THROUGHPUT FOR PACKET RADIO NETWORKS WITH DYNAMIC POWER LEVELS

by

Jeffrey Neil Marcus

S.B. Massachusetts Institute of Technology
(1982)

Submitted to the Department of
Electrical Engineering and Computer Science
in Partial Fulfillment of the
Requirements of the Degree of

MASTER OF SCIENCE

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 1984

c Massachusetts Institute of Technology 1984

Signature of Author: J. Marcus
Department of Electrical Engineering and Computer Science,
June 18,1984

Certified by: ______________________
Thesis Supervisor

Accepted by: ______________________
Chairman, Departmental Committee on Graduate Students

THROUGHPUT FOR PACKET RADIO NETWORKS WITH DYNAMIC POWER LEVELS

by

Jeffrey Neil Marcus

Submitted to the Department of Electrical Engineering and Computer Science in partial fulfillment of the requirements of the Degree of Master of Science

ABSTRACT

A packet radio network consists of a set of computers whose intercommunication needs are served by mobile radios. In this work, relatively simple models are used to determine attainable throughputs in such networks under a variety of routing and scheduling strategies. In particular, we assume uniform traffic requirements between all pairs of nodes and an idealized capture effect.

Through the derivation and numerical solution of a differential equation, it is shown that a large network of nodes which are uniformly spaced along a line and which uses slotted-Aloha channel access can attain a throughput of 2.05/e if its transmitters can dynamically adjust their power, as opposed to a 1/e throughput attainable in fixed-power networks.

An upper bound of 2 in line networks is then derived and a collision-free scheduling algorithm, Distance-based Time Division Multiple Access (DTDMA), which can approach this bound in uniformly spaced line networks, is introduced. A modified form of DTDMA for use in networks whose nodes are placed on a regular grid is shown to yield a throughput near $\sqrt{n}$, where n is the number of nodes in the network.

Cellular DTDMA (CDTDMA), in which radios are organized into cells, is introduced for use in networks whose nodes are randomly placed along a line. It is shown that a throughput approaching the upper bound of 2 can be attained in such networks when the number of radios per cell is large. An approach based upon the law of large numbers is used to show that a modified form of CDTDMA for use in networks whose nodes are placed randomly on a plane yields a throughput of $O(\sqrt{n/\ln n})$ when the number of radios per cell is large and suggestions are made to improve this performance.

Thesis Supervisor: Professor Robert G. Gallager

Title: Professor of Electrical Engineering

ACKNOWLEDGEMENTS

I would like to express my appreciation to Professor Robert Gallager, my thesis advisor, whose insight and patience contributed greatly towards both the quality of the thesis and the experience of doing it.

I also wish to express gratitude to my parents, Robert and Barbara, who have encouraged me throughout my education and, indeed, throughout my life.

The support from the National Science Foundation contract number NSF-ECS-8310698 and the Advanced Research Projects Agency of the Department of Defense contract number ONR-N000114-75-C-1183 has been greatly appreciated.

Finally, I should mention my friends in the Wednesday Lunch Club, who have introduced me to Thai food and have otherwise improved the quality of my life during the period I worked on this thesis.

TABLE OF CONTENTS

LIST OF FIGURES

1. INTRODUCTION

1.1 The Packet Radio Network Model

A packet radio network (PRNET) consists of a set of computers whose intercommunication needs are served by mobile radios (PR's). Packets of data originating at some computer in the network may be transmitted to their destinations directly or may be forwarded towards their destinations by intermediate radios, called repeaters. In this paper, we assume that the repeater roles are assumed by radios which also serve host computers so that there are no distinctions between the two, although in a real network this may not be true.

Figure 1.1.1 depicts several nodes in a PRNET and illustrates the important characteristics of the model we impose on the network. First of all, we assume that antennae are omnidirectional and that with every transmission there exists a "radius of transmission". We say that every PR within this radius "hears" the transmission. If a PR hears more than one transmission simultaneously, the packets collide and no packet is received successfully. No PR outside this radius hears the transmission. In other words, as long as the received power of some transmission exceeds some threshold, the receiver is within the transmission radius and hears the transmission; it is successfully received if no other transmission Thus, in Figure 1.1.1, A's transmission to B is heard by C and C cannot successfully receive a packet.

