing traffic among multiple data flows with low mutual coupling. But if the response is too quick it can lead to oscillatory path selections, referred to as *route thrashing*, that are detrimental to network performance. Congestion-avoidance routing is most prone to route thrashing if it does not split traffic, but route thrashing is possible with traffic splitting as well. The traffic-splitting algorithm considered here is designed to exploit the benefits of the feedback mechanism while avoiding route thrashing.

Changes in the primary and secondary designations and the splitting probability for a destination occur only in response to changes in relevant path metrics calculated at the node. Changes in the path metrics in turn occur only in response to some instances of one of three events: receipt of a PROP packet, receipt of an acknowledgment packet, or a change in the value of a node-contention metric $NC_{Tx}(n)$. Any change in path metrics is stored at the node, but the change only prompts a change in the traffic-splitting information for a destination if at least $NumSettling * TimeBwUpdate$ seconds have elapsed since the last change in the splitting probability for the destination.

The changes that occur in the traffic-splitting information are illustrated by considering a circumstance in which initially link A is designated as the primary link from the node to a certain destination, link B is designated as the secondary link to the destination, and some splitting probability p_{curr} between zero and 0.5 is specified for traffic to the destination. The paths associated with the links A and B are denoted as P_A and P_B, respectively.

Suppose a change in path metrics once again results in paths through links A and B, again denoted as P_A and P_B, that are the lowest-cost paths to the destination. The new splitting probability p_{new} is given by

$$p_{new} = \max\{0, \min\{1, \tilde{p}_{new}\}\} \tag{8.8}$$

where

$$\tilde{p}_{new} = \begin{cases} p_{curr} + \text{sgn}(\Delta)\ f_{diff}(|\Delta|), & \text{if } \ p_{curr} + \text{sgn}(\Delta)\ f_{diff}(|\Delta|) \geq minSplit \\ 0, & \text{if } \ p_{curr} + \text{sgn}(\Delta)\ f_{diff}(|\Delta|) < minSplit \end{cases},$$

$$f_{diff}(\Delta) = \begin{cases} 0, & \Delta < \Delta_{min} \\ \frac{(\Delta - \Delta_{min})}{(\Delta_{max} - \Delta_{min})}, & \Delta \geq \Delta_{min} \end{cases},$$

$$\Delta = R(P_A) - R(P_B),$$

$minSplit$ is a parameter between 0 and 0.5, Δ_{max} and Δ_{min} are parameters with $\Delta_{max} > \Delta_{min} > 0$. If $p_{new} > 0.5$, link B is designated as the primary link, link A is designated as the new secondary link, and the splitting probability is set equal to $1 - p_{new}$.

The effect of the change in the splitting probability is to increase the fraction of the traffic for the destination that is routed on the more attractive of the primary and secondary paths (i.e., the one with the lower path metric of the two), and the magnitude of the change is determined by the difference in the path metrics associated with the primary and secondary links. The swap in primary and secondary designations occurs if needed to ensure that the primary link is always used with a probability of at least 0.5. The parameters $NumSettling$, $minSplit$, Δ_{max}, and Δ_{min} are chosen to provide a settling time such that the data flows can reach a quasi-steady state before another routing change is introduced. The choice of values for the parameters provides a tradeoff between responsiveness to changes in traffic load and susceptibility to route thrashing.

Suppose instead that links A and B are the current respective primary and secondary links and that a change in path metrics results in a path through link C, denoted as P_C, that has a lower cost than one or both of P_A and P_B. Two tests are applied based on parameters γ and ν, where $0 \le \gamma \le 0.5$ and $\nu > 0$. If both

$$p_{curr} \le \gamma$$

and

$$R(P_B) - R(P_C) \ge \nu,$$

link C replaces link B as the secondary link for the destination and the new splitting probability is determined by equations (7)-(9) with

$$\Delta = R(P_A) - R(P_C).$$

If either of the two tests fails, however, no change occurs.

Traffic splitting has secondary benefits in addition to balancing traffic among multiple paths and providing for mitigation of route thrashing. Frequent use of both the primary and secondary paths ensures that the nodes along both paths have up-to-date path metrics, since the channel-access acknowledgment packets include current path-metric information. Furthermore, if the congestion is the result of a single source-destination pair that generates more traffic than a single path can support, the traffic demand can be satisfied only by splitting it along multiple paths that have low mutual coupling [45].

8.5 Performance Results

In this section we compare the performance of a CA routing protocol with the performance of a (modified) min-hop routing protocol. The CA routing protocol uses the distributed Bellman-Ford algorithm for distance-vector routing, with path metrics disseminated using the PROP packets described in Chapter 4. Path metrics are formed from the new link traffic-contention metric and link-quality metric described in Sections 8.4.1 and 8.4.2, respectively, and the traffic-splitting forwarding protocol described in Section 8.4.3 is used. The min-hop protocol is a modification of the min-hop version of the LRR protocol described in Chapter 4. The link metric of that protocol is modified here by that the addition of the link-quality metric described in Section 8.4.2. Thus the modified min-hop protocol is biased away from links of marginal link quality in comparison with the min-hop protocol of Section 8.4.2, which provides a fairer comparison with the CA protocol. The physical layer described in Chapter 2 and the MAC-layer protocol described in Chapter 3 are used with both the CA and min-hop routing protocols. The networks using either routing protocol employ an information rate of 100 kbits/s and a chip rate of 4 Mchips/s in the examples in this section. Each network uses three traffic channels.

A node attempts to forward a packet a maximum of six times with either the CA or min-hop protocol, and it drops the packet if all six forwarding attempts are unsuccessful. The CA routing protocol uses the traffic-splitting algorithm of Section 8.4.3 for each forwarding attempt. In contrast, the min-hop routing protocol uses the

Info. rate	100kbps
Chip rate	4Mchips/s
$numAveragingSlots$	150
α	0.15
$minDiff$	0.07
$maxDiff$	2
$minSplit$	0.03
$NumSettling$	5
$thresh_1$	0.75
$thresh_2$	0.9
$hops_1$	1.3
$hops_2$	2.3
$maxHops$	3.3
$maxHopsQuality$	3.3
γ	0.2
ν	0.6

Table 8.3 Parameters used in the examples.

simpler forwarding algorithm described in Chapter 4 with the first four transmission attempts on the primary link and the remaining two on the secondary link. The link metric is incremented by 0.2 for each failed transmission attempt in the min-hop protocol as described in Chapter 4. (The link metric in the CA protocol is not altered due to the success or failure of transmission attempts.) The other parameters of the CA protocol for the examples in this section are given in Table 8.3.

8.5.1 Performance with Five-Node Networks

The benefits of the CA routing protocol using traffic splitting are illustrated by comparing its performance and the performance of the min-hop routing protocol with the performance of an idealized traffic-splitting routing protocol. The performance with the three routing protocols is considered for the two five-node networks shown in Figs. 8.2 and 8.13. In both networks, traffic is generated from node 1 to node 4, and from node 1 to node 2. The two networks differ only in that node 5 uses an omnidirectional antenna in the first network and has three sectors in the second network.

The idealized traffic-splitting routing protocol is referred to as the *ideal-p splitting protocol*. Its use in either of the two example networks results in the following behavior. If a packet originating at node 1 has node 4 as its destination, it traverses the route 1-5-3-4 with a probability of p and it traverses the route 1-3-4 with a probability of $1 - p$. For the two networks and the specified traffic scenario, the ideal-p splitting protocol thus exploits the only available degree of freedom to improve performance compared with min-hop routing. Moreover, we consider its performance under the assumption that it employs the (genie-aided) best choice of a fixed value of p for the generation rates of the two data sources. In this sense it thus exploits the degree of freedom in the optimal manner. It provides a benchmark for the potential performance that can be attained with traffic-splitting routing, and comparison of its performance with the performance of the actual (practical) CA protocol thus allows us to characterize the effectiveness of the actual protocol. (In the two figures discussed below, genie-aided selection of the parameter p in ideal-p splitting for each generation rate is approximated by considering values of p in the set $\{0, 0.1, 0.2, \cdots, 0.9, 1\}$ and selecting the one that results in the best performance for that generation rate.)

The end-to-end completion percentage is shown in Fig. 8.14 for the network of Fig. 8.2 and in Fig. 8.15 for the network of Fig. 8.13. The CA routing protocol results in significantly better end-to-end completion percentages than the min-hop protocol in either network. For example in the first network, CA routing supports a generation rate of 74 packets/s with a 90% completion percentage, whereas the same completion percentage is achieved with min-hop routing only for a much lower generation rate. Similarly, the CA protocol permits a much higher generation rate than does min-hop routing in the second network under the constraint of a 90% completion percentage. Moreover, the performance of the CA protocol is nearly identical to that of ideal-p splitting in the first network, and it is much closer to that of ideal-p splitting than is min-hop routing in the second network. The generation rate supported with a 90% completion percentage for the CA protocol is only about 87% of the 104 packets/s generation rate for ideal-p splitting in the latter network, however. The difference is a consequence of some route thrashing in the CA protocol that results in less than optimal utilization of the 1-5-3-4 path.

8.5.2 Performance with Thirty-Node Networks

In this subsection, we compare the performance of the two routing protocols for each of eight different scenarios involving networks of 30 nodes. The scenarios are differentiated by the total number of nodes with three-sector directional antennas, the number of *advantaged nodes* (defined below), and the number of advantaged nodes with directional antennas. The distinguishing characteristics of the scenarios are shown in Table 8.4. The networks in scenarios 4-8 (which include several nodes with directional antennas) are referred to in what follows as *heterogeneous networks*. Six randomly generated network topologies are considered for each scenario.

Two of the scenarios (scenarios one and four) do not include any advantaged nodes. Scenario one includes only nodes with omnidirectional antennas, and scenario four includes twelve nodes with directional antennas. Each of the six topologies for scenario one is generated by randomly placing each of the 30 nodes in an 8 km by 8 km square region according to a uniform distribution. Each topology for scenario four is derived from one of the topologies for scenario one by choosing twelve of the 30 nodes at random to have three-sector directional antennas. (Each node with directional antennas has a fixed sector orientation in azimuth.) The remaining nodes have omnidirectional antennas. For each topology in either scenario, eight source-destination pairs are selected at random with the constraint that the distance between the source node and the destination node is greater than 4.8 km and that no node can participate in more than one source-destination pair. The source nodes generate packets according to independent Poisson processes with the same rate.

In many circumstances involving wireless ad hoc networks, some nodes are positioned so that their links with other nodes enjoy better propagation conditions than most of the links in the network. A node in this situation is referred to as an advantaged node. As an example, a node that is located in an elevated position (such as on a hill) is likely to exhibit the characteristics of an advantaged node. In each of scenarios three through eight, the network includes either one or two advantaged nodes. The propagation characteristics for each advantaged node are modeled in the examples by employing a path-loss coefficient of 2.7 rather than 3.0 in Equation (5.1) for transmissions in either direction on each link involving the advantaged node.

The six topologies for each scenario that includes one advantaged node are derived ultimately from the randomly generated topologies for scenario one. For example, each topology for scenario two is obtained from a topology of scenario one by designating one randomly selected node as the (sole) advantaged node. The same node selection from the same topology of scenario one is employed as the three-sector advantaged node for a topology of scenario three. The same approach is used to obtain the six topologies for scenarios five and six from the topologies of scenario four (which are derived in turn from the topologies for scenario one). The topologies of scenario seven are generated by a random placement of omnidirectional nodes, three-sector nodes, and advantaged nodes that is independent of the placement used in scenarios one through six. Each topology of scenario eight is obtained by substituting three-sector advantaged nodes in place of the two omnidirectional advantaged nodes in a topology of scenario seven.

For some of the scenarios with advantaged nodes, the resulting number of nodes with directional antennas can differ among the six topologies. If a node with directional antennas in a topology of scenario four is selected as the (omnidirectional) advantaged node in scenario five, the resulting topology in scenario five will have only eleven nodes with directional antennas. If instead an omnidirectional node is selected in a topology of scenario four, the resulting topology in scenario five will have twelve nodes with directional antennas. In a similar manner, the number of nodes with directional antennas can be either eleven and twelve in the individual topologies of scenario six. The number of nodes with directional antennas can range between ten and twelve in the topologies of scenario seven and between twelve and fourteen in the topologies of scenario eight.

In some applications of ad hoc networks, performance is considered acceptable if a target percentage of the packets generated in the network are received successfully at their respective destinations (i.e., a target aggregate end-to-end completion percentage is achieved by the network). In such instances, it is desirable that the target completion percentage is achieved for the highest possible generation rate (and hence the greatest possible throughput). Motivated by this requirement, one measure of performance we consider for each topology of a given scenario is the (maximum) aggregate generation rate that is achievable for a specified aggregate completion percentage. For a specified aggregate completion percentage, the achievable aggregate generation rate will differ in general among the six topologies for a given scenario and routing protocol. Thus we consider both the achievable aggregate generation rate for each topology and the average of the achievable aggregate generation rates over the six topologies. The latter is referred to as the *average generation rate* in the remainder of the chapter, and it is given as the average generation rate for a specified (aggregate) completion percentage. The average end-to-end delay of successfully received packets is also considered for each topology using the generation rate that achieves the specified completion percentage for that topology.

In other applications of ad hoc networks, such as some mobile networks, poorer than expected short-term performance can be tolerated as long as the average performance over an extended period of time meets a specified performance criterion. In such instances, each short-term configuration of the nodes constitutes a distinct network topology, and the long-term average performance corresponds to an average over the various topologies. Motivated by this requirement, the second measure of performance we consider for a given scenario is the average over the six topologies of the aggregate end-to-end completion percentage that results for a specified aggregate

Scenario	Number of Directional Antenna Nodes	Number of Advantaged Nodes	Advantaged Node Antenna
1	0	0	N/A
2	0	1	omni-directional
3	1	1	three-sector directional
4	12	0	N/A
5	11 or 12	1	omni-directional
6	12 or 13	1	3-sector directional
7	10-12	2	omni-directional
8	12-14	2	3-sector directional

Table 8.4 Scenarios for the 30-node networks.

generation rate. This measure is referred to as the *average completion percentage* in the remainder of the chapter, and it is given as the average completion percentage for a specified (aggregate) generation rate. The average delay over all six topologies is also considered for the specified generation rate. (Note that for a specified generation rate, the completion percentage will differ in general among the six topologies for a given scenario and routing protocol.)

8.5.2.1 Thirty-Node Networks with No Advantaged Nodes

Heterogeneous Network with No Advantaged Nodes

The performance of the two routing protocols in a heterogeneous network with no advantaged nodes is illustrated by considering scenario four in Table 8.4. The average generation rate over the six topologies is shown for this scenario in Fig. 8.16 as a function of the per-topology completion percentage. Clearly the CA routing protocol is able to achieve any specified completion percentage at a significantly higher generation rate than the min-hop routing protocol. For example, min-hop routing results in a 90% completion percentage for each of the six topologies with an average generation rate of 111.7 packets/s over the six topologies. In contrast, the average generation rate is 156.2 packet/s for the same completion percentage if CA routing is used, an improvement of 40% in the generation rate (and the throughput). If the performance criterion is a 95% completion percentage instead, it is satisfied for an average generation rate of 142.2 packet/s with the CA protocol but only with an average generation rate of 102.9 packets/s with the min-hop protocol. Thus the CA protocol provides an improvement of 38.2% improvement in throughput under this criterion.

Not only does the CA protocol result in better average performance over the six topologies, it results in better performance for each topology. This is illustrated in Table 8.5, which shows the generation rate at which 90% and 95% completion percentages are achieved for each of the topologies. Suppose the performance criterion is that the network must achieve a 90% completion percentage, for example. The CA routing protocol satisfies the criterion for a greater throughput than the min-hop routing protocol with each topology, and the improvement in performance is 101% for one of the topologies, 67% for another, 29-35% for three others, and 13% for the sixth. If instead the performance criterion is that the network must achieve a 95% completion percentage, the CA routing protocol results in an increase in the throughput of 87% for one of the topologies, 59% for another, 30-35% for three others, and 16% for the sixth. Thus the degree of performance improvement with CA routing is substantial but highly dependent on the topology.

The use of the CA protocol results in consistently better throughput for each source-destination pair as well for each of the six topologies in scenario four. The end-to-end throughput for each individual source-destination pair is shown in Figs. 8.18-8.23 for the six topologies and both routing protocols. The generation rate for each topology results in a 95% completion percentage for the topology. (Note that the completion percentage differs for different source-destination pairs within a given

Topology	Min-Hop Routing		CA Routing	
	95%	90%	95%	90%
1	122.1	134.5	160.2	174.1
2	60.5	64.75	96.3	107.9
3	142.2	155.6	164.2	176
4	125.2	137.8	169.9	185.4
5	90.5	96.4	119	130
6	77.1	81.2	144.1	163.6
Average over 6 topologies	102.9	111.7	142.2	156.2

Table 8.5 Per-topology generation rate (in packets/s) for 95% and 90% completion percentages in scenario four.

topology.) The throughput for every source-destination pair is greater with the CA protocol than with the min-hop protocol. The increase in the throughput ranges from a few percent for some source-destination pairs to more than two-fold for other pairs. For some topologies (such as the first topology), there is a wide range of values for the per-pair increase in throughput. For other topologies (such as the sixth one), the increase in throughput is similar for all source-destination pairs.

The CA protocol also results in better delay performance than the min-hop protocol if a specified completion percentage is required. The delay performance is illustrated in Table 8.6, which gives the average end-to-end delay for each of the six topologies for the respective generation rates that result completion rates of 90% and 95% for each topology. For five of the six topologies, the use of the CA protocol results in better delay performance than use of the min-hop protocol for a specified completion percentage, even though the throughput for the topology is much greater with the CA protocol. The CA protocol results in a larger average delay than the min-hop protocol for the fourth topology. Thus for the latter topology, the choice between the two routing protocols provides a tradeoff between throughput and average delay for a given completion percentage.

The CA routing protocol also results in superior performance with respect to the average network performance for a specified generation rate. This is illustrated by considering the average completion percentage (over the six topologies) that is achieved for a given generation rate. An average completion percentage of 95% is achieved with the min-hop protocol if the generation rate for each topology is 79 packets/s. In contrast, the same average completion percentage is achieved with the CA protocol if the generation rate for each topology is 122.2 packets/s. Thus the CA protocol permits a 54.6% higher generation rate than the min-hop protocol under this criterion. An average completion percentage of 90% is achieved with the min-hop and CA protocols for respective generation rates of 90.6 packets/s and 144.9 packets/s, so that the CA protocol permits a 60% higher generation rate than the min-hop protocol.

The CA protocol also results in better delay performance than the min-hop protocol for a specified generation rate. This is illustrated in Fig. 8.17 in which the average end-to-end delay over the six topologies is shown as a function of the generation rate. Clearly the average delay is much less for any given generation rate with the CA protocol than with the min-hop protocol. Moreover, the generation rates that result in comparable average completion percentages for the two protocols results in similar delay performance. For example, respective generation rates for the CA and min-hop protocols are 79 packets/s and 122.2 packet/s (resulting in an average completion percentage of 95% with each protocol), yet they result in almost identical values of average delay. If the generation rates are considered that result in an average completion percentage of 90% with each protocol, the average delay is only slightly larger in the network using the CA protocol than in the network using the min-hop protocol. Thus the CA protocol achieves a given average completion percentage with a much higher throughput and a similar average delay compared with the min-hop protocol.

All-Omnidirectional Network with No Advantaged Nodes

The performance of the two routing protocols in a network with no advantaged nodes and no directional antennas is illustrated by considering scenario one in Table

Topology	Min-Hop Routing		CA Routing	
	95%	90%	95%	90%
1	969.7	1254.7	840.35	1165.7
2	1515.7	1670.6	1025.6	1350.8
3	1089	1517	909.9	1199
4	809.5	1117.4	1003	1291
5	1052.4	1188.5	913.3	1201.4
6	981	1360	480	950

Table 8.6 Per-topology average end-to-end delay (in ms) for per-topology 95% and 90% completion percentages in scenario four.

8.4. The average generation rate over the six topologies is shown for this scenario in Fig. 8.24 as a function of the per-topology completion percentage. The benefit provided by the use of directional antennas is apparent for either routing protocol by comparing the results in Fig. 8.24 with those in Fig. 8.16. For example, min-hop routing results in a 90% completion percentage for each of the six topologies with an average generation rate of 38.9 packets/s in scenario one in which all the nodes use omnidirectional antennas. In scenario four in which twelve of the nodes have three sectors each, in contrast, the average generation rate is 111.7 packets/s. The presence of directional antennas at the twelve nodes thus results in an increase of 187% in performance. Similarly, the average generation rate in scenario one is 52.1 packets/s if the CA protocol is used, whereas the average generation rate with the same protocol is 156.2 packets/s in scenario four. Thus the presence of the directional antennas improves the performance by 200% with the CA routing protocol.