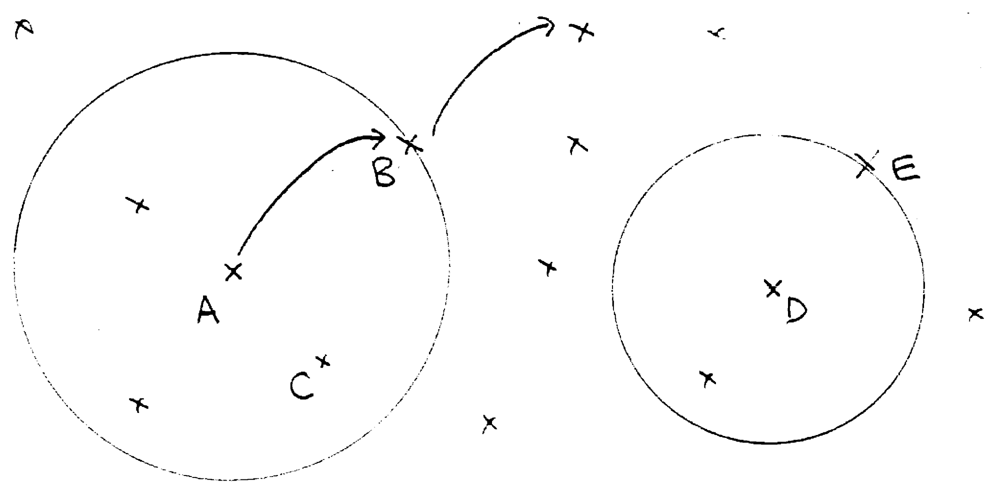

Fig. 1.1.1 A Packet Radio Network

However, D can successfully transmit to E since E is outside the radius. Note also that radio B may not be the final destination of the packet from A but may be acting as a repeater and may subsequently forward the packet to another radio.

The concept of a well-defined radius of transmission is a simplification of the "capture effect" described in [Robe72]. A packet can be successfully received as long as the ratio of its received power to the received power of any other packet exceeds the "capture ratio". We make this simplification so as to simplify our analysis and believe that the results obtained with this model correspond qualitatively to those that would be obtained if the capture effect were included.

Our goal in this paper is to develop and analyze routing and scheduling algorithms for PRNETS. We cannot accomplish this by simply applying work which has been done for other types of networks because, as is clear from the above description, PRNETS are different. Their use of radio repeaters distinguishes them from cellular radio networks, which use fixed transmission towers to forward messages between radios. Additionally, satellite networks and ground radio networks which use a single station to retransmit all messages between radios do not allow more than one successful transmissions at a time. As shown in Figure 1.1.1, this is untrue for a PRNET. The most general difference is that all the networks mentioned above only have a single receiver: the transmission tower in cellular radio networks, the satellite in satellite networks, and the central station in ground radio networks. Thus, implementation and analysis of algorithms is inherently simpler in these networks than it is in packet radio networks, which contain many units which can act as receivers. Finally, radio mobility and the effect of transmission power on the network's connectivity cause the network's topology to be fluid, unlike wire and satellite networks, which have a relatively fixed topology. All these characteristics will be taken into account in our work.

1.2 Definition of Some Terms

Before discussing previous work in the field and introducing our own, we define several terms especially relevant to our work. Other terms and relevant notation will be defined as they are introduced in subsequent chapters.

The performance measure we use to evaluate the algorithms proposed in this paper is the "network throughput". This is defined as the expected number of packets successfully received at their destinations per unit time under a given set of conditions. The particular unit of time is a "slot", which is defined as the time required to transmit a packet. We assume that all packets require the same time, i.e, a slot, for transmission.

At times, we will discuss performance in terms of the "order" of some quantity rather then its precise value. For example, if the throughput cannot be determined exactly but it is known that as the network grows large, it is proportional to n, the number of nodes, then we say that the throughput is $O(n)$. In general, if $g(x)=O(f(x))$, $x \to x_0$, then

$$\lim_{x \to x_0} |f(x)/g(x)|=c, \text{ c some constant}$$

In our discussions, x_0 will generally be infinity and the functional form $f(x)$ will not be used explicitly. However, it will be clear from the context that we generally use this type of analysis in cases where the network size, n, tends to infinity.