The CA routing protocol exploits the presence of directional antennas more effectively than the min-hop routing protocol, and so conversely, the performance benefits of CA routing over min-hop routing are more modest in an network with all omnidirectional antennas than if some nodes have directional antennas. Yet the CA protocol results in better performance than the min-hop protocol even if there are no directional antennas present, as is illustrated in Fig. 8.24. For example, min-hop routing results in a 90% completion percentage for each of the six topologies with an average generation rate of 38.9 packets/s over the six topologies. In contrast, the average generation rate is 52.1 packet/s for the same completion percentage if CA routing is used, an improvement of 33.9% in the generation rate (and the throughput). If the performance criterion is a 95% completion percentage instead, it is satisfied for an average generation rate of 47.7 packet/s with the CA protocol but only with an average generation rate of 37.4 packets/s with the min-hop protocol. Thus the CA protocol provides an improvement of 27.5% improvement in throughput under this criterion.

There is substantial variation in the performance from other topology to another with either protocol in scenario one, as was also seen in the results for scenario four. The CA protocol results in comparable or better performance for each topology than the min-hop protocol, but as observed for scenario four, the performance improvement it provides differs substantially for different topologies in scenario one. This is illustrated in Table 8.7, which shows the generation rate at which 90% and 95% completion percentages are achieved for each of the topologies. Suppose the performance criterion is that the network must achieve a 90% completion percentage, for example. The CA routing protocol satisfies the criterion for a greater throughput than the min-hop routing protocol with each topology, and the improvement in performance is 105% for one of the topologies, approximately 50% for each of two others, and only about 17% for each of two others. For the remaining topology (the fourth one in the table), the two routing protocols result in comparable performance.

The superior delay performance of the CA protocol observed in the heterogeneous network is even more pronounced if the network includes only nodes with omnidirectional antennas. The per-topology average delay is illustrated in Table 8.8, which gives the average end-to-end delay for each of the six topologies with the respective generation rates that result in average completion rates of 90% and 95% for each topologies. For all six topologies, the use of the CA protocol results in better delay performance than use of the min-hop protocol, even though the throughput is much greater with the CA protocol. The CA protocol results in a dramatic reduction in the average delay for some topologies, and it results in a modest reduction for other topologies.

The CA routing protocol also results in superior performance with respect to the average network performance for a specified generation rate if every node has an omnidirectional antenna. An average completion percentage of 95% over the six topologies in scenario one is achieved with the min-hop protocol if the generation rate for each topology is 29.5 packets/s. In contrast, the same average completion percentage is achieved with the CA protocol if the generation rate for each topology is 38.8 packets/s. Thus the CA protocol permits a 31.5% higher generation rate than the min-hop protocol under this criterion. An average completion percentage of 90% is achieved with the min-hop and CA protocols for respective generation rates of 31.4 packets/s and 45.7 packets/s, so that the CA protocol permits a 45.5% higher generation rate than the min-hop protocol.

The superior delay performance of the CA protocol is also seen for a specified generation rate in the network containing only nodes with omnidirectional antennas. This is illustrated in Fig. 8.25 in which the average end-to-end delay over the six topologies for scenario one is shown as a function of the generation rate. The average delay is much less for any given generation rate with the CA protocol than with the min-hop protocol. Moreover, the generation rates that result in comparable average completion percentages for the two protocols result in much better delay performance with the CA protocol than with the min-hop protocols. (This is in contrast with the circumstance for scenario four, in which the two protocols result in similar delay performance under this criterion.) For example, respective generation rates for the CA and min-hop protocols of 38.8 packets/s and 29.5 packet/s result in an average completion percentage of 95% with each protocol. The average delay is 1675 ms with the min-hop protocol but only 875 ms with the CA protocol. Thus under the criterion of a 95% average completion percentage, the CA protocol results in a 31.5% greater throughput and a 48% lower average delay than the min-hop protocol. Similarly, if the respective generation rates are chosen to result in a 90% average completion percentage with each protocol, the CA protocol results in a 45.5% greater throughput and a 29% lower average delay than the min-hop protocol.

8.5.2.2 Thirty-Node Networks with One or Two Advantaged Nodes

The ability of the two routing protocols to accommodate the presence of advantaged nodes is examined by considered several scenarios involving one or two advantaged nodes. The lower path-loss coefficient for links involving an advantaged node results in links that are viable at a greater distance than is possible for links involving other nodes. If both the advantaged node and the other node in a link use an omnidirectional antenna, for example, the signal-to-noise ratio for an information bit is 11.5 dB at a distance of 4530 m in the absence of multiple-access interference. (In comparison, the same signal-to-noise ratio is only achieved at a distance of 1350 m if neither node is advantaged.) As a result, each advantaged node in our examples is able to sustain a viable link with every other node in the network. Thus the advantaged node is able to serve as the relay node in a two-hop path between any pair of nodes in the network.

Heterogeneous Network with One Omnidirectional Advantaged Node

It is apparent from the observations above that if there is one advantaged node and min-hop routing is employed, each source-destination pair is likely to utilize the advantaged node heavily. Consequently, the advantaged node will become heavily congested. Moreover, the extended link range of the advantaged node results in suppression of other transmissions over a wide area (via its CTS transmission) whenever it is receiving data. It also results in interference with other transmissions over a wide area whenever the advantaged node is transmitting. The consequences of this are seen by considering the performance of the routing protocols in scenario five

Topology	Min-Hop		CA routing	
	95%	90%	95%	90%
1	39	40.6	55.2	60.1
2	28.3	29.3	32.8	34.3
3	28.8	29.9	47.8	59.6
4	56.6	58.8	56.2	60.13
5	28.6	29.5	42.9	45.7
6	43.5	45.7	51	52.3
Average over 6 topologies	37.4	38.9	47.7	52.1

Table 8.7 Per-topology generation rate (in packets/s) for 95% and 90% completion percentages in scenario one.

	Min-Hop		CA routing	
	95%	90%	95%	90%
Topology 1	2140	2590	1205	1910
Topology 2	3250	3445	2055	2610
Topology 3	2920	3365	485	2053
Topology 4	2065	2525	1865	2325
Topology 5	3055	3250	1760	2220
Topology 6	1960	2275	1865	1890

Table 8.8 Per-topology average end-to-end delay (in ms) for per-topology 95% and 90% completion percentages in scenario one.

(in which the heterogeneous network includes one advantaged node using an omnidirectional antenna). The performance is illustrated in Fig. 8.26, which shows the average aggregate generation rate over the six topologies for this scenario as a function of the per-topology completion percentage. Each topology is identical to the correspondingly numbered topology of scenario four, except for the fact that one node is designated as advantaged with an omnidirectional antenna. The advantaged node acts as one of the eight source nodes in topologies one, two, and six; it acts as one of the eight destination nodes in topology three; and it serves only as a relay node in topologies four and five.

The presence of the advantaged node clearly has a detrimental effect on the performance of the network. Moreover, the impact on the performance is severe if min-hop routing is employed. It results in a 90% completion percentage for each of the six topologies with an average generation rate of only 43 packets/s. In contrast, an average generation rate of 111.7 packets/s is achieved under the same performance criterion with min-hop routing in scenario four (in which there is no advantaged node). Thus the presence of the advantaged node causes a 61.5% degradation in performance. The decrease in the per-topology generation rate ranges from 49% (for topology two) to 70% (for topology three), as seen by comparing Tables 8.5 and 8.9. Similarly, the presence of the advantaged node results in a 60.3% reduction in the average generation rate (from 102.9 packet/s to 40.9 packets/s) if the completion percentage for each topology is 95%.

The impact of the advantaged node on the network's performance is much less pronounced if the network employs the CA routing protocol, since the protocol effectively exploits alternative (non-minimal hop count) paths that do not utilize the advantaged node. The use of alternative paths has the direct benefit of relieving congestion at the advantaged node. This in turn has the indirect benefit of reducing the long-range interference and the wide-area suppression of other transmissions due to activity at the advantaged node. The result is illustrated by considering the average generation rate if the performance criterion is a 90% completion percentage for each topology. In this instance, the average generation rate is 156.2 packet/s in scenario four but only 130.6 packets/s in scenario five. Thus the presence of the advantaged node results in a relatively small 16.4% degradation in performance. Similarly, the average generation rate is reduced from 142.2 packets/s to 115 packets/s (a decrease of 19.9% if a 95% completion percentage is required for each topology.) The decrease in the per-topology generation rate ranges between 6% (for topology three) and 25% (for topology six).

The more effective response of the CA protocol to the presence of the advantaged node results in a large difference in the performance with the two protocols. If the performance criterion is a 90% completion percentage for each topology, the CA protocol results in 203.3% higher average throughput than the min-hop protocol. The increase in throughput differs substantially for the individual topologies (ranging from 141% for topology one to 262% for topology six), but the gains are large in each instance. If instead the required completion percentage for each topology is 95%, the average throughput with CA protocol is 181.5% better than with the min-hop protocol and the increase is large for each topology.

The differing impact of the advantaged node on the throughput with the two protocols is also reflected in its impact on the delay performance with the protocols, which is shown in Table 8.10. The presence of the advantaged node results in a large increase in the per-topology average delay if the min-hop protocol is used. The per-topology average delay at a 90% completion percentage for each topology is increased by an amount ranging from 30% (for topology one) to 156% (for topology five). If the completion percentage for each topology is 95%, the increase in the average delay ranges from 36% (for topology one) to 139% (for topology six). In contrast, the delay performance of the network with the CA protocol is better in the network with the advantaged node for most of the topologies if the completion percentage for each topology is 90%. Five of the topologies experience a decrease in the average delay, and the decrease ranges from 5% (topology two) to 30% (topology one). Only topology three experiences an increase in the average delay, and the increase is 38%. If the per-topology completion percentage is 95%, the average delay also decreases for five of the topologies if the advantaged node is present. The decreases range between 15% (for topology one) and 52% (for topology five). Only topology six experiences an increase under this performance criterion, and the increase is 44%.

As a result of the differing impact of the advantaged node on the delay performance with the two protocols, the delay performance with the CA protocol is markedly better than the performance with the min-hop protocol for each of the six topologies. If the generation rate used with each topology and routing protocol is chosen to result in a 90% completion percentage for that topology and protocol, the average delay with CA routing is between 21% and 73% lower than the corresponding average delay with min-hop routing. These extremes occur for topologies three and six, respectively. If the performance criterion is a 95% completion percentage for each topology, the average delay is between 56% lower (for topology three) and 79% lower (for topology two) with the CA protocol than with the min-hop protocol. Thus the ability of the CA protocol to exploit non-minimum-hop paths effectively results in both much higher throughput and much lower delay than with the min-hop protocol. In topology six of scenario five, for example, the throughput is 262% greater and the average delay is 73% lower with CA routing than with min-hop routing if a completion percentage of 90% is achieved with each protocol.

The CA routing protocol also results in superior performance with respect to the average network performance for a specified generation rate if an advantaged node is present. An average completion percentage of 95% over the six topologies is achieved with the min-hop protocol if the generation rate for each topology is 35.8 packets/s. In contrast, the same average completion percentage is achieved with the CA protocol if the generation rate for each topology is 96.4 packets/s. Thus the CA protocol permits a generation rate almost three times that permitted by the min-hop protocol under this criterion. An average completion percentage of 90% is achieved with the min-hop and CA protocols for respective generation rates of 33.2 packets/s and 116.4 packets/s, so that the CA protocol permits more than three and one-third times the generation rate with the min-hop protocol.

The CA protocol also results in better delay performance than the min-hop protocol in scenario five for a specified generation rate. This is illustrated in Fig. 8.27 in which the average end-to-end delay over the six topologies is shown as a function of the generation rate. Clearly the average delay is much less for any given generation rate with the CA protocol than with the min-hop protocol. Moreover, the generation rates that result in comparable average completion percentages for the two protocols result in much better delay performance with the CA protocol than with the min-hop protocols. (This is again in contrast with the circumstance for scenario four, in which the two protocols resulted in similar delay performance under this criterion.) For example, respective generation rates for the CA and min-hop protocols of 33.2 packets/s and 96.4 packet/s result in an average completion percentage of 95% with each protocol. The average delay is 1535 ms with the min-hop protocol but only 540 ms with the CA protocol. Thus under the criterion of a 95% average completion percentage, the CA protocol results in a 190% greater throughput and a 65% lower average delay than the min-hop protocol. Similarly, if the respective generation rates are chosen to result in a 90% average completion percentage with each protocol, the CA protocol results in a 225% greater throughput and a 59% lower average delay than the min-hop protocol.

Heterogeneous Network with One Three-Sector Advantaged Node

The performance of the network improves for most topologies with either routing protocol if the advantaged node has a three-sector directional antenna instead of an omnidirectional antenna, but the performance is still poorer in most instances than if there were no advantaged nodes. This is illustrated in Fig. 8.28, which shows the average aggregate generation rate over the six topologies for scenario six as a function of the per-topology completion percentage. Each topology is identical to the correspondingly numbered topology of scenario five, except for the fact that the sole advantaged node has a three-sector directional antenna.

The use of a three-sector advantaged node in the network with min-hop routing results in an average generation rate of 58 packets/s if the specified completion percentage is 90%, which represents a 35% improvement in the performance compared with the use of an omnidirectional antenna at the advantaged node. But the performance is still 48% poorer than in scenario four (in which there are no advantaged

	Min-Hop		CA routing	
	95%	90%	95%	90%
Topology 1	59.5	63.3	139.1	152.4
Topology 2	31.3	33.1	81.6	90.7
Topology 3	44.27	46.6	142.8	165.1
Topology 4	47.3	49.9	128.5	153.8
Topology 5	30.3	31.7	92.2	99.8
Topology 6	32.4	33.7	105.7	122
Average over 6 topologies	40.85	43.05	115	130.6

Table 8.9 Per-topology generation rate (in packets/s) for 95% and 90% completion percentages in scenario five.

	Min-Hop		CA routing	
	95%	90%	95%	90%
Topology 1	1315	1615	710	815
Topology 2	2805	2955	810	1287
Topology 3	1640	2105	720	1645
Topology 4	1645	1960	480	10455
Topology 5	2370	3045	770	1030
Topology 6	2345	2910	685	775

Table 8.10 Per-topology average end-to-end delay (in ms) for per-topology 95% and 90% completion percentages in scenario five.

nodes). Four of the six topologies exhibit better performance under this criterion if the advantage node has a three-sector directional antenna instead an omnidirectional antenna, and the other two topologies exhibit poorer performance. The change in the per-topology generation rate ranges from -34% (for topology four) to 100% (for topology two), as seen by comparing Tables 8.9 and 8.11. Similarly, the use of a directional antenna at the advantaged node results in a 37.5% increase in the average generation rate if the completion percentage for each topology is 95%, and the per-topology generation rates change by between -27% and 106% under that performance criterion. The average generation rate is still 47% less than in scenario four, however.

The impact of sectorization at the advantaged node on the network's performance is similar if the network employs the CA routing protocol. The average generation rate is increased by 30% and 3% for respective per-topology completion percentages of 90% and 95%. But the performance under the two respective criteria is still slightly worse than if there were no advantaged nodes.

The superiority of the performance of the CA protocol in scenario six is nearly as great as it is in scenario five. If the performance criterion is a 90% completion percentage for each topology, the CA protocol results in 157% higher average throughput than the min-hop protocol. The increase in the per-topology throughput ranges from 84% for topology two to 278% for topology five. If instead the required completion percentage for each topology is 95%, the average throughput with the CA protocol is 145% better than with the min-hop protocol and the increase is large for each topology. The use of the CA protocol results in consistently better throughput for each source-destination pair as well for each of the six topologies in scenario six. This is seen by considering the end-to-end throughput for each individual source-destination pair and both routing protocols, which is shown in Figs. 8.30-8.35 for the six topologies and a 95% completion percentage for each topology. (As noted before, the completion percentage may differ for different source-destination pairs within a given topology.)

The per-topology average delay in scenario six is shown in Table 8.12 for completion percentages of 90% and 95% for each topology. The introduction of sectorization at the advantaged node results in better delay performance for some topologies and poorer delay performance for others, and this occurs with either routing protocol. The delay performance with min-hop routing is much poorer than occurs without an advantaged node. The delay performance with CA routing is similar to or better than the delay that would occur without an advantaged node, in contrast, for each topology and either the 90% completion percentage or the 95% completion percentage.

In scenario six the presence of the advantaged node once again results in markedly better delay performance with the CA protocol than with the min-hop protocol for each of the six topologies. If the generation rate used with each topology and routing protocol is chosen to result in a 90% completion percentage for that topology and protocol, the average delay with CA routing is between 8% and 71% lower than the corresponding average delay with min-hop routing. These extremes occur for topologies four and five, respectively. If the performance criterion is a 95% completion percentage for each topology, the average delay is between 42% lower (for topology two) and 81% lower (for topology five) with the CA protocol than with the min-hop protocol. Thus the superior ability of the CA protocol to exploit non-minimum-hop paths effectively is equally beneficial whether the advantaged node has an omnidirectional antenna or a three-sector directional antenna. In topology six of scenario five, for example, the throughput is 278% greater and the average delay is 71% lower with CA routing than with min-hop routing if a completion percentage of 90% is achieved with each protocol.

The CA routing protocol also results in superior performance with respect to the average network performance for a specified generation rate if a three-sectored advantaged node is present. An average completion percentage of 95% over the six topologies is achieved with the min-hop protocol if the generation rate for each topology is

39.6 packets/s. In contrast, the same average completion percentage is achieved with the CA protocol if the generation rate for each topology is 122.7 packets/s. Thus the CA protocol permits a generation rate three times that permitted by the min-hop protocol under this criterion. An average completion percentage of 90% is achieved with the min-hop and CA protocols for respective generation rates of 43.7 packets/s and 139.7 packets/s, so that the CA protocol permits more than three and one-fourth times the generation rate with the min-hop protocol.

The CA protocol also results in better delay performance than the min-hop protocol for a specified generation rate. This is illustrated in Fig. 8.29 in which the average end-to-end delay over the six topologies is shown as a function of the generation rate. Once again, the average delay is much less for any given generation rate with the CA protocol than with the min-hop protocol. Moreover, the generation rates that result in comparable average completion percentages for the two protocols results in much better delay performance with the CA protocol than with the min-hop protocols. Thus the CA routing protocol results in much greater throughput with a much lower average delay in scenario six for a specified generation rate.

Heterogeneous Network with Two Advantaged Nodes

The introduction of a second advantaged node into the network results better performance with either routing protocol than occurs if there is only a single advantaged node. This is illustrated by considering the performance in scenarios seven and eight. In both scenarios, the network contains two advantaged nodes, which have omnidirectional antennas in scenario seven and three-sector directional antennas in scenario eight. (The randomly generated topologies for these two scenarios do not correspond to the topologies considered in scenarios one through six, which precludes per-topology comparisons of the results for scenarios one through six with the results for scenarios seven and eight.)

The second advantaged node is of particular benefit in the network using min-hop routing. It presence guarantees a second two-hop path for most source-destination

	Min-Hop		CA routing	
	95%	90%	95%	90%
Topology 1	43.7	46.6	134.8	147.6
Topology 2	63.9	67.1	117.4	123.6
Topology 3	73.3	79.6	171.2	187.1
Topology 4	75	77.7	151.5	175.6
Topology 5	32.3	33.3	116.6	126.1
Topology 6	41.4	42.8	119.9	130.7
Average over 6 topologies	54.9	57.9	135.3	148.5

Table 8.11 Per-topology generation rate (in packets/s) for 95% and 90% completion percentages in scenario six.

	Min-Hop		CA routing	
	95%	90%	95%	90%
Topology 1	1800	2135	710	915
Topology 2	1095	1495	625	835
Topology 3	1335	1380	705	1068
Topology 4	1080	1175	615	1075
Topology 5	3025	3140	570	900
Topology 6	1745	2315	520	805

Table 8.12 Per-topology average end-to-end delay (in ms) for per-topology 95% and 90% completion percentages in scenario six.

pairs (in addition to the two-hop path using the first advantaged node as the relay). Thus it alleviates some of the congestion that min-hop routing produces at the sole advantaged node of scenarios five and six. This is illustrated in Tables 8.13 and 8.14 which show the per-topology generation rate for specified per-topology completion percentages in scenarios seven and eight, respectively. A comparison of Tables 8.9 and 8.13 shows that the average performance with min-hop routing is improved for either completion percentage if a second omnidirectional advantaged node is introduced into scenario five. The average generation rate is increased from 43 packets/s to 75 packets/s (a 74% increase) for a 90% completion percentage, and it is increased from 40 packets/s to 70 packets/s (a 75% increase) for a 95% completion percentage.