The networks we investigate in this paper include "line networks", in which all radios (nodes) are located on a line, "ring networks", in which they are located on the perimeter of a circle, and "planar networks", in which they are located on a plane. In each of these cases, we deal with both "regular networks", in which nodes are placed deterministically and in a geometrically regular fashion, and "random networks", in which nodes are placed randomly.

In order to specify a given network's mode of operation, we must specify its "traffic matrix", which specifies the average number of packets per unit of time which must be sent between each origin-destination pair of nodes. We are particularly concerned here with a "uniform traffic matrix", in which an equal number of packets must travel between each such pair per unit of time.

It is also necessary to specify the "channel access scheme" used in network operation. One scheme considered in this paper is "slotted-Aloha", introduced in [Robe72]. In slotted-Aloha, all transmissions begin at discrete points in time, referred to as slot boundaries. A radio, upon receiving a packet to be transmitted (either from another radio or from its host computer), sends it off at the next slot. If there is a collision at the receiver and the packet is thus not succssfully received, the transmitter re-transmits the packet a random number of slots later. Another scheme is "Time Division Multiple Access" (TDMA), in which each slot is reserved for one transmission between two specific radios and thus, there are no collisions.

Finally, the method of routing packets between origin and destination must be specified. One method of routing is to simply transmit a packet from its origin to its destination in one hop. For obvious reasons, this is called one-hop routing. Alternatively, multi-hop routing, which requires repeaters, may be used.

1.3 Previous Work

Several workers have studied the problem of routing in packet radio networks using various approaches. [Kahn78] proposes several distributed algorithms for implementing routes], along with a survey of issues involved in the architecture and operation of PRNETS. Baker and Ephremides [Bake81] propose a distributed algorithm for the self-organization of a network into "clusters", each of which contains a "cluster head" that acts as a central controller. These papers differ from our work in that they are concerned with implementation of workable routes while we emphasize analysis of different routing strategies and assume that they can be implemented somehow.

A more analytical approach is taken by Kleinrock and Silvester in [Klei78], in which the PRNET is modelled as a random graph. Slotted-Aloha channel access is assumed and transmission radii are assumed to be fixed and equal for every transmission. It is claimed that this radius should be chosen such that each radio can be heard by roughly 6 neighbors so as to maximize throughput. The derivation of this result requires the assumption that the network is connected, i.e, that each radio can be heard by at least one neighbor, and simulation is used to justify the validity of this assumption. An inconsistency in this paper is discovered by Takagi and Kleinrock in [Taka84], and the optimal number of neighbors is revised to 8.

Silvester and Kleinrock [Silv83a] consider line, ring, and planar networks whose nodes are regularly spaced and determine network throughput as a function of the degree of connectivity (the number of potential receivers for each transmitter) assuming slotted-Aloha operation. The major conclusions of this work are:

(1) In a ring network with an optimal degree of connectivity, a throughput of 2/e can be attained.

(2) In a line network, a throughput of c/e, where c is some constant, can be attained almost independently of connectivity.

(3) In a two-dimensional network, a throughput on the order of the square root of the number of nodes in the network is achievable.

These conclusions are substantiated and made more precise in our work under a different set of assumptions about network topology and transmitter operation.

Finally, Silvester and Kleinrock [Silv83b] consider networks whose nodes are randomly placed and obtain results for network throughput in a one-hop routing environment in which a traffic matrix is assumed in which each radio only transmits to one other radio. Results are obtained for transmitter-receiver pairs chosen randomly and for those chosen to optimize throughput. These results are not directly applicable to our work because we consider networks in which each transmitter must communicate with every other receiver in the network, i.e., a uniform traffic matrix. Thus, transmitter-receiver pairs cannot be chosen either arbitrarily or randomly but must be chosen so as to satisfy the traffic requirements, either through one-hop or multi-hop routing. However, in contrast to the previous work mentioned, [Silv83b] does assume that transmission radii are adjustable. This assumption is also made in our work.

1.4 Motivation and Outline of this Research

In this paper, we propose several algorithms for routing and scheduling