A comparison of Tables 8.11 and 8.14 shows that the introduction of a second three-sector advantaged node into scenario six results in an increase of 66% in the average generation rate for a 90% completion percentage and an increase of 60% for a 95% completion percentage. As was previously observed for the network containing a single advantaged node, the average performance of min-hop routing in the presence of two advantaged nodes is better if the two nodes have three-sector directional antennas instead of omnidirectional antennas and it is better for five of the six topologies. In spite of the benefits provided by the second advantaged node, the average performance of the min-hop protocol in either scenario seven or scenario eight is still poorer than in a network without an advantaged node.

The network using CA routing also benefits from the addition of a second advantaged node, though the benefit is much smaller than for the network using min-hop routing. The introduction of a second omnidirectional advantaged node into scenario five results in an increase of 8% in the average generation rate for a 90% completion percentage and an increase of 7% for a 95% completion percentage, as shown by a comparison of Tables 8.9 and 8.13. The introduction of a second three-sector advantaged node into scenario six results in an increase of 13% in the average generation rate for a 90% completion percentage and an increase of 10% for a 95% completion percentage, as shown by a comparison of Tables 8.11 and 8.14. The average performance of CA routing in the presence of two advantaged nodes is better if the two nodes have three-sector directional antennas instead of omnidirectional antennas. Furthermore with CA routing, the average generation rate at either specified completion percentage is somewhat greater with two three-sector advantaged nodes than without any advantaged nodes. Thus the use of CA routing in scenario eight results in the best average performance under these criteria among all of the eight scenarios considered.

The performance of CA routing is better than the performance of min-hop routing if there are two advantaged nodes present. The percentage difference in their performance is smaller than in the presence of only one advantaged node, but it is larger than if there were no advantaged nodes present. If the performance criterion is a 90% completion percentage for each topology, the CA protocol results in an 89% higher average throughput than the min-hop protocol in scenario seven. The increase in the per-topology throughput ranges from 38% for topology one to 139% for topology two. If instead the required completion percentage for each topology is 95%, the average throughput with CA protocol is 76% better than with the min-hop protocol and the increase is large for each topology. In scenario eight, the CA protocol results in 77% higher average throughput than the min-hop protocol if the performance criterion is a 90% completion percentage for each topology. The increase in the per-topology throughput ranges from 42% for topology three to 120% for topology four. If instead the required completion percentage for each topology is 95%, the average throughput with CA protocol is 69% better than with the min-hop protocol, and once again, the increase is large for each topology.

All-Omnidirectional Network with One Advantaged Node

The effect of introducing an advantaged node into the network with all omnidirectional antennas depends on whether or not the advantaged node uses sectorization. This is illustrated by considering the performance in scenarios two and three. In

	Min-Hop		CA routing	
	95%	90%	95%	90%
Topology 1	113.8	127.3	160	176.1
Topology 2	51.3	54	105.2	128.9
Topology 3	64.3	68.1	129.2	145.3
Topology 4	69.6	73.6	140.9	155.7
Topology 5	52.7	55.4	96	119.7
Topology 6	66.4	71	105.5	122.3
Average over 6 topologies	69.7	74.9	122.9	141.4

Table 8.13 Per-topology generation rate (in packets/s) for 95% and 90% completion percentages in scenario seven.

	Min-Hop		CA routing	
	95%	90%	95%	90%
Topology 1	90.30	94.8	157.4	191.1
Topology 2	110.1	120	148.2	162.4
Topology 3	105.8	114.4	145.9	162.5
Topology 4	75.6	81.3	165.9	178.7
Topology 5	66.5	72	121.5	135.4
Topology 6	81.2	90.7	155.2	185
Average over 6 topologies	88.2	95.5	149	169.2

Table 8.14 Per-topology generation rate (in packets/s) for 95% and 90% completion percentages in scenario eight.

both scenarios, the network contains a single advantaged node, which has an omnidirectional antenna in scenario two and a three-sector directional antenna in scenario three. Each topology is identical to the correspondingly numbered topology of scenario four, except for the fact that one node is advantaged. The same node serves as the advantaged node in both scenario two and scenario three. It acts as one of the eight source nodes in topologies one, two, and six; it acts as one of the eight destination nodes in topology three; and it serves only as a relay node in topologies four and five.

The presence of the advantaged node with an omnidirectional antenna has a detrimental effect on the performance of most of the topologies with min-hop routing, though the average percentage decrease in performance is much less than between scenarios four and six. This is illustrated by comparing the results in Table 8.7 with the results in Table 8.15 which shows the per-topology generation rate for specified per-topology completion percentages in scenario two. The introduction of the advantaged node results in a decrease in the average generation rate from 39 packets/s to 30 packets/s (a 23% decrease) for a 90% completion percentage. The per-topology generation rate is increased by 7% for topology three, and it remains constant for topology two. For the remaining four topologies, it exhibits a decrease of between 3% (for topology five) and 51% (for topology four). The average generation rate is decreased from 37 packets/s to 29 packets/s (a 22% decrease) for a 95% completion percentage.

The introduction of the omnidirectional advantaged node also degrades the average performance of the network with the CA protocol, though the penalty is less than with the min-hop protocol. The average generation rate decreased from 52 packets/s to 46 packets/s (a 12% decrease) for a 90% completion percentage. The per-topology generation rate is increased by 38% for topology two, and it exhibits a decrease ranging between 4% (for topology six) and 37% (for topology one) for the other five topologies. The average generation rate is decreased from 48 packets/s to 39 packets/s (a 19% decrease) for a 95% completion percentage. Furthermore, the generation rate decreases for all six topologies under this performance criterion.

The performance advantage of CA routing over min-hop routing is more pronounced in scenario two with an omnidirectional advantaged node than in scenario one without an advantaged node. If the performance criterion is a 90% completion percentage for each topology, the CA protocol results in 53% higher average throughput than the min-hop protocol in scenario two. The increase in the per-topology throughput ranges from 28% for topology one to 67% for topology three. If instead the required completion percentage for each topology is 95%, the average throughput with CA protocol is 34% better than with the min-hop protocol and the increase ranges between 10% and 61% for the six topologies.

The introduction of a three-sector directional antenna at the advantaged node results in a marked improvement in the performance with min-hop routing. This is illustrated by comparing the results in Table 8.15 with the results in Table 8.16 which shows the per-topology generation rate for specified per-topology completion percentages in scenario three. The introduction of sectorization at the advantaged node results in an increase in the average generation rate from 30 packets/s to 49 packets/s (a 63% increase) for a 90% completion percentage. The per-topology generation rate exhibits an increase that ranges between 3% for topology three and 117% for topology two. The average generation rate is increased from 29 packets/s to 47 packets/s (a 62% increase) for a 95% completion percentage, and the per-topology generation rate remains the same or increases for each topology under this criterion.

The average performance with min-hop routing in the presence of a three-sector advantaged node is also superior to its average performance in the absence of an advantaged node. This is illustrated by comparing the results in Tables 8.7 and Table 8.16. The presence of the three-sector advantaged node results in an increase in the average generation rate from 39 packets/s to 49 packets/s (a 26% increase) for a 90% completion percentage. The average generation rate is increased from 37 packets/s to 47 packets/s (a 27% increase) for a 95% completion percentage. The performance for individual topologies varies widely among the topologies for either performance criterion, however. Some topologies exhibit a large improvement in performance and others exhibit a large decrease in performance due to the presence of the three-sector advantaged node.

The introduction of a three-sector directional antenna at the advantaged node results in a more modest improvement in the average performance of the network with CA routing. The introduction of sectorization at the advantaged node results in an increase in the average generation rate from 46 packets/s to 58 packets/s (a 26% increase) for a 90% completion percentage. The change in the per-topology generation rate ranges between -6% for topology six and 55% for topology three. The average generation rate is increased from 39 packets/s to 55 packets/s (a 41% increase) for a 95% completion percentage, and the change in the per-topology generation rate ranges between -4% for topology six and 112% for topology two.

The average performance with CA routing in the presence of a three-sector advantaged node is only slightly better than its average performance in the absence of an advantaged node. The presence of the three-sector advantaged node results in an increase in the average generation rate from 52 packets/s to 58 packets/s (a 12% increase) for a 90% completion percentage. The average generation rate is increased from 48 packets/s to 55 packets/s (a 15% increase) for a 95% completion percentage. Once again, however, some topologies exhibit a large improvement in performance while others exhibit a large decrease in performance.

The performance advantage of CA routing over min-hop routing is much smaller in scenario three than in either scenario one or scenario two. If the performance criterion is a 90% completion percentage for each topology, the CA protocol results in an 18% higher average throughput than the min-hop protocol in scenario three. The increase in the per-topology throughput ranges from 8% for topology four to 36% for topology five. If instead the required completion percentage for each topology is 95%, the average throughput with CA protocol is 17% better than with the min-hop protocol and the increase ranges between 7% and 37% for the six topologies.

The results demonstrate that the presence of an advantaged node has a much less severe impact on the performance of an all-omnidirectional network with min-hop routing than the performance of a heterogeneous network with min-hop routing. This is a consequence of the low generation rate supported by the all-omnidirectional network without advantaged nodes. Even if an advantaged node is introduced and

	Min-Hop		CA routing	
	95%	90%	95%	90%
Topology 1	29.1	30	34.3	38.4
Topology 2	28	29	30.92	46.6
Topology 3	30.5	31.8	47.3	53.1
Topology 4	27.9	28.9	40.3	45.3
Topology 5	27.9	28.9	31.3	39.3
Topology 6	29	30.3	46.6	50.34
Average over 6 topologies	28.7	29.8	38.45	45.5

Table 8.15 Per-topology generation rate (in packets/s) for 95% and 90% completion percentages in scenario two.

	Min-Hop		CA routing	
	95%	90%	95%	90%
Topology 1	29.3	30.5	35.78	37.24
Topology 2	60.8	63.3	65.7	69.8
Topology 3	62	65.3	77	82.1
Topology 4	58	60	62.1	65
Topology 5	32.5	33.6	44.4	45.7
Topology 6	40.35	42	45.3	46.6
Average over 6 topologies	47.2	49.1	55.05	57.8

Table 8.16 Per-topology generation rate (in packets/s) for 95% and 90% completion percentages in scenario three.

most of the network traffic is routed through it, it does not suffer excessive congestion. Thus the network performance is not degraded severely.

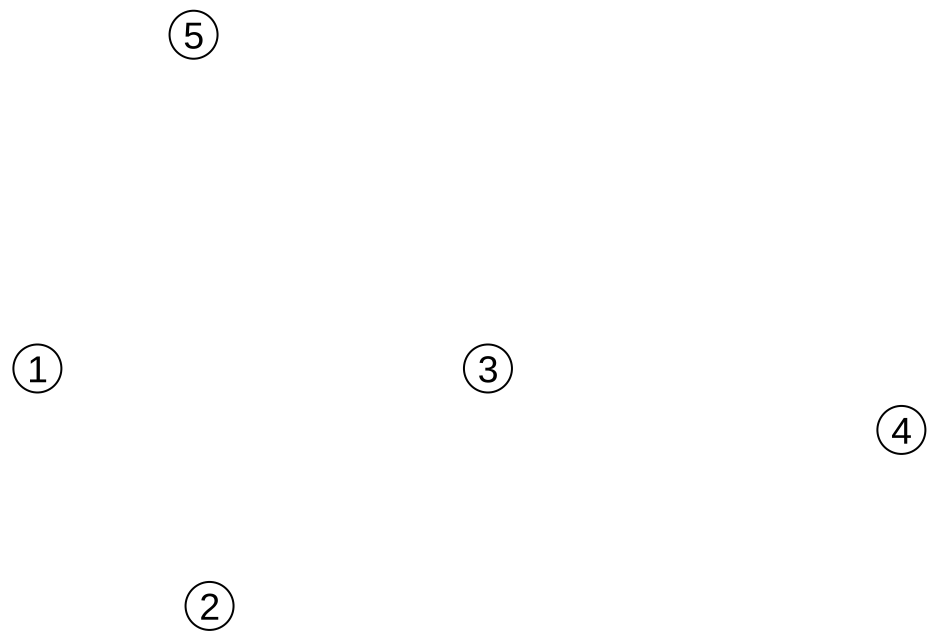

Figure 8.1 Five-node network with all omni-directional antennas.

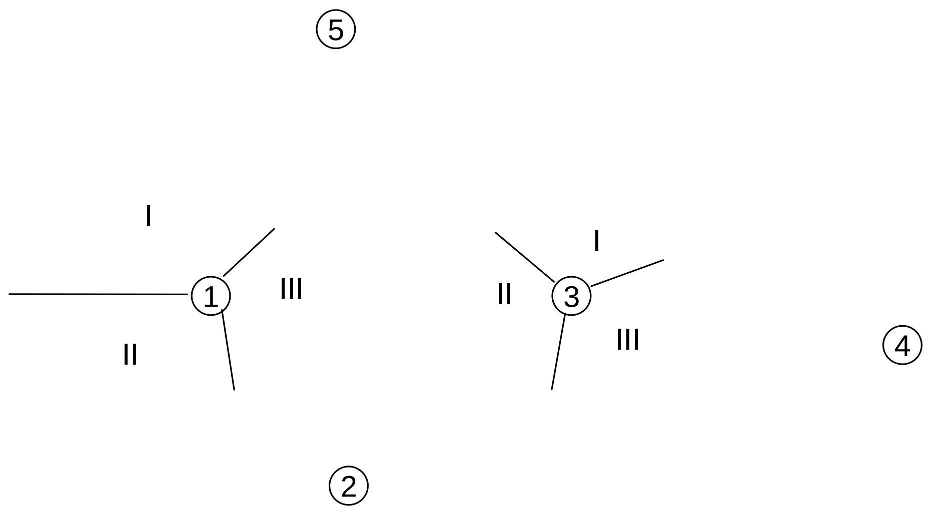

Figure 8.2 Five-node network with directional antennas at nodes 1 and 3.

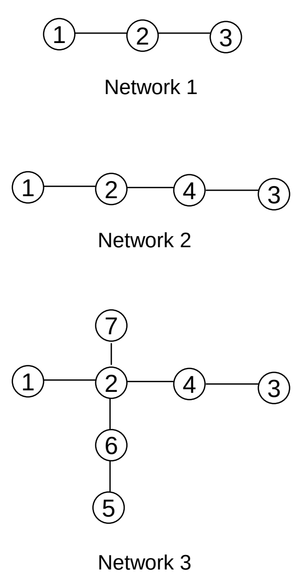

Figure 8.3 Networks used to motivate the metric used to measure total contention at a node.

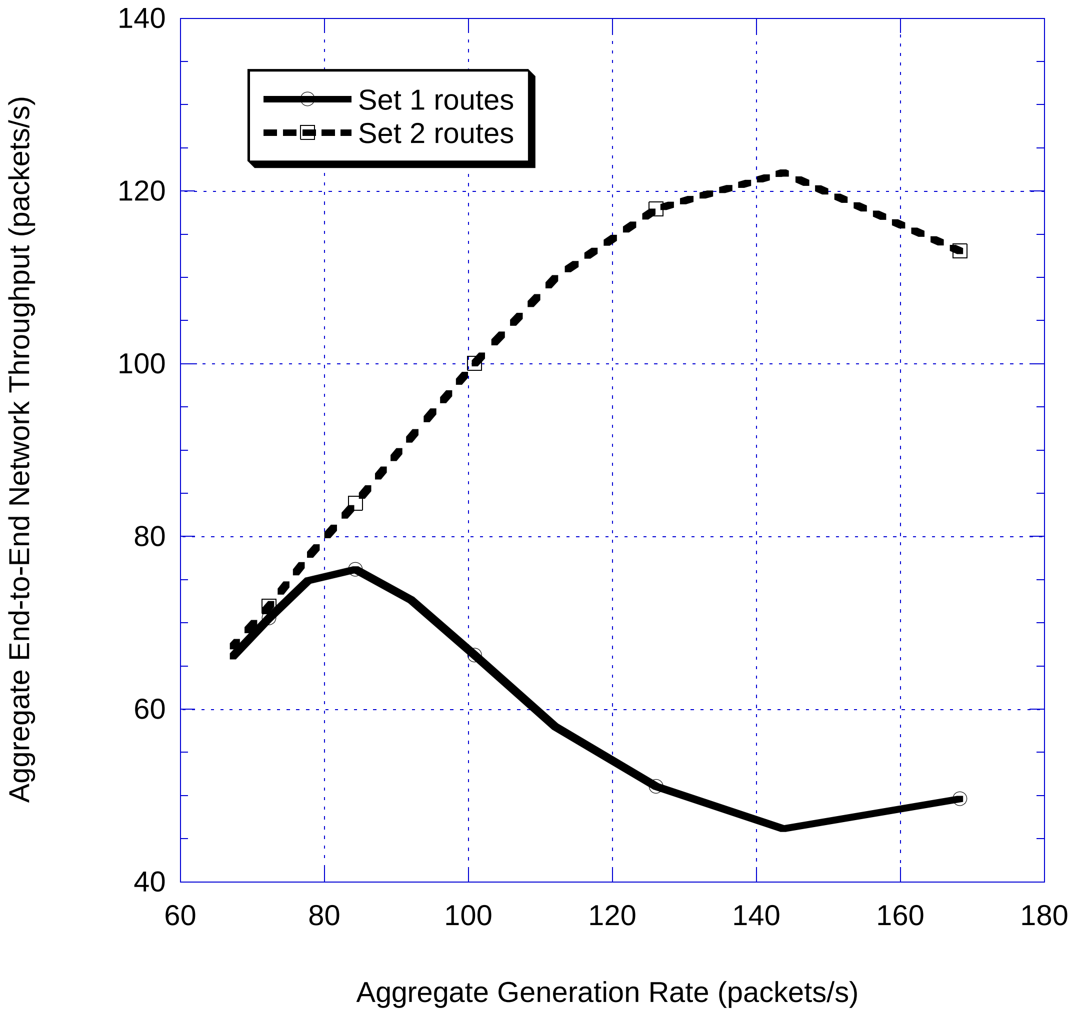

Figure 8.4 Performance of source routing with the routes of sets one and two.

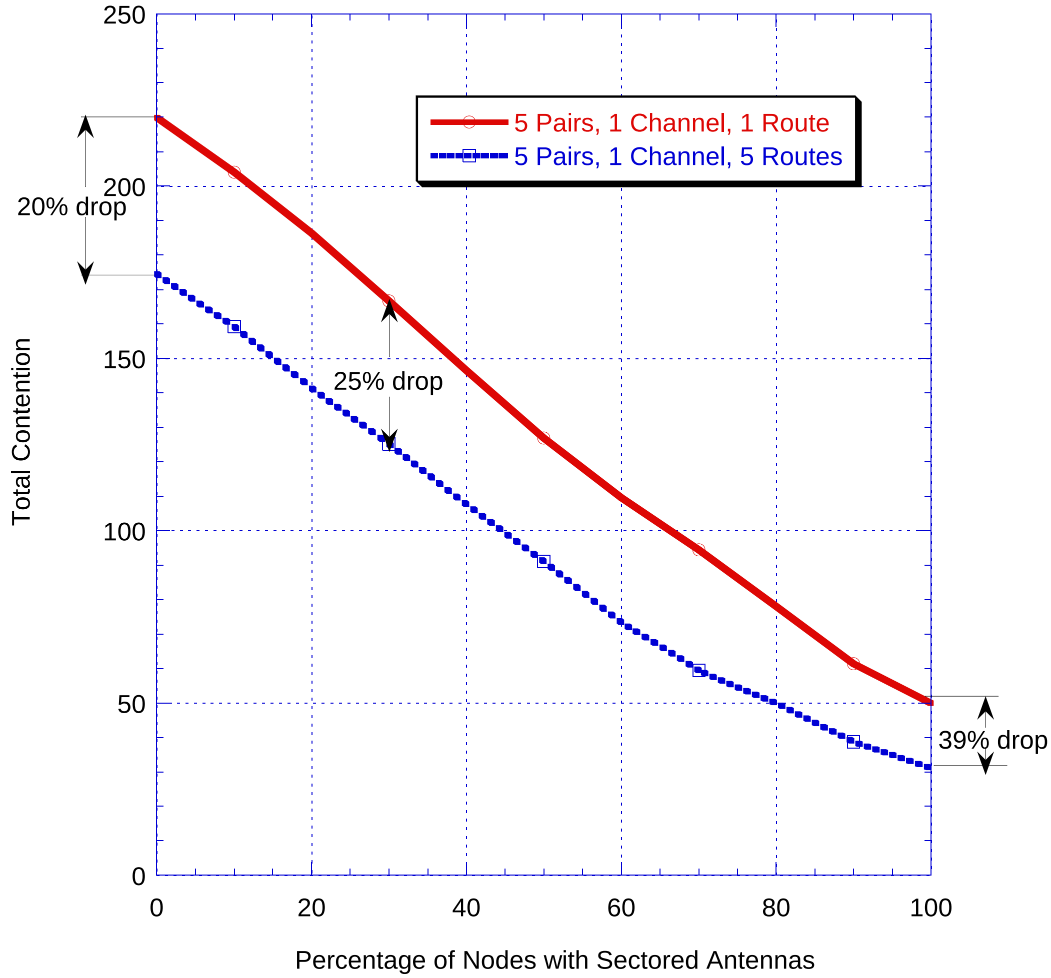

Figure 8.5 Total-contention metric for a 30-node network employing one traffic channel with five source-destination pairs.

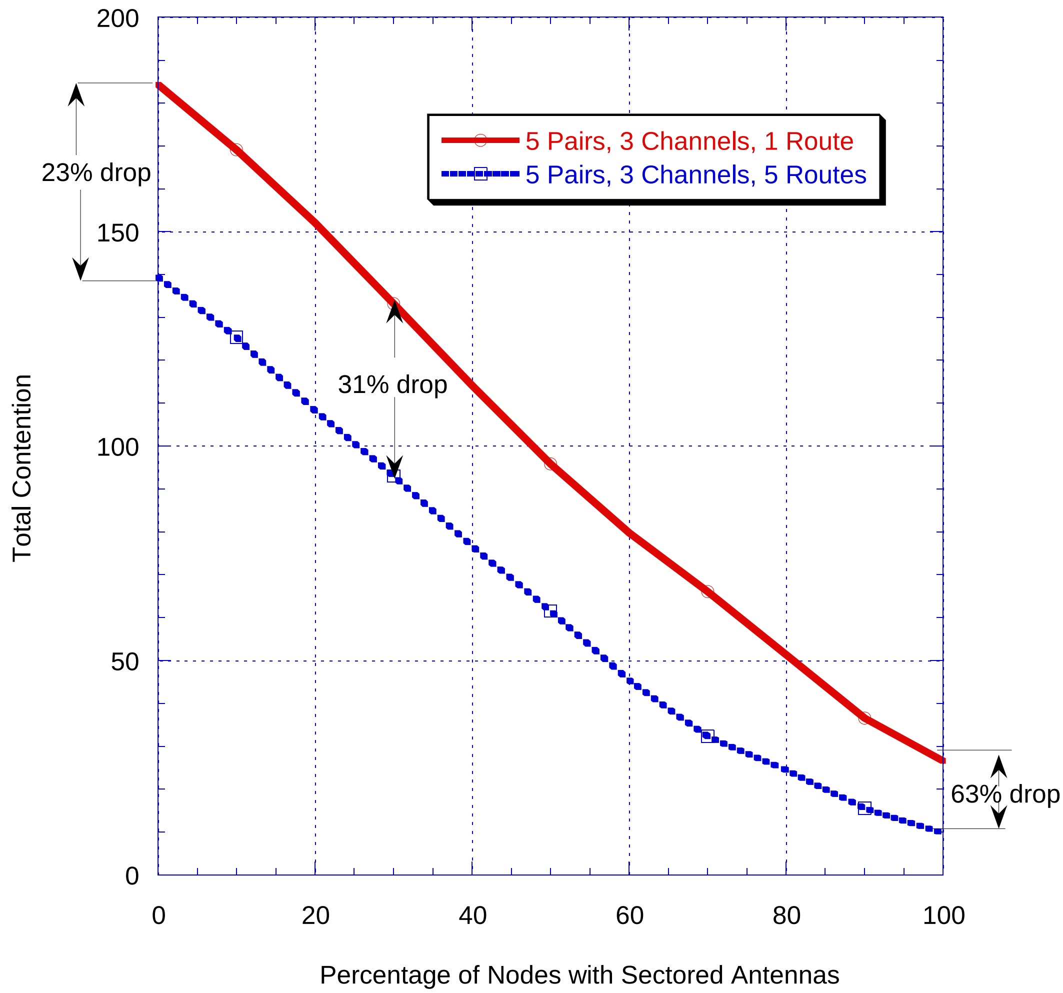

Figure 8.6 Total-contention metric for a 30-node network employing three traffic channels with five source-destination pairs.

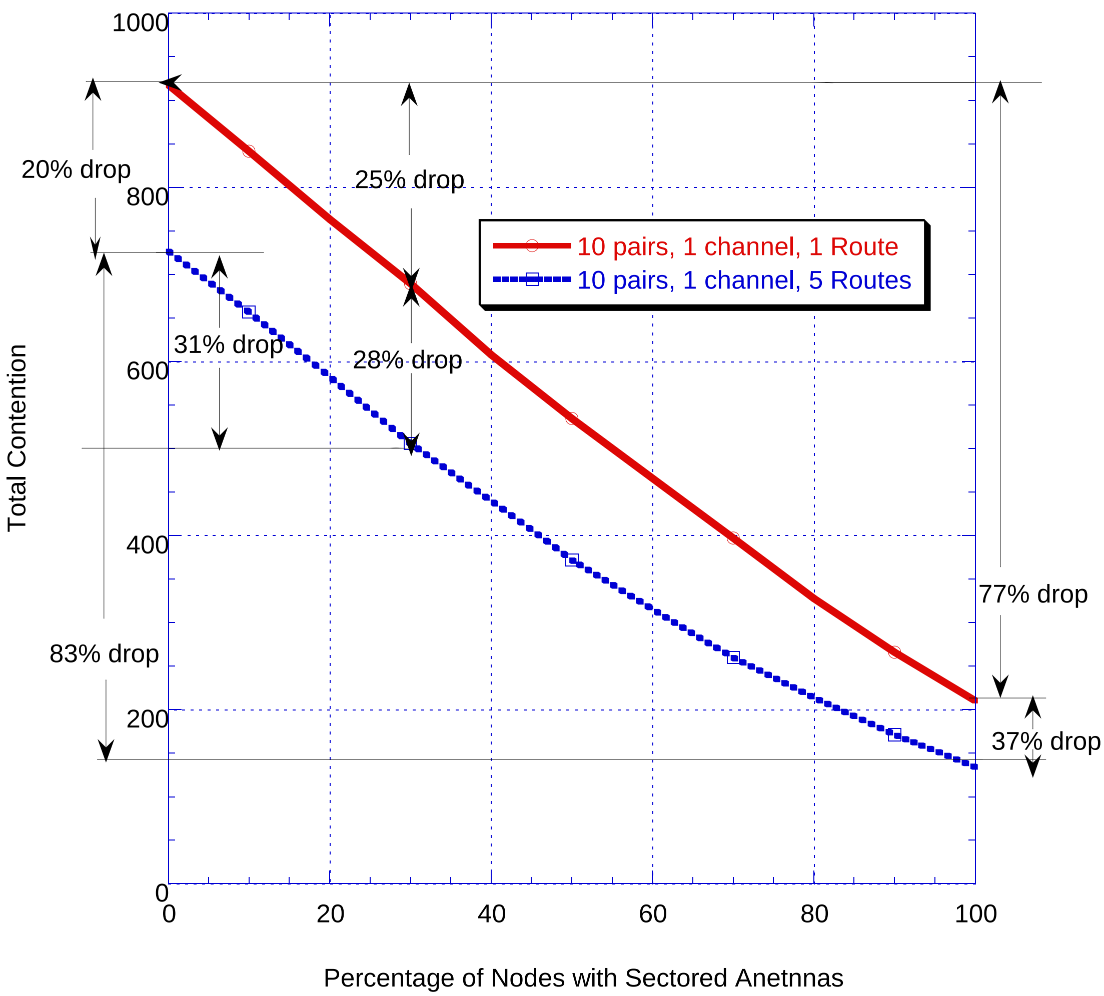

Figure 8.7 Total-contention metric for a 30-node network employing one traffic channel with ten source-destination pairs.

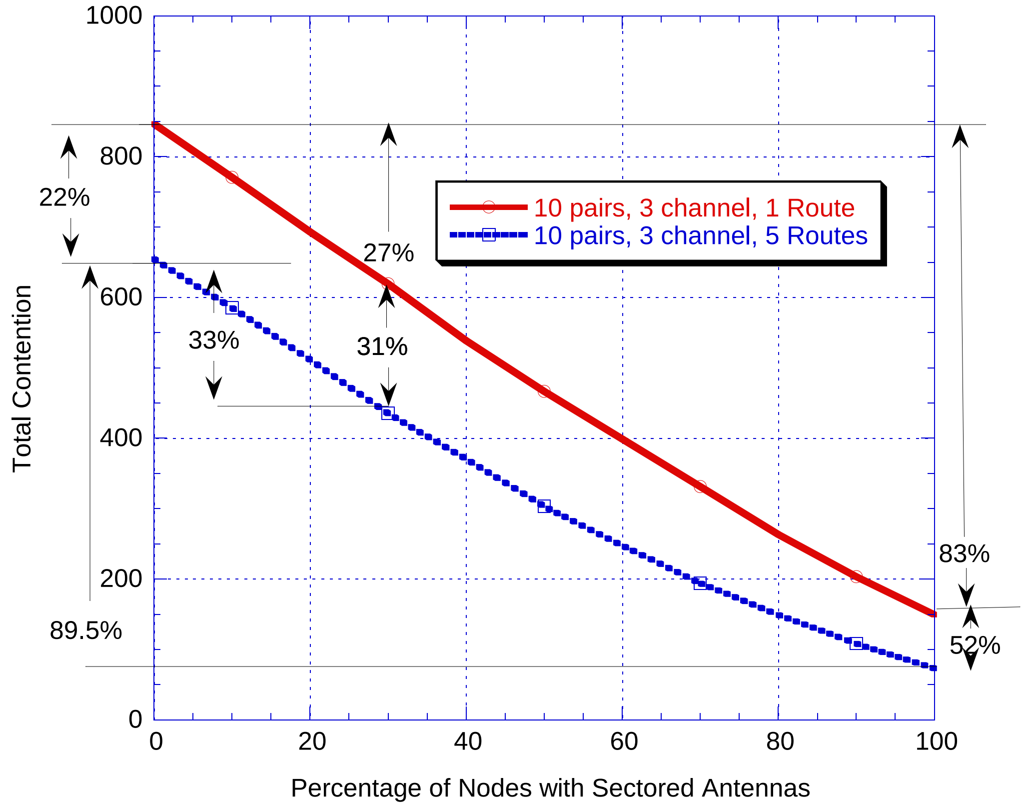

Figure 8.8 Total-contention metric for a 30-node network employing three traffic channels with ten source-destination pairs.

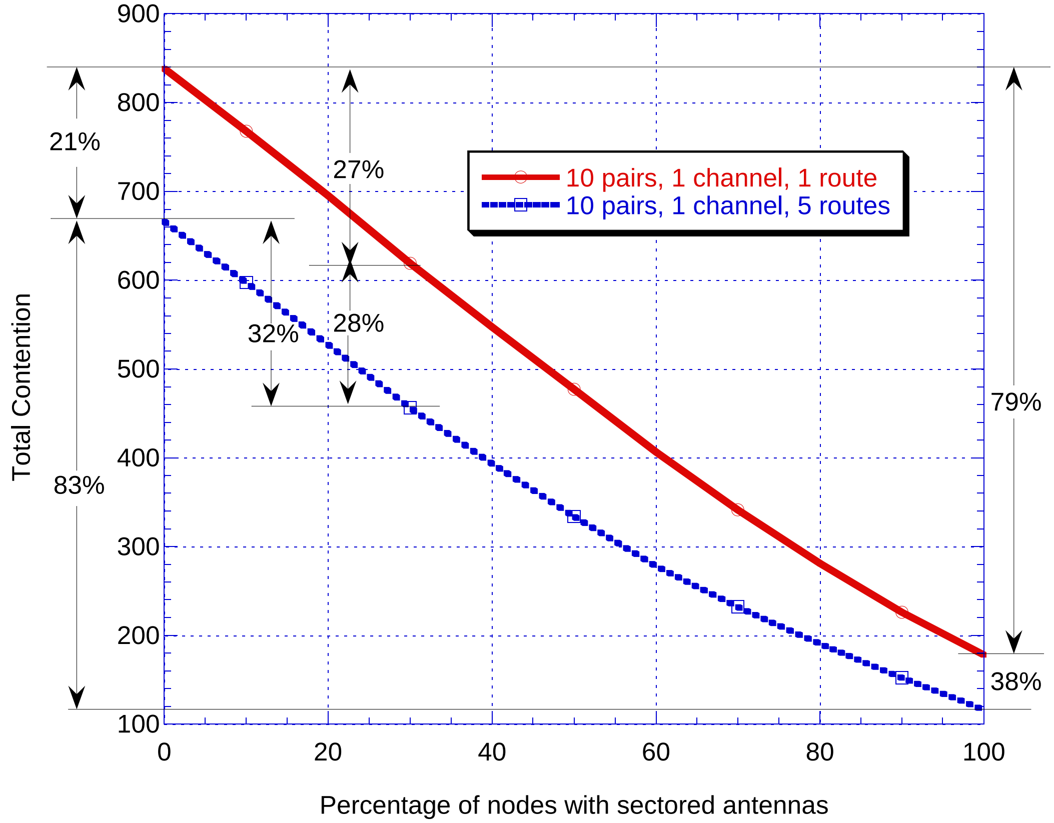

Figure 8.9 Total-contention metric for a 40-node network employing one traffic channel with ten source-destination pairs.

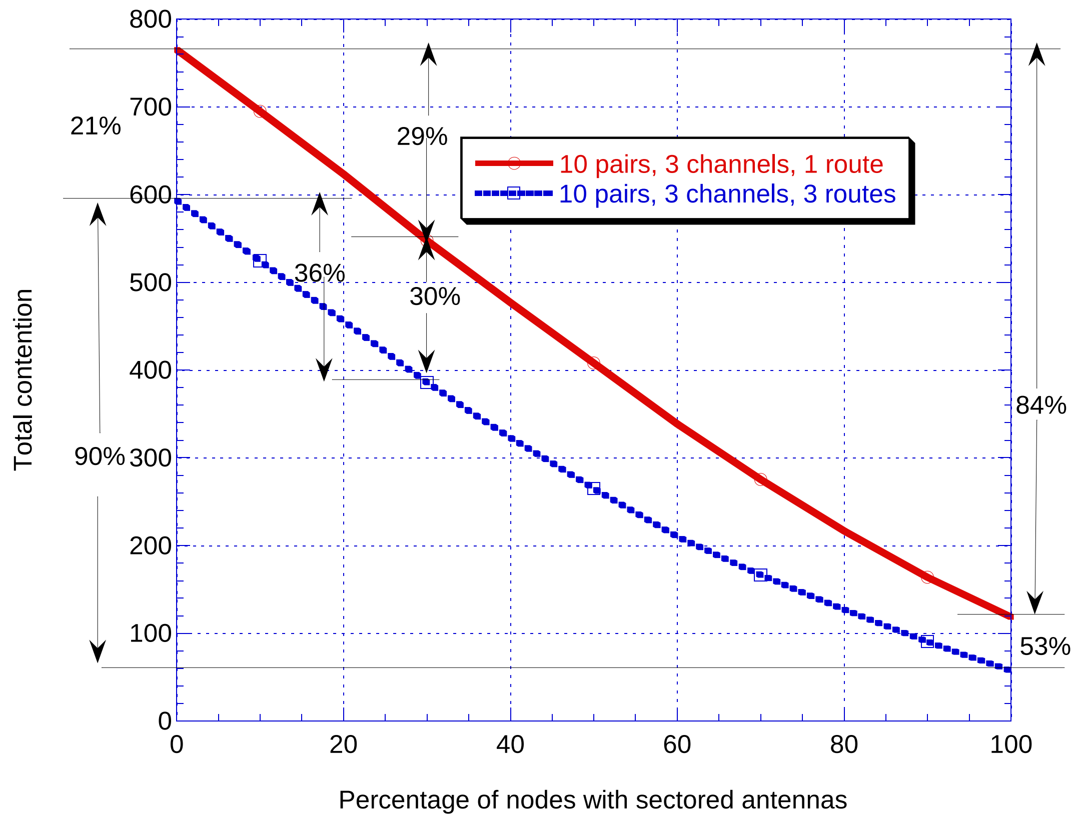

Figure 8.10 Total-contention metric for a 40-node network employing three traffic channels with ten source-destination pairs.

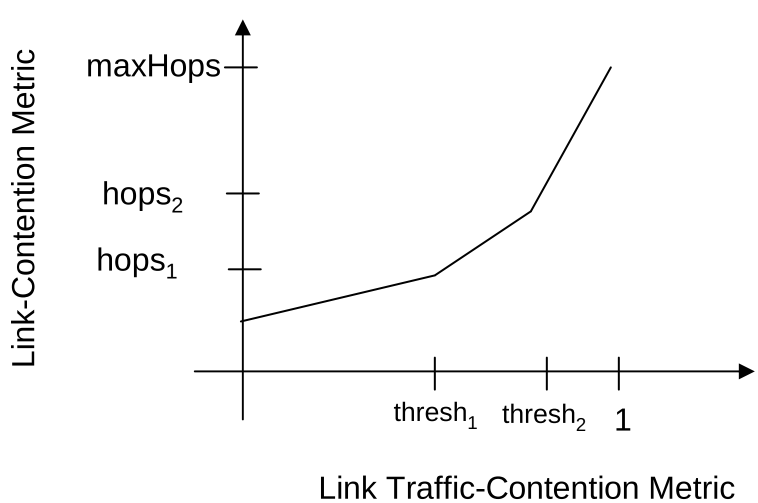

Figure 8.11 Function that maps the LC metric to the link's TC metric.

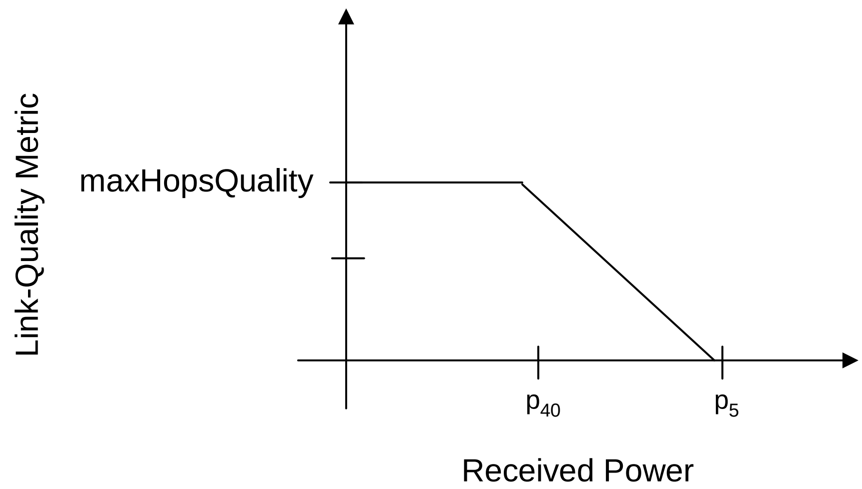

Figure 8.12 Function that maps the received power to link-quality metric.

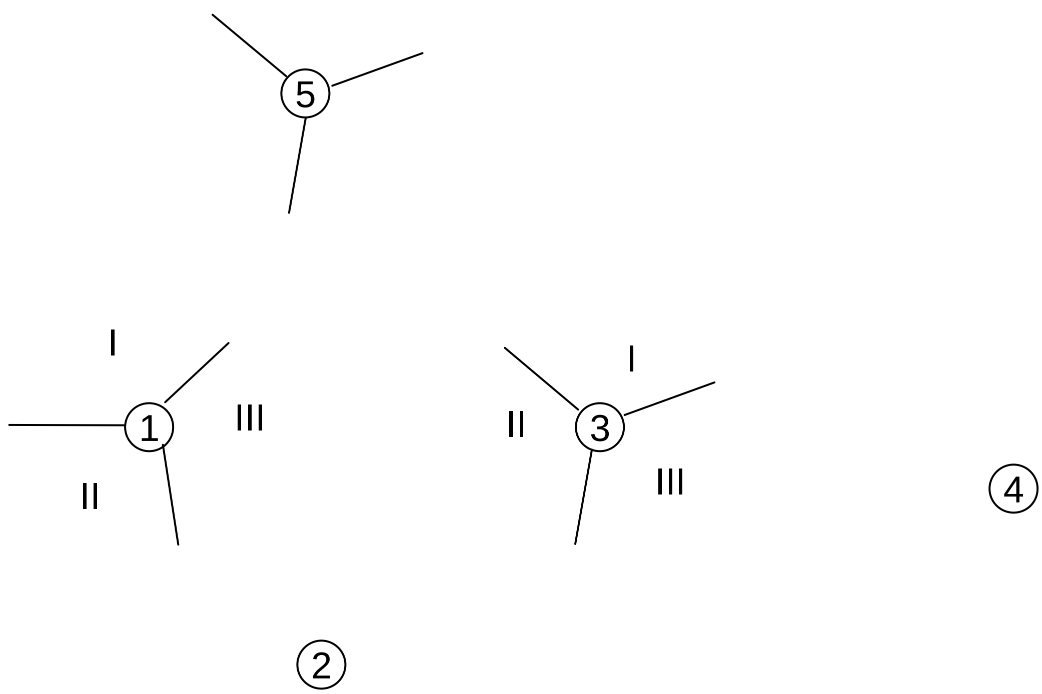

Figure 8.13 Five-node network with directional antennas at nodes 1, 3, and 5.

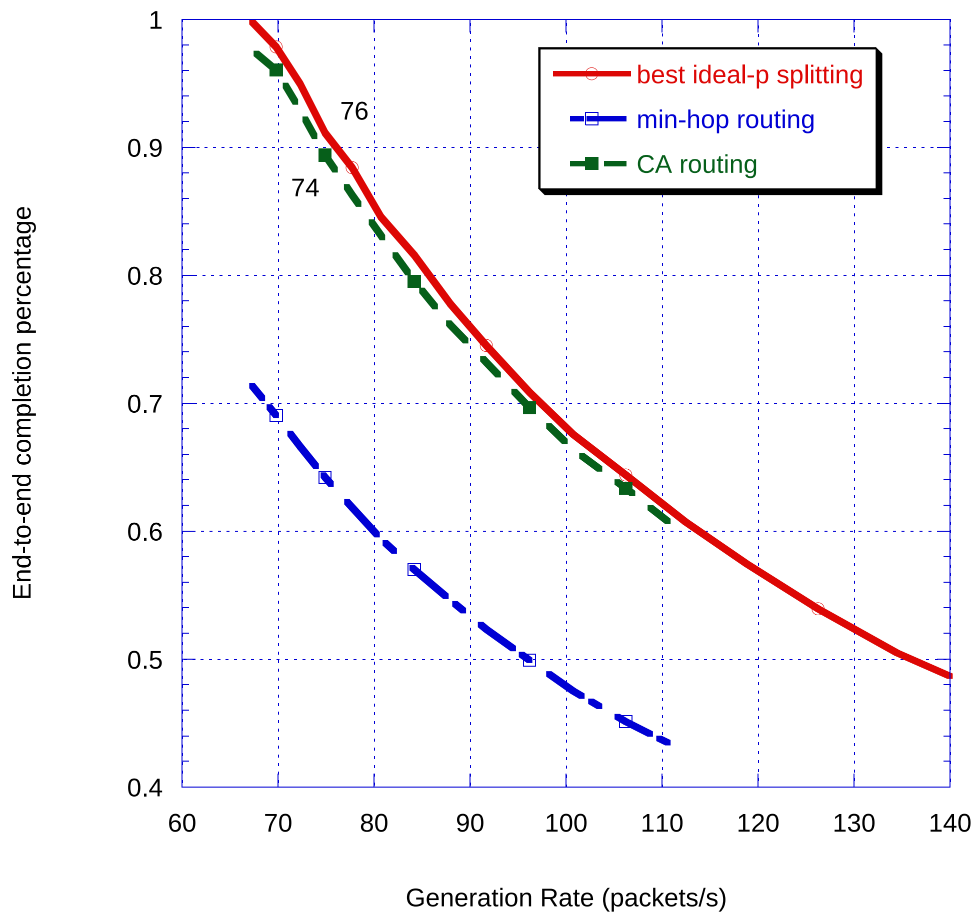

Figure 8.14 End-to-end completion percentage for the five-node network of Fig. 8.2.

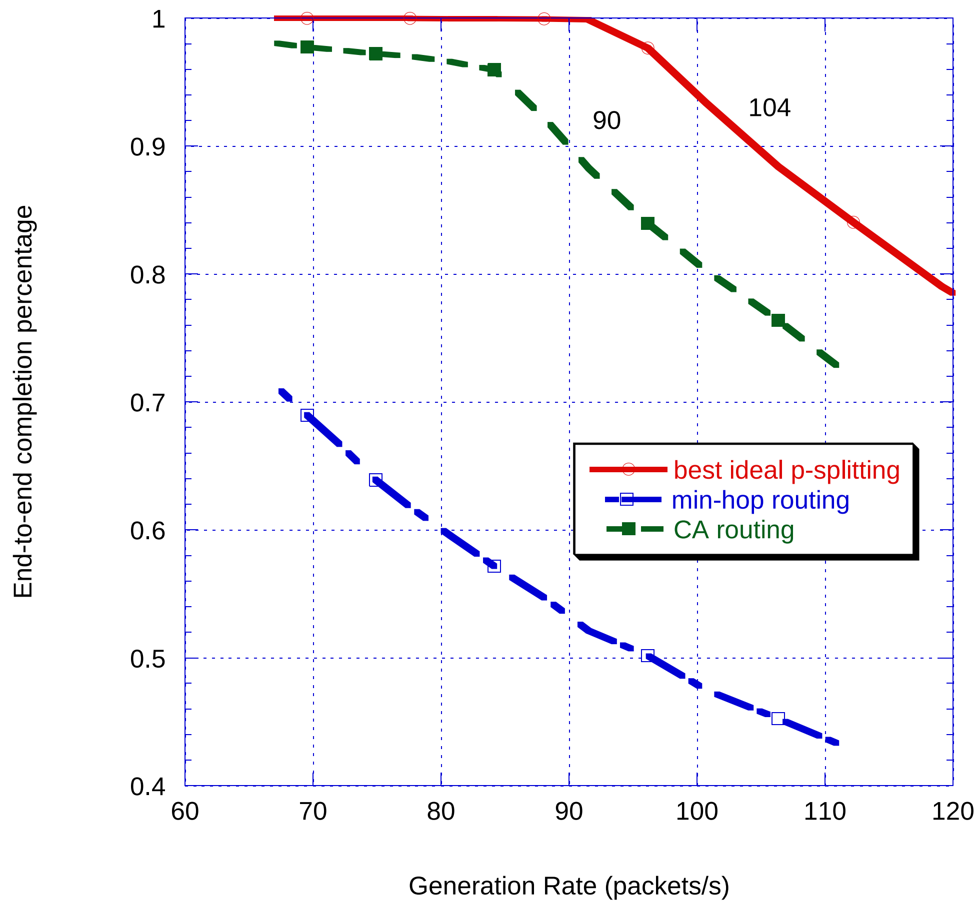

Figure 8.15 End-to-end completion percentage for the five-node network of Fig. 8.13.

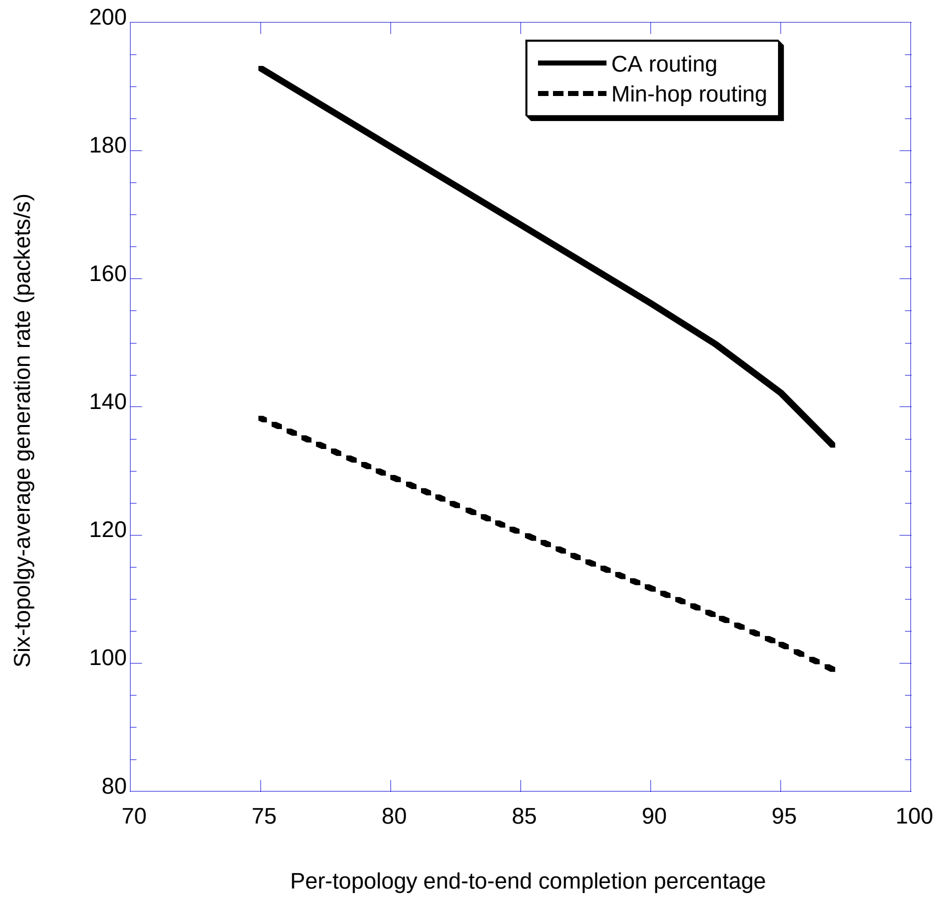

Figure 8.16 Average generation rate as a function of the per-topology completion percentage in scenario four.

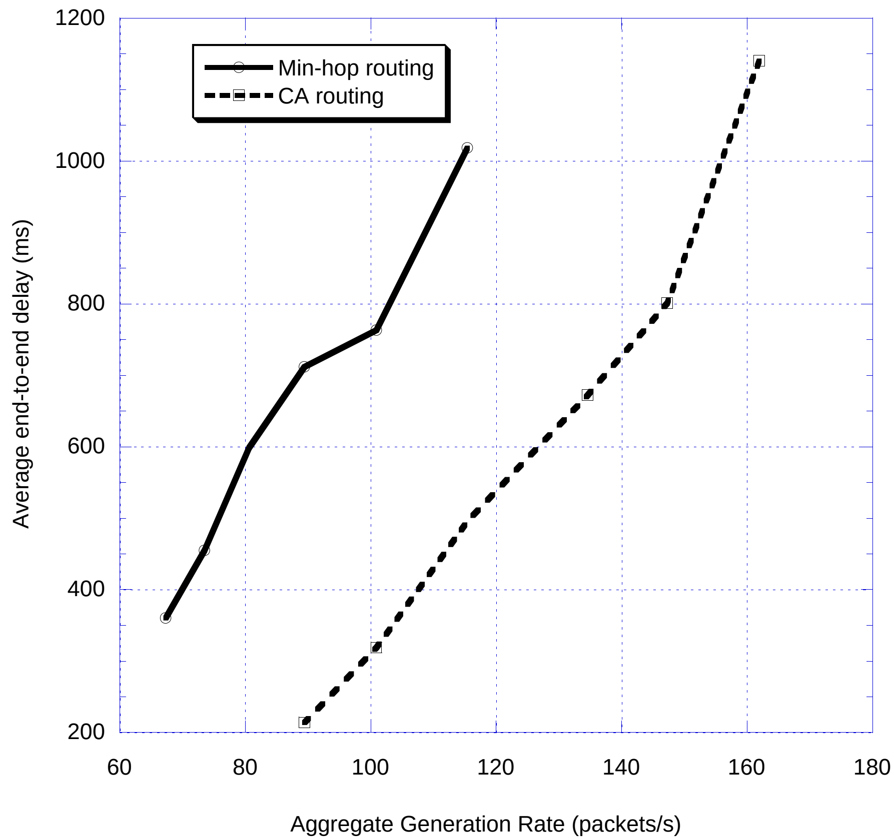

Figure 8.17 Average end-to-end delay over the six topologies as a function of the generation rate in scenario four.

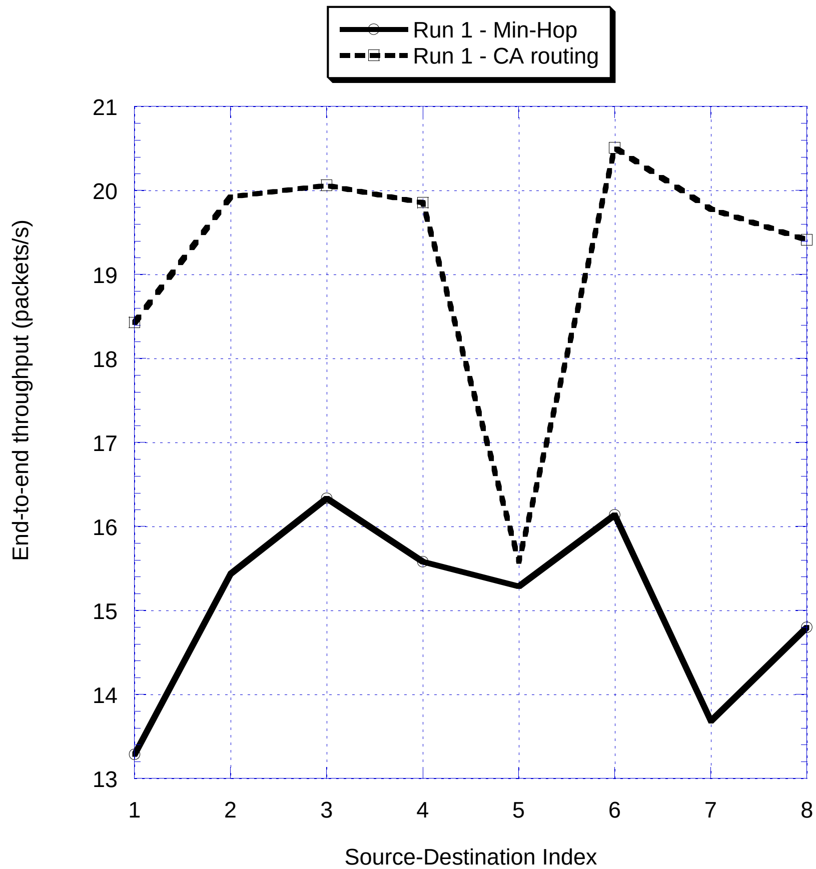

Figure 8.18 End-to-end throughput achieved by each source-destination pair in the first topology of scenario four.

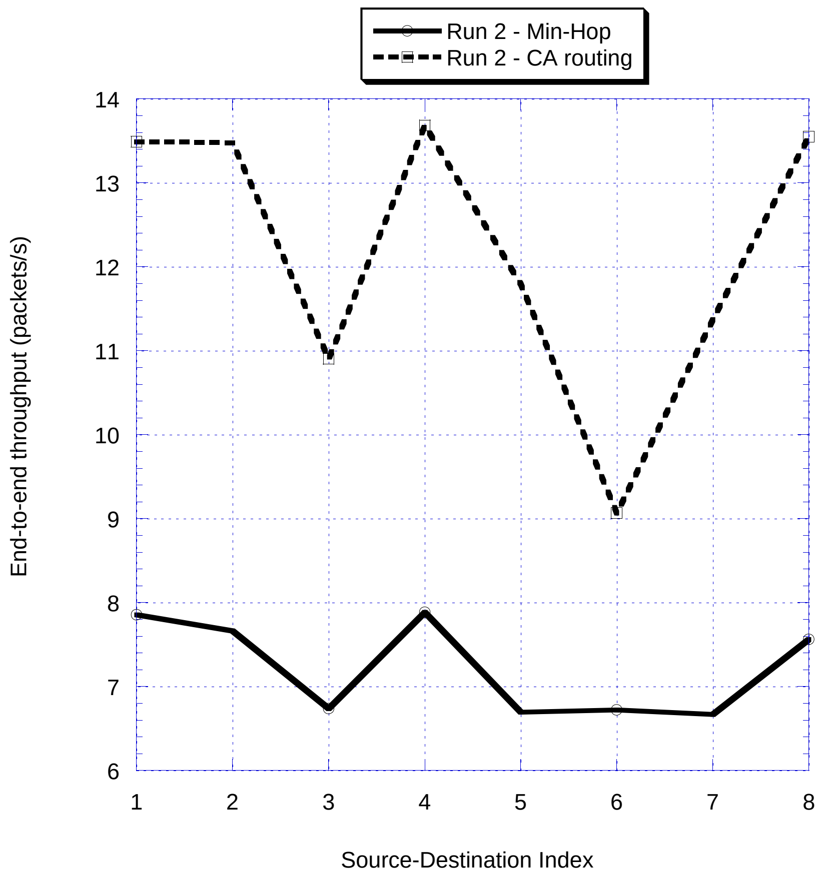

Figure 8.19 End-to-end throughput achieved by each source-destination pair in the second topology of scenario four.

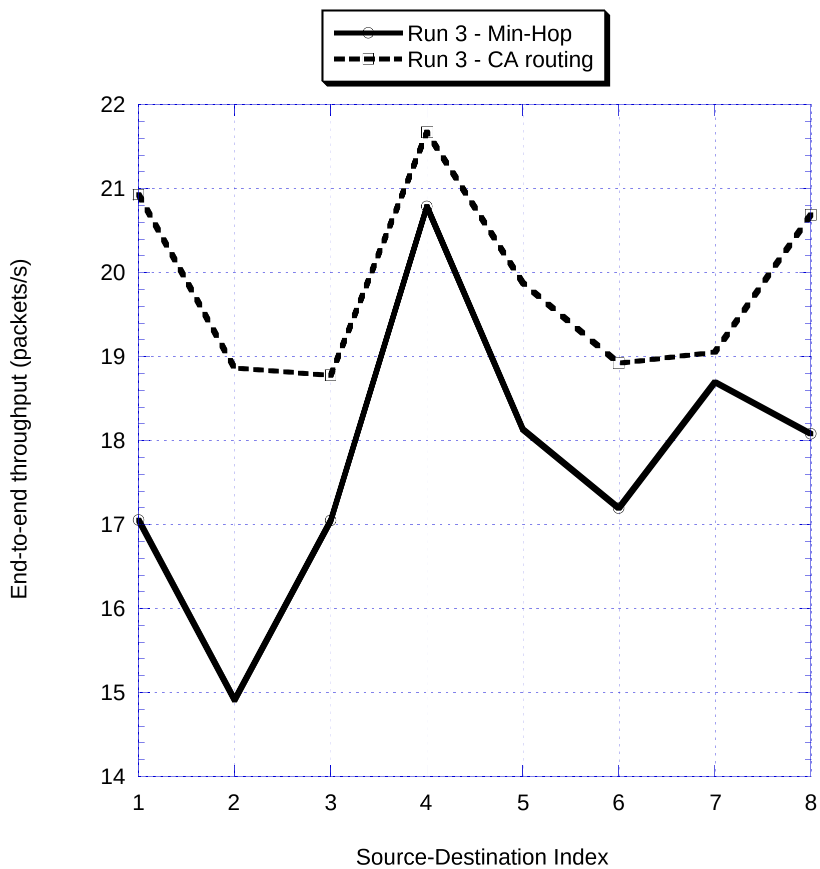

Figure 8.20 End-to-end throughput achieved by each source-destination pair in the third topology of scenario four.

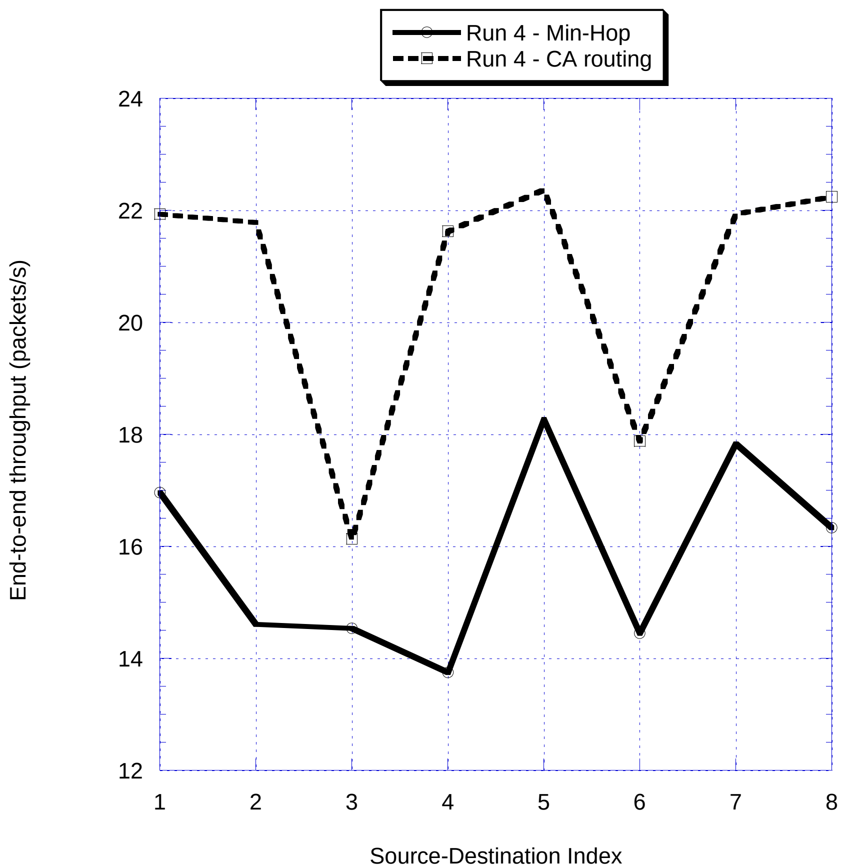

Figure 8.21 End-to-end throughput achieved by each source-destination pair in the fourth topology of scenario four.

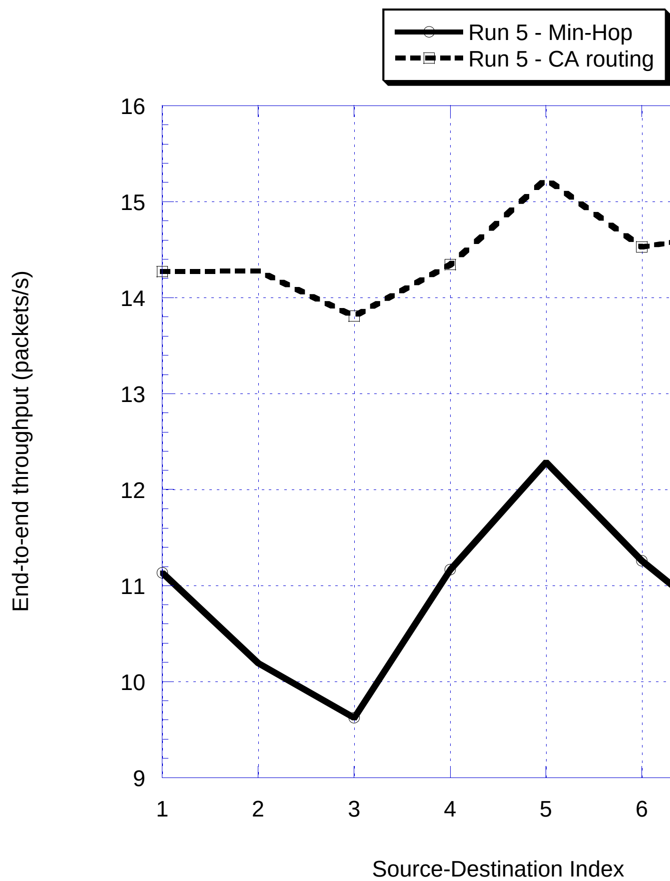

Figure 8.22 End-to-end throughput achieved by each source-destination pair in the fifth topology of scenario four.

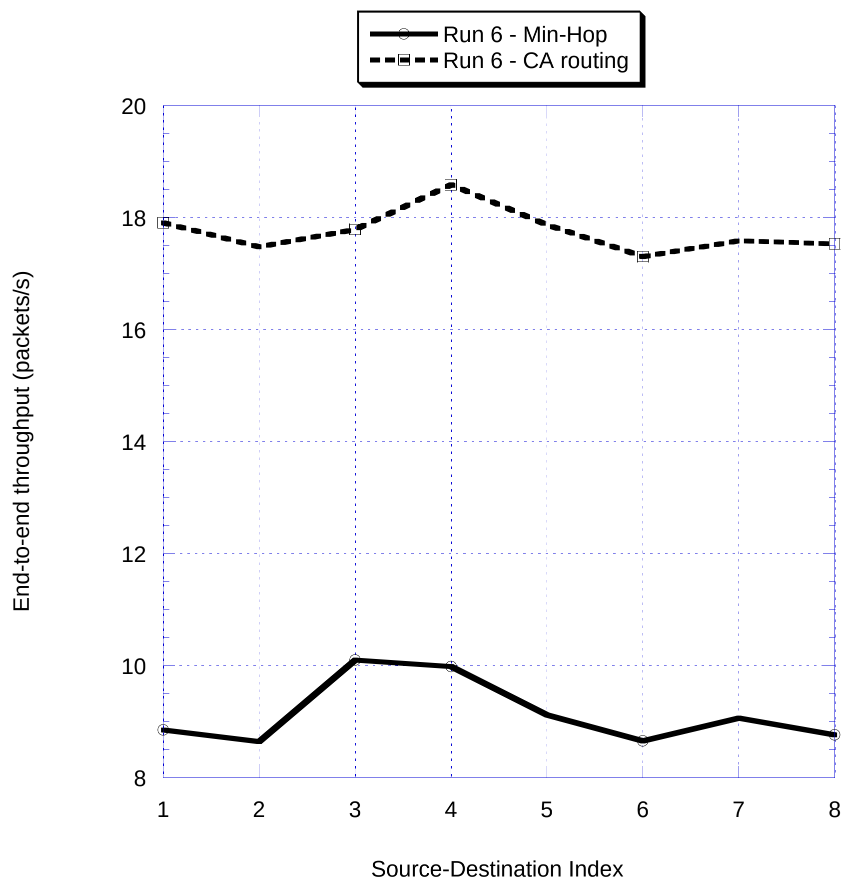

Figure 8.23 End-to-end throughput achieved by each source-destination pair in the sixth topology of scenario four.

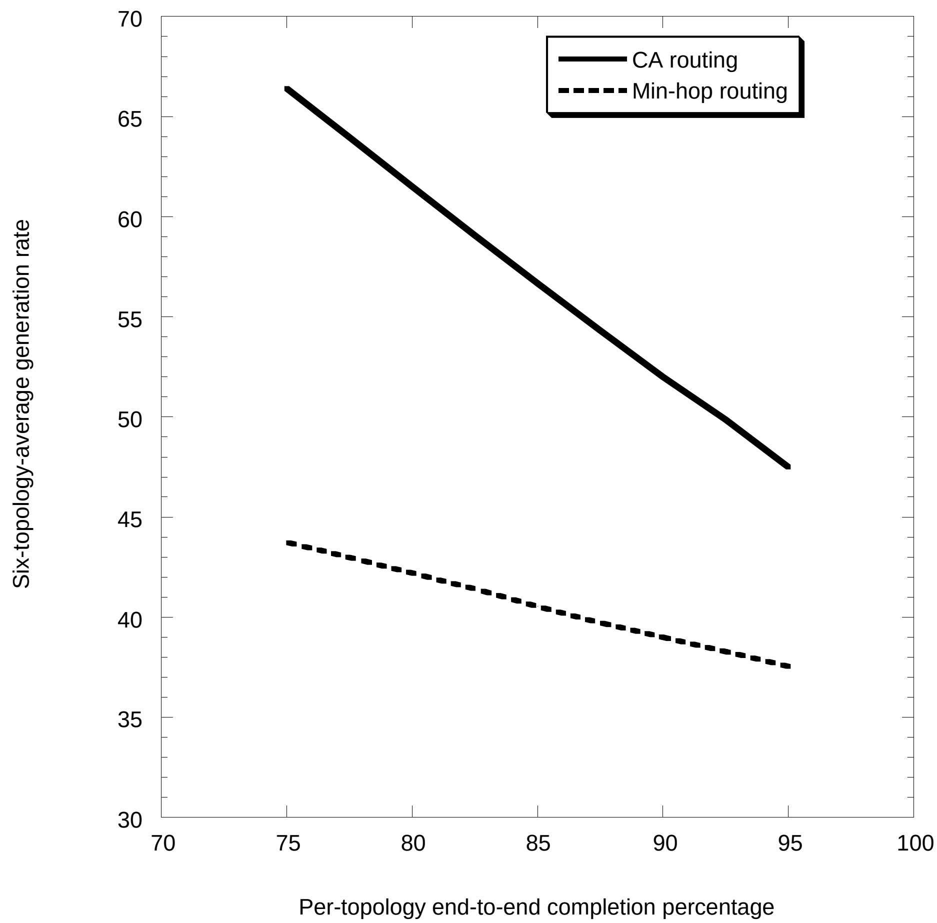

Figure 8.24 Average generation rate as a function of the per-topology completion percentage in scenario one.

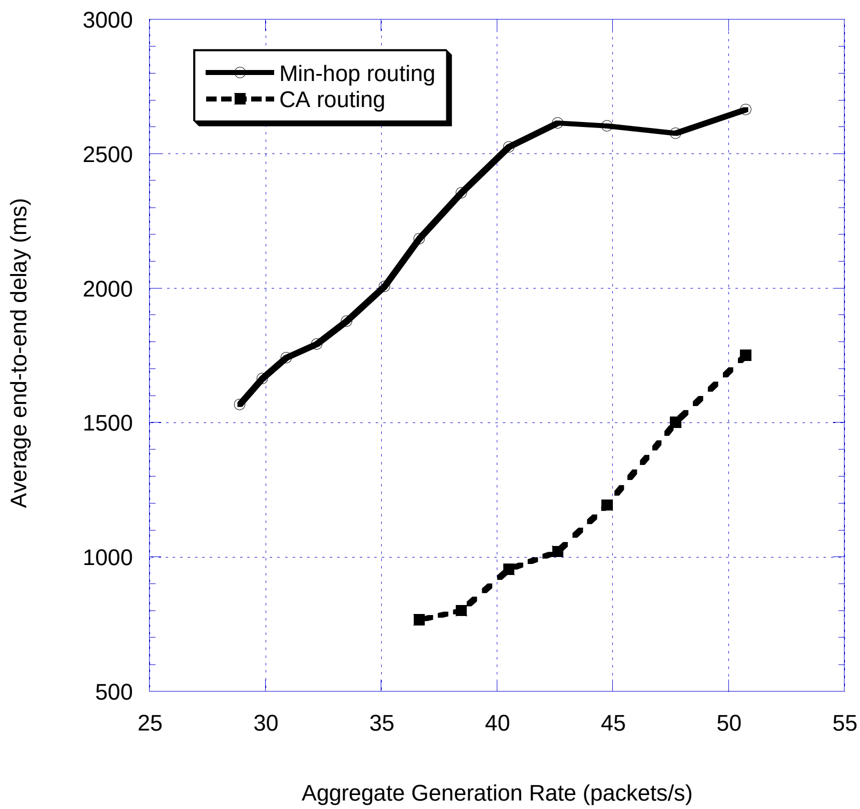

Figure 8.25 Average end-to-end delay over the six topologies as a function of the generation rate in scenario one.

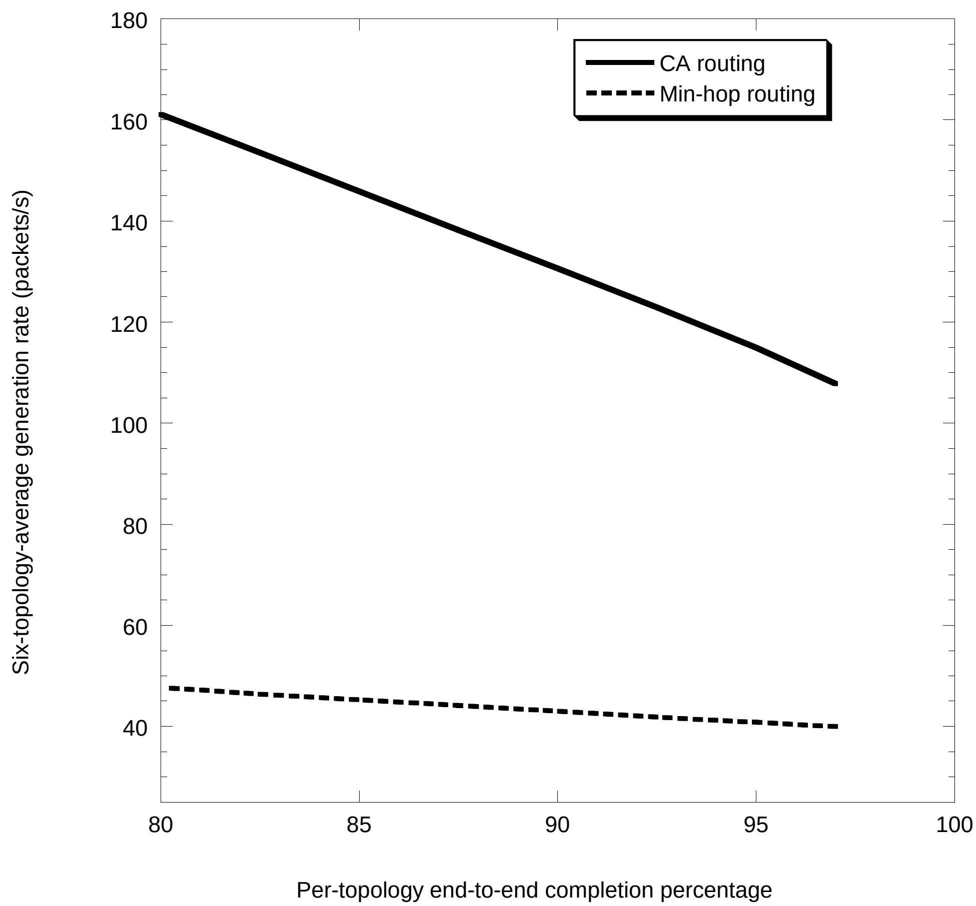

Figure 8.26 Average generation rate as a function of the per-topology completion percentage in scenario five.

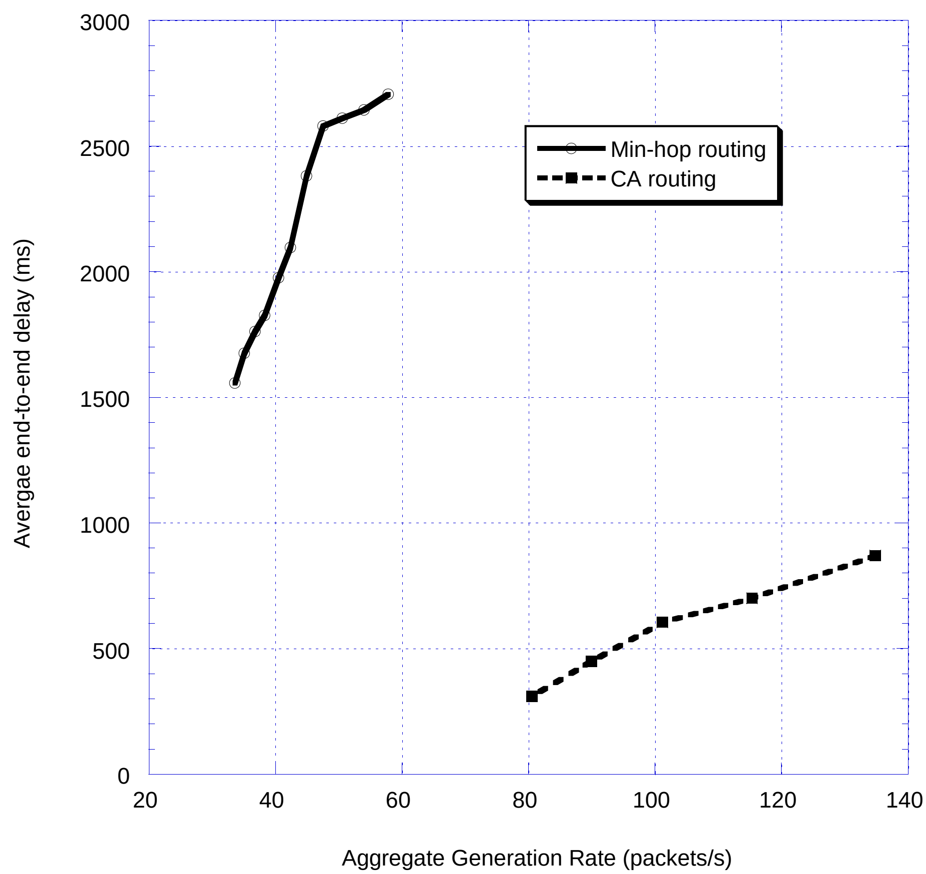

Figure 8.27 Average end-to-end delay over the six topologies as a function of the generation rate in scenario five.

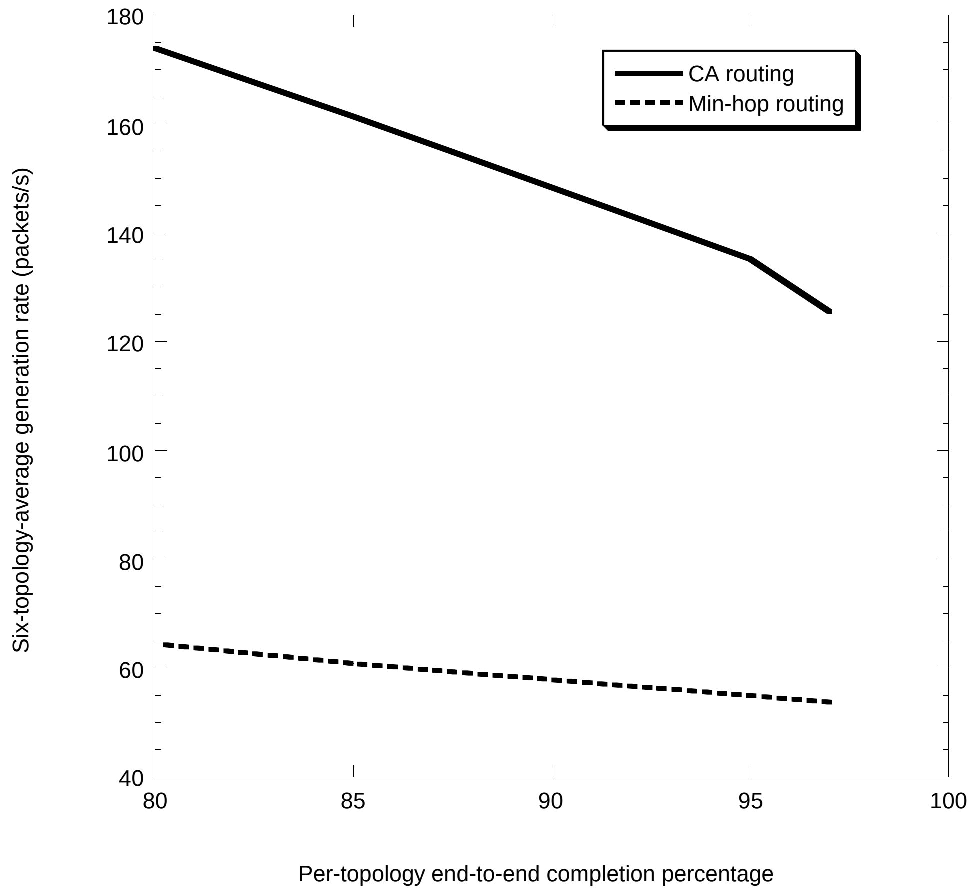

Figure 8.28 Average generation rate as a function of the per-topology completion percentage in scenario six.

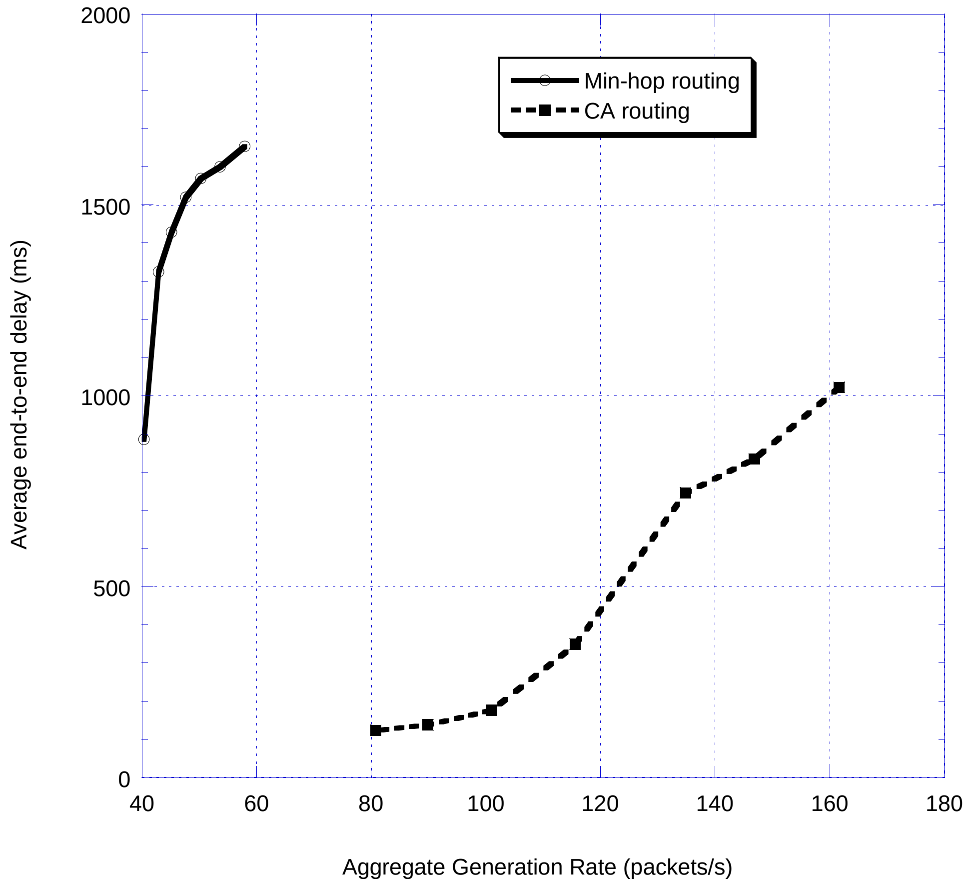

Figure 8.29 Average end-to-end delay over the six topologies as a function of the generation rate in scenario six.

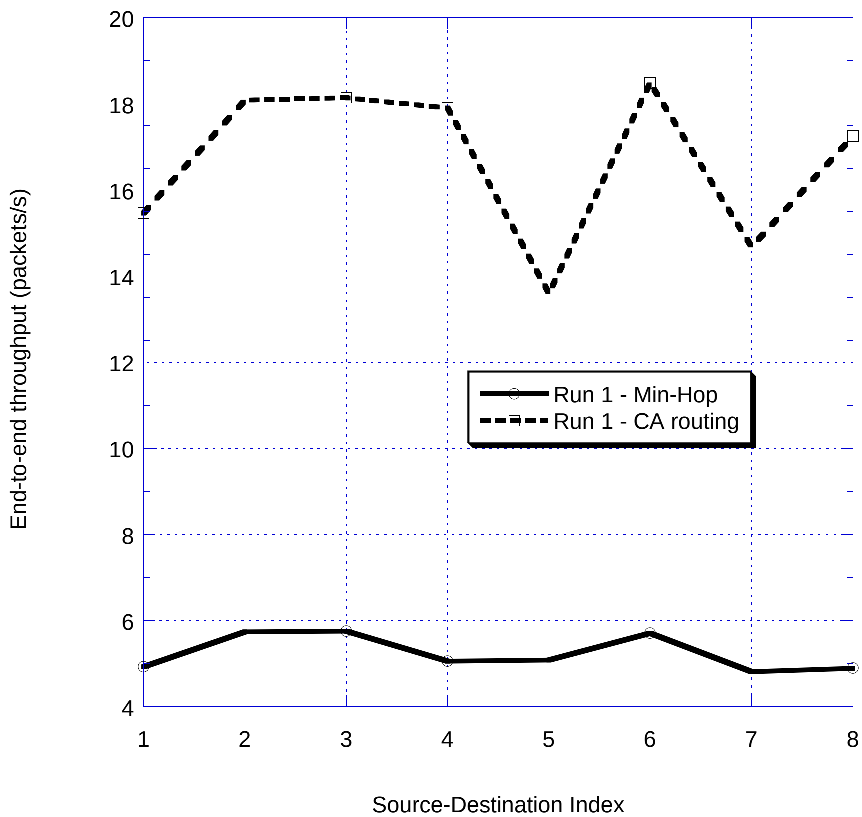

Figure 8.30 End-to-end throughput achieved by each source-destination pair in the first topology of scenario six.

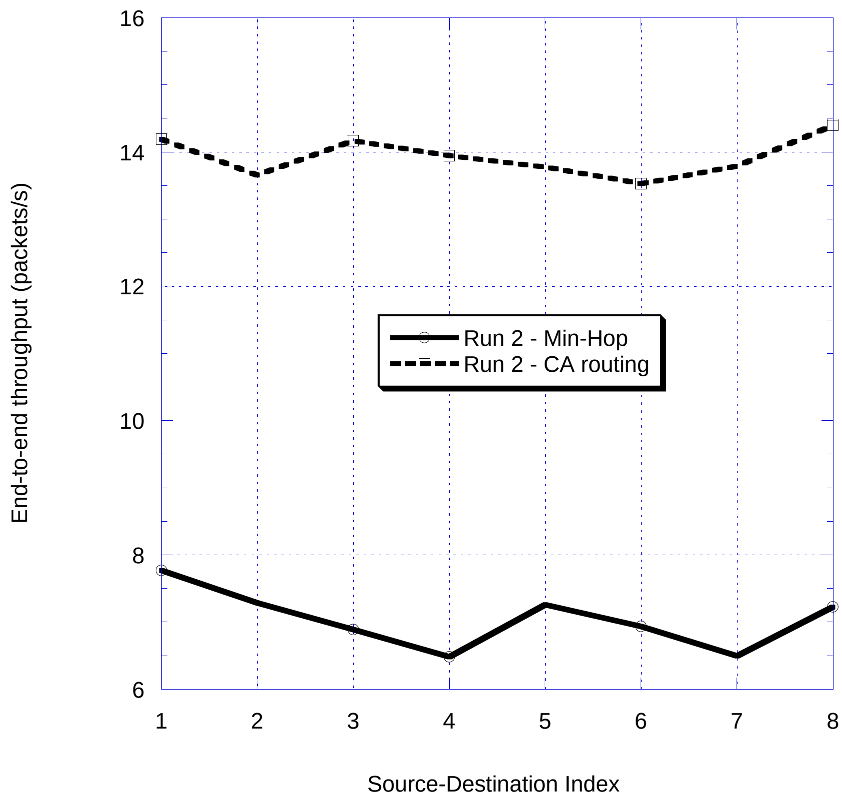

Figure 8.31 End-to-end throughput achieved by each source-destination pair in the second topology of scenario six.

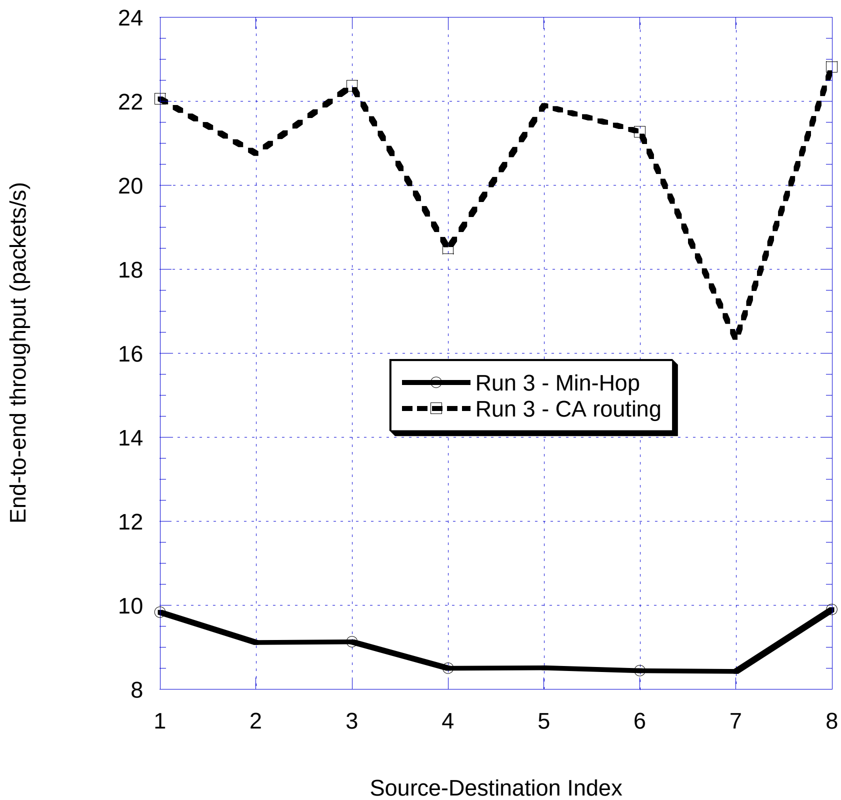

Figure 8.32 End-to-end throughput achieved by each source-destination pair in the third topology of scenario six.

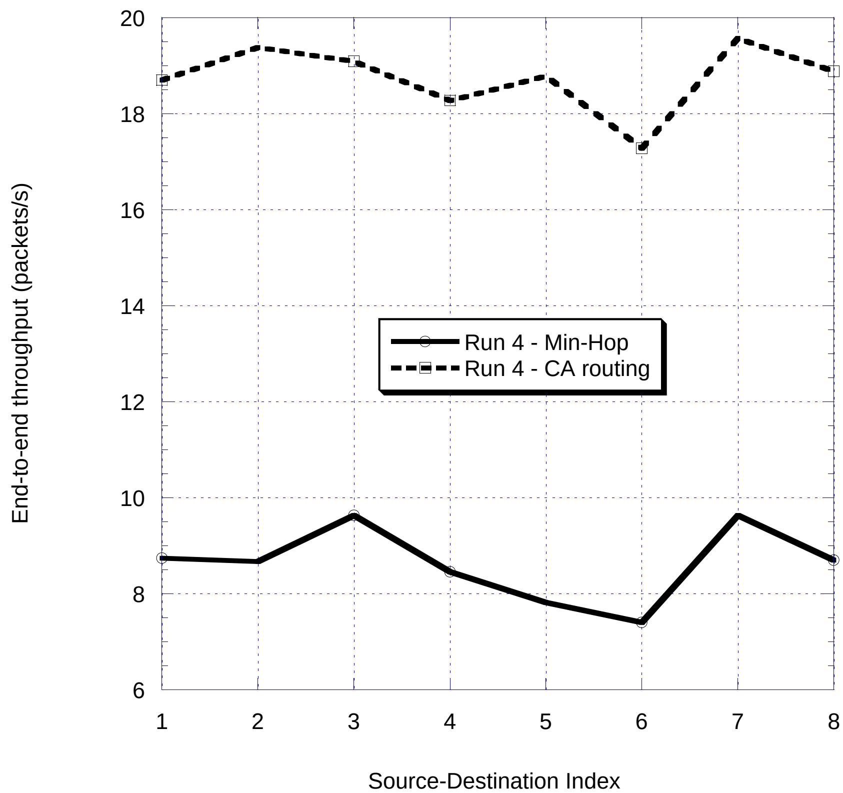

Figure 8.33 End-to-end throughput achieved by each source-destination pair in the fourth topology of scenario six.

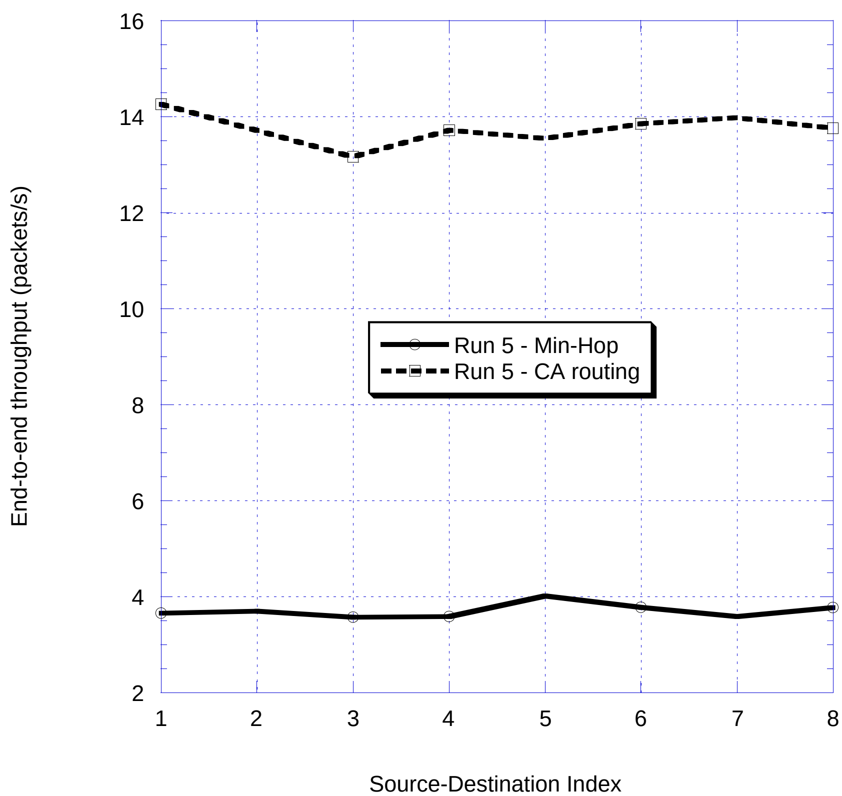

Figure 8.34 End-to-end throughput achieved by each source-destination pair in the fifth topology of scenario six.

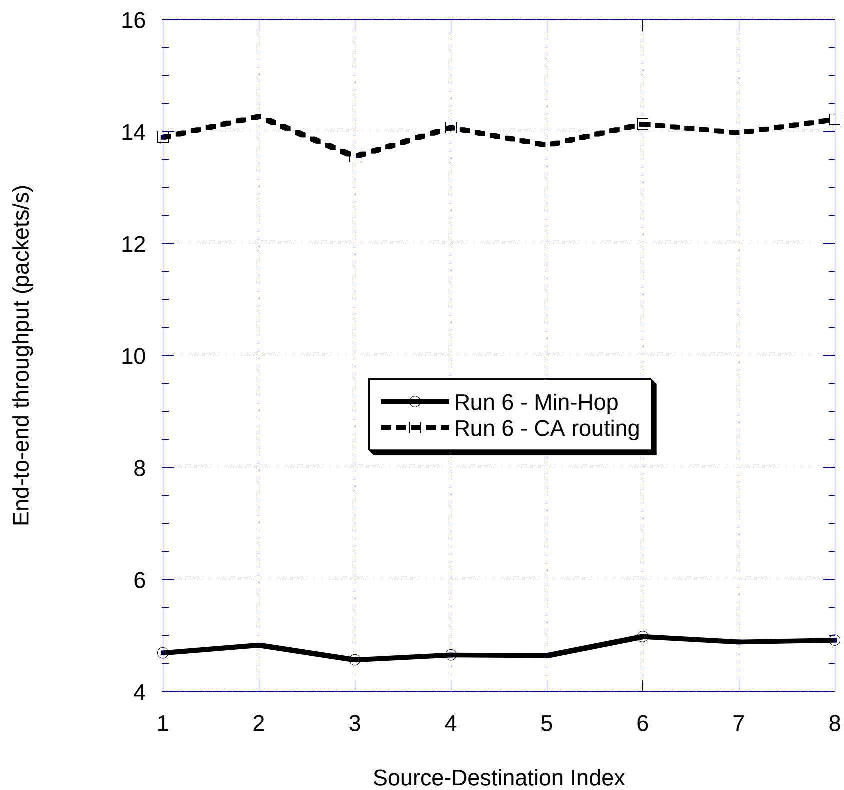

Figure 8.35 End-to-end throughput achieved by each source-destination pair in the sixth topology of scenario six.

CHAPTER 9

CONCLUSIONS

In this dissertation we have present the design of channel-access and routing protocols for use in a heterogeneous wireless ad hoc network that consists of a mix of nodes with directional and omni-directional antennas. The protocols are designed to translate the physical-layer benefits of directional antennas into improved network performance. They account for the unique characteristics nodes employing directional antennas, and they are effective in exploiting the additional degrees of freedom that result from the presence of directional antenna nodes. They are simple and yet lead to substantial performance gains in a wide variety of network scenarios.

In chapter 3 of the dissertation, we present a MAC protocol and a packet transmission scheduling algorithm that is designed to account for the co-site limitation of the directional antenna system as well as to exploit the interference-mitigation capabilities of directional antennas. In chapter 6 we demonstrate that the characteristics of link-level communications using directional antennas results in a greater probability of occurrence of the receiver blocking problem than if omnidirectional antennas are used. It is shown that this has a detrimental effect on network throughput unless the MAC protocol is designed to address the problem explicitly. We then present a new MAC protocol that incorporates a simple mechanism that exploits the availability of the control channel in order to mitigate the receiver blocking problem. It is shown that the use of the modified MAC protocol results in a significant improvement in network performance.

In chapter 7, we illustrated the tension between two counteracting objectives in the design of channel-access protocols in wireless ad hoc networks: spatial reuse of the frequency channels, and protection against multiple-access interference. We demonstrate that a MAC protocol that implements a conservative approach to reuse of the channel results in a poor trade-off between the two objectives under some commonly arising conditions, and we demonstrate that a MAC protocol with an aggressive approach to channel reuse provides a poor-tradeoff under a complementary set of conditions. Based on this observation, we present three variants of a simple, practical MAC protocol that uses a dynamic approach to channel reservation which tailors the aggressiveness of its spatial reuse of traffic channels to the local conditions in the network. The performance of each variant of this selective-reuse protocol is compared for a variety of environments with the performance of two "conventional" RTS-CTS channel-access protocols use respective conservative and aggressive approaches to traffic-channel reservation. It is shown that the selective-reuse protocol yields significantly better performance for a wide range of processing gains of the DS signal, a wide range of network topologies and a wide range of network densities. In particular, two of the variants of the selective-reuse protocol result in uniformly superior performance to either conventional protocol over the full range of conditions that are considered.

In chapter 8, we demonstrated that the presence of nodes with directional antennas decreases the mutual coupling between the different paths in the network. Following this, we presented a new traffic-contention metric for use with the routing protocol and a new forwarding protocol that employs traffic splitting. The traffic-contention metric allows the routing protocol to identify multiple routes between a particular source-destination pair that are not heavily congested. The forwarding protocol utilizes the path metrics determined by the routing protocol to split traffic among two routes with low mutual coupling if such routes can be identified. It is shown that in heterogeneous networks in which some of the nodes have directional antennas, the joint routing and forwarding approach exploits the decoupling of paths in the network and provides substantial improvements in network performance compared to a scheme that simply selects minimum-hop routes. The network performance improvements that result from the new protocols are particularly effective in networks in which some of the nodes have directional antennas. Significant network performance gains are also shown for networks in which all nodes employ omnidirectional antennas, though the performance gains are not as large. Finally, the CA routing protocol is particularly beneficial in a heterogeneous network with one or two advantaged nodes.

APPENDICES

Appendix A

Detailed Description of Baseline MAC Protocol

A.1 Data Structures

Each node in the network maintains a *Channel Usage (CU) table*, which is a list of entries of the form <*sector-number, channel number, receive, transmit*>. The table is used to store information about the availability of the control channel and the traffic channels in each sector of the node. For example, an entry of the form <3, 2, YES, NO> indicates that in sector three, traffic channel two is available for reception but not for transmission. Nodes with directional antennas (and thus with multiple sectors) additionally maintain a *Neighbor Sector (NS) table*, which contains an entry of the form <neighbor, sector-to-use> for each neighboring node. The entry indicates the sector in which communication with the neighbor is possible, and hence, which transceiver should be used in the communication. Finally, every node maintains an *Outstanding Packets (OP) table* containing outgoing packets that have been dispatched from the network layer to the MAC layer.

A.2 Packet Scheduler

The network layer at each node maintains two queues of packets scheduled for transmission: the *control queue* and the *data queue*. The control queue contains *PROP* packets, which are network-layer control packets broadcast to all neighboring nodes periodically for use in building routing tables [22]. The data queue contains unicast, routed data packets.

The scheduler is initially in the idle state awaiting an interrupt. One type of interrupt occurs due to the arrival of a data packet from the transport layer. In this case the packet is added to the tail of the data queue and control is passed to a common dispatcher routine which is described below. A second type of interrupt occurs due to the generation of a prop packet by the control plane of the network layer. In this case, all packets already in the control queue are first deleted. Since the prop packet is a broadcast packet, one copy of the packet is scheduled for each of the n sectors of the node. Hence, n copies of the packet are inserted in the queue, each with a label indicating the sector for which the copy is intended. The packets are inserted in the queue in a random order to ensure fairness in the delay of the prop packets transmitted into the various sectors. Following this, control is passed to the dispatcher routine.

A third type of interrupt occurs when either the MAC layer informs the scheduler of a change in the information stored in the CU table or a back-off timer expires. Control is passed to the dispatcher routine to determine if any queued packets can thus be transmitted. A fourth type of interrupt occurs when the MAC layer informs the scheduler of an event related to a specific packet. The interrupt may concern an (outgoing) packet in the OP table. In this case, the interrupt indicates either that the packet was transmitted successfully (including receipt of an acknowledgment if the packet is from the data queue) or that the packet was not transmitted successfully. If the packet was not transmitted successfully, it is time-stamped according to an exponential back-off algorithm [42], an interrupt is scheduled with the corresponding time delay, and the packet is queued again. If the interrupt concerns a packet received from another node, the packet is sent to the transport layer if it is intended for the local node. But if it is a packet to be forwarded, it is added to the data queue. Control is again passed to the dispatcher routine.

The dispatcher routine first determines if there are any (PROP) packets to be transmitted in the control queue. If there are packets in the control queue, each packet in the control queue starting with the head-of-line packet is checked to determine if it can be transmitted. A PROP packet can be transmitted if the control channel and at least one data channel are both available for transmission in the sector into which the packet is to be transmitted. If a PROP packet can be transmitted, the network layer chooses one data channel at random from among the available ones in the corresponding sector, and it updates the CU table to indicate that the control channel and the chosen data channel cannot be used for reception in any sector. Then the packet is sent to the MAC layer along with a specification of the data channel chosen for the transmission. Following this the network layer examines the next packet in the control queue and applies the same tests. When all the packets in the control queue have been tested, the network-layer process returns to the idle state. Note that multiple packets from the control queue may be dispatched (into different sectors) before the network-layer process returns to the idle state.

If there are no packets in the control queue when the dispatcher is called, or if none of the packets in the control queue can be transmitted, the dispatcher determines if there are packets in the data queue. If there are none, the process returns to the idle state. If there are packets in the data queue, however, the dispatcher selects a packet from the queue according to the scheduling discipline. (Non-priority head-of-line queuing is used in the results given here.) Based on the destination of the packet, the NS table is used to determine the sector into which the packet should be transmitted (which we shall refer to as sector S). The network layer then determines if the control channel and at least one data channel are both available for both transmission and reception in sector S. If so, a *traffic list* of all the traffic channels available for both transmission and reception in the sector is generated, and the CU table is updated to indicate that the control channel and all the data channels in the traffic list are not available for transmission or reception in any of the sectors. The network layer then removes the packet from the queue and sends it to the MAC layer together with the traffic list. It also adds the data packet to the OP table and then returns to the idle state. If instead the packet can not be transmitted due to the lack of an available channel, the data queue is searched according to the scheduling discipline until the first data packet that can be transmitted (if any) is found.

Note that at most one packet from the data queue is dispatched before the network-layer process returns to the idle state since the control channel is blocked for both transmission and reception in all sectors. (This does not preclude multiple data-packet transmissions that partially overlap in time due to sequential instantiations of the network-layer process if the transmissions are into distinct sectors and use distinct traffic channels, however. This is made apparent by the following description of the MAC protocol.)

A.3 Baseline MAC Protocol

First consider the behavior of the MAC layer when a packet is dispatched from the network layer. If a data packet is received from the network layer, the MAC layer is given the traffic list associated with the packet and it determines the neighbor to which the packet is to be transmitted. The traffic list is included in a RTS it transmits into the corresponding sector on the control channel using the neighbor's receiver-directed spreading code. The MAC layer then waits for either a valid CTS packet to arrive from the neighbor on the control channel in the same sector or for a timeout to occur. If a timeout occurs, the CU table is updated to release the control channel and all the traffic channels in the traffic list specified in the RTS, and the MAC layer returns to the idle state. (When the MAC layer updates the CU table to release either a control or a traffic channel, the network layer is informed through an interrupt. And the receiver for each sector is tuned to the control channel whenever the MAC layer returns to the idle state.) If a valid CTS is received instead, the MAC layer updates the CU table to release the control channel and all the traffic channels in the data packet's traffic list except the one traffic channel designated in the CTS. It then transmits the data packet into the corresponding sector on the designated traffic channel using the neighbor's receiver-directed spreading code, and it then waits for an acknowledgment or negative acknowledgment from the neighbor to arrive on the same traffic channel or for a timeout to occur. In either case it then updates the CU table to release the traffic channel and sends notice of the event to the network layer. The MAC layer then returns to the idle state.

If a PROP packet is received from the network layer, the MAC layer is given the traffic channel to be used for transmission of the packet. A RTS is transmitted into the corresponding sector on the control channel using the common spreading code, and the CU table is then updated to release the control channel. The PROP packet is transmitted into the corresponding sector on the designated traffic channel using the common spreading code immediately upon completion of the transmission of the RTS. The CU table is then updated to release the traffic channel, notice is sent to the network layer, and the MAC layer returns to the idle state.

Now consider the behavior of the MAC protocol when it receives an interrupt from the physical layer. If the interrupt from the physical layer is due to the arrival of a RTS for a data packet, the MAC layer determines if the control channel is available for transmission. If so, a list of the traffic channels available for both transmission and reception in the sector is generated, and a channel that is shared in common by the traffic list specified in RTS and the locally generated list is chosen at random. (If there is no channel shared in common, the RTS is ignored and the MAC layer returns to the idle state.) Then the CU table is updated to indicate that the control channel is not available for reception in any sector and that the selected traffic channel is not available for either transmission or reception in any sector. Following this, a CTS addressed to the relevant node and specifying the chosen traffic channel is transmitted into the corresponding sector on the control channel using the common spreading code. The CU table is then updated to release the control channel for reception. Following this the receiver in the sector is tuned to the selected traffic channel, and the MAC layer waits for the data packet to arrive or for a timeout to occur. If a timeout occurs, the CU table is updated to release the selected traffic channel and the MAC layer returns to the idle state. If instead the data packet is received, it is sent to the network layer. If the data packet decodes correctly, an acknowledgment packet is transmitted to the other node using the same traffic channel on which the data packet arrived. If the data packet does not decode correctly, a negative acknowledgment is transmitted on the traffic channel instead. In either case, the neighbor's receiver-directed spreading code is used, and in either case, the MAC layer then returns to the idle state.

If the interrupt from the physical layer is due to the arrival of a RTS for a PROP packet, the MAC layer determines the traffic channel on which the subsequent PROP will be transmitted. It then updates the CU table to indicate that the designated traffic channel is not available for transmission in any sector. (If the traffic channel is not available for reception, the MAC layer instead ignores the RTS and returns to the idle state.) The MAC layer waits for the PROP packet to arrive on the designated traffic channel in the corresponding sector or for a timeout to occur. If a timeout occurs, the CU table is updated to release the traffic channel and the MAC layer returns to the idle state. If instead the PROP packet is received, it is sent to the network layer. The CU table is updated to release the traffic channel and the MAC layer returns to the idle state. Finally, if the interrupt received from the physical layer is due to a CTS packet not destined for the node (i.e., an overheard CTS), the CU table is updated to indicate that the traffic channel specified in the CTS is not available for either transmission or reception in the corresponding sector for the amount of time specified in the CTS. Following this the MAC layer returns to the idle state.

Appendix B

Detailed Description of NCTS MAC Protocol

Each node maintains a *NCTS table* to augment the channel-usage table. The NCTS table contains information about the state of the traffic channels at each neighboring node, and the information is obtained from NCTS packets received on the control channel. Each entry in the table specifies the identifier for a neighboring node and one of the traffic channels, the earliest time at which the local node may use the traffic channel to transmit to the neighboring node (*time-to-use*), the *channel-use counter*, and a Boolean information-source designator. The channel-use counter is used to update the time-to-use in the manner described below, and the information-source designator indicates if the information in the table entry has been derived from an overheard NCTS or from a NCTS addressed to the local node. The scheduler and dispatcher differ only slightly from those in the baseline MAC protocol. There is one additional class of time interrupts that can invoke the scheduler (described below), and the dispatcher checks the NCTS table in addition to the CU table when determining the availability of traffic channels for transmission of a data packet.

Suppose a NCTS is transmitted in response to a RTS that contains the RTS packet's traffic list. The NCTS includes the following information: the address of the node that sent the corresponding RTS, the time at which each of the designated traffic channels will be available at the node sending the NCTS, and the number of NCTS packets that have been sent by the local node for each traffic channel in the list since the the current blocking interval began for that traffic channel. The latter entry is denoted as *Num NCTS*. The value of *Num NCTS* serves as a estimate of the number of nodes that may request the use of the traffic channel when it becomes available. (If a node transmits a RTS to a neighbor and does not receive either a CTS or a NCTS, the node employs an exponential back-off for retransmissions to that neighbor.)

Suppose an overheard NCTS packet is received at a node/sector. For each channel in the packet's traffic list, the entry corresponding to the traffic channel and the node/sector that transmitted the NCTS packet is examined in the NCTS table of the local node/sector. If the table entry is current (i.e., its time-to-use value exceeds the current time) and its information-source designator is set to "response NCTS," , the entry is not altered. Otherwise, the channel-use counter is set to the value of *Num NCTS* specified for that traffic channel in the NCTS packet, and the information-source designator is set to "overheard NCTS." Furthermore, the time-to-use in the entry is set to the time of availability listed for the channel in the NCTS packet plus an offset constant times a randomly selected integer between one and the channel-use counter. (The offset constant is the sum of the RTS and CTS transmission times and the largest possible round-trip propagation delay, and the random offset is designed to mitigate contention when a blocked channel becomes available simultaneously at two or more nodes blocked by the same overheard CTS.)

Suppose instead that the node/sector receives a NCTS packet in response to a RTS that it has transmitted. Again the appropriate NCTS table entry is examined for each channel in the packet's traffic list. The channel-use counter is set to the value of *Num CTS* in the NCTS packet, and the information-source designator is set to "response NCTS." The node further determines if its NCTS table has another current entry for the same traffic channel at any other neighboring nodes with its channel-user counter equal to one. If so, the channel-use counter in the table entry that is being modified is incremented by one. If the resulting channel-use counter is equal to one, the time-to-use is set to the channel availability time specified in the NCTS packet. Otherwise, the time-to-use is set to the channel availability time plus the same offset constant as above times a randomly selected integer between one and one less than the channel-use counter. In each case, the data packet for which the MAC layer had attempted transmission is placed back into the OP table in the same location from which it was previously removed. It also selects the smallest value of time-to-use among the NCTS table entries that have been altered, and it sets an interrupt timer to expire at that time for an interrupt that invokes the scheduler.

REFERENCES

1. L. C. Godara, "Application of antenna arrays to mobile communications, part I: Performance improvement, feasibility and system consideration," *Proc. IEEE*, vol. 85, no. 7, pp. 1031–1060, July 1997.

2. L. C. Godara, "Application of antenna arrays to mobile communications, part II: Beam-forming and direction-of-arrival considerations," *Proc. IEEE*, vol. 85, no. 8, pp. 1195–1245, Aug. 1997.

3. J. C. Liberti and T. S. Rappaport, *Smart antennas for wireless communications: IS-95 and third generation CDMA applications*. Upper Saddle River, NJ: Prentice Hall, 1999.

4. G. J. Foschini and M. J. Gans, "On limits of wireless communication in a fading environment when using multiple antennas," *Wireless Pers. Commun.*, vol. 6, no. 3, pp. 311–335, Mar. 1998.

5. M. B. Pursley, "The role of spread spectrum in packet radio networks," *Proc. IEEE*, vol. 85, no. 7, pp. 1031–1060, July 1997.

6. T. S. Yum and K. W. Hung, "Design algorithms for multihop packet radio networks with multiple directional antennas stations," *IEEE Trans. Commun.*, vol. 40, no. 11, pp. 1716–1724, Nov. 1992.

7. R. Ramanathan *et al.*, "Ad hoc networking with directional antennas: A complete system solution," *IEEE J. Select. Areas Commun.*, vol. 23, no. 3, pp. 496–506, Mar. 2005.

8. M. Takai *et al.*, "Directional virtual carrier sensing for directional antennas in mobile ad hoc networks," in *Proc. ACM MobiHoc*, Lausanne, Switzerland, June 2002, pp. 188–193.

9. R. R. Choudhury *et al.*, "Using directional antennas for medium access control in ad hoc networks," in *Proc. 2002 ACM MOBICOM*, Atlanta, GA, Sept. 2002, pp. 59–70.

10. Y. B. Ko, V. Shankarkumar, and N. Vaidya, "Medium access control protocols using directional antennas in ad hoc networks," in *Proc. IEEE INFOCOM*, Atlanta, GA, Mar. 2000, pp. 13–21.

11. *Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specification*, IEEE Std. 802.11, Aug. 1997.

12. V. Bharghavan *et al.*, "MACAW: A medium access protocol for wireless LANs," in *Proc. 1994 ACM SIGCOMM*, London, UK, Sept. 1994, pp. 212–225.

13. Y.-B. Ko and N. Vaidya, "Location-aided routing (LAR) in mobile ad hoc networks," *ACM/Baltzer Wireless Networks (WINET)*, vol. 6, no. 4, pp. 307–322, July 2000.

14. A. Nasipuri *et al.*, "On-demand routing using directional antennas in mobile ad hoc networks," in *Proc. IEEE International Conference on Computer Communication and Networks*, Las Vegas, NV, Oct. 2000, pp. 535–541.

15. S. Roy *et al.*, "A network-aware MAC and routing protocol for effective load balancing in ad hoc wireless networks with directional antennas," in *Proc. 4th ACM Internat. Symposium on Mobile Ad Hoc Networking and Computing*, Annapolis, MD, June 2003, pp. 88–97.

16. A. R. Raghavan and C. W. Baum, "An unslotted multichannel channel-access protocol for distributed direct-sequence networks," *Mobile Networks and Applications*, vol. 5, no. 1, pp. 49–56, Mar. 2000.

17. M. Steenstrup, Ed., *Routing in Communications Networks.* Englewood Cliffs, NJ: Prentice Hall, 1995.

18. M. B. Pursley and H. B. Russell, "Routing in frequency-hop packet radio networks with partial-band jamming," *IEEE Trans. Commun.*, vol. 41, no. 7, pp. 1117–1124, July 1993.

19. M. B. Pursley and H. B. Russell, "Network protocols for frequency-hop packet radios with decoder side information," *IEEE J. Select. Areas Commun.*, vol. 12, no. 4, pp. 612–621, May 1994.

20. J. Jubin and J. D. Tornow, "The DARPA packet radio network protocols," *Proc. IEEE*, vol. 75, no. 1, pp. 21–32, Jan. 1987.

21. A. R. Raghavan, C. W. Baum, and H. B. Russell, "A distance-vector routing protocol with consistency checking for mobile distributed direct-sequence packet radio networks," in *Proc. 1999 IEEE Military Commun. Conf.*, Atlantic City, NJ, Nov. 1999, pp. 780–784.

22. J. Westcott and J. Jubin, "A distributed routing design for a broadcast environment," in *Proc. 1982 IEEE Military Commun. Conf.*, vol. 3, Boston, MA, Oct. 1982, pp. 10.4.1–5.

23. D. L. Noneaker, A. R. Raghavan, and C. W. Baum, "The effect of automatic gain control on serial, matched-filter acquisition in direct-sequence packet radio communications," *IEEE Trans. Veh. Technol.*, vol. 50, no. 4, pp. 1140–1150, July 2001.

24. M. B. Pursley and D. J. Taipale, "Error probabilities for spread-spectrum packet radio communications with convolutional codes and Viterbi decoding," *IEEE Trans. Commun.*, vol. 35, no. 1, pp. 1–12, Jan. 1987.

25. I. M. Onyszchuk, “On the performance of convolutional codes,” Ph.D. dissertation, California Institute of Technology, Pasadena, CA, 1990.

26. F. A. Tobagi and L. Kleinrock, “Packet switching in radio channels: Part II: The hidden terminal problem in carrier sense multiple-access and the busy-tone solution,” *IEEE Trans. Commun.*, vol. COM-23, no. 12, pp. 1417–1433, Dec. 1975.

27. V. Bharghavan, “Performance evaluation of algorithms for wireless medium access,” in *Proc. 1998 IEEE Performance and Dependability Symposium*, Raleigh, NC, 1998, pp. 1597–1603.

28. A. Muqattash and M. Krunz, “CDMA-based MAC protocol for wireless ad hoc networks,” in *Proc. 2003 ACM MobiHoc*, Annapolis, MD, June 2003, pp. 153–164.

29. H. Zhai *et al.*, “A dual-channel MAC protocol for mobile ad hoc networks,” in *Proc. 2004 IEEE Globecom*, Dallas, TX, Nov. 2004, pp. 27–32.

30. S. Ramanathan, “A unified framework and algorithm for channel assignment in wireless networks,” *Wireless Networks*, vol. 5, no. 2, pp. 81–94, May 1999.

31. M. Krunz, A. Muqattash, and S. J. Lee, “Transmission power control in wireless ad hoc networks: Challenges, solutions, and open issues,” *IEEE Network*, vol. 18, no. 5, pp. 8–14, Sept. 2004.

32. J. Monks, V. Bharghavan, and W. W. Hwu, “A power controlled multiple access protocol for wireless packet networks,” in *Proc. 2001 IEEE Infocom Conf.*, Anchorage, AK, Apr. 2001, pp. 219–228.

33. A. Muqattash, M. Krunz, and W. E. Ryan, “Solving the near-far problem in CDMA-based ad hoc networks,” *Ad Hoc Networks*, vol. 1, pp. 435–453, Nov. 2003.

34. T. A. ElBatt and A. Ephremides, “Joint scheduling and power control for wireless ad-hoc networks,” in *Proc. 2002 IEEE Infocom Conf.*, New York, NY, June 2002, pp. 976–984.

35. M. B. Pursley and C. S. Wilkins, “Adaptive transmission for direct-sequence spread-spectrum communications over multipath channels,” *International J. Wireless Inform. Networks*, vol. 7, no. 2, pp. 487–495, Sept. 2001.

36. A. Swaminathan, D. L. Noneaker, and H. B. Russell, “The receiver blocking problem in a DS mobile ad hoc network with directional antennas,” in *Proc. 2004 IEEE Military Commun. Conf.*, Monterey, CA, Nov. 2004, pp. 920–926.

37. S. J. Lee and M. Gerla, "Split multipath routing with maximally disjoint paths in ad hoc networks," in *Proc. 2001 IEEE Internat. Conf. Commun.*, Helsinki, Finland, June 2001, pp. 3206–3210.

38. H. Hassanein and A. Zhou, "Routing with load balancing in wireless ad hoc networks," in *Proc. 4th ACM Internat. Workshop on Modeling, Analysis and Simulation of Wireless and Mobile Systems*, Rome, Italy, July 2001, pp. 89–96.

39. L. Chen and W. Heinzelman, "QOS-aware routing based on bandwidth estimation for mobile ad hoc networks," *IEEE J. Select. Areas Commun.*, vol. 23, no. 3, pp. 561–572, Mar. 2005.

40. S. J. Lee and M. Gerla, "Dynamic load-aware routing in ad hoc networks," in *Proc. 2001 IEEE Internat. Conf. Commun.*, Helsinki, Finland, June 2001, pp. 3201–3205.

41. M. B. Pursley, H. B. Russell, and J. S. Wysocarski, "Route adaptation in response to changes in network activity in mobile frequency-hop networks," in *Proc. 2004 IEEE Military Commun. Conf.*, Monterey, CA, Nov. 2004, pp. 1597–1603.

42. D. Bertsekas and R. Gallager, *Data Networks*, 2nd ed. Upper Saddle River, New Jersey: Prentice Hall, 1992, ch. 5.

43. M. R. Pearlman *et al.*, "On the impact of alternate path routing for load balancing in mobile ad hoc networks," in *Proc. MOBIHOC 2000*, Boston, MA, June 2000, pp. 150–156.

44. D. B. Johnson and D. A. Maltz, "Dynamic source routing in ad hoc wireless networks," in *Mobile Computing*, T. lmielinski and H. Korth, Eds. Kluwer Academic Publishers, 1996, ch. 5, pp. 153–181.

45. A. Khanna and J. Zinky, "The revised ARPANET routing metric," in *Symposium Proc. Commun. Architectures and Protocols*, Austin, TX, Sept. 1989, pp. 45–56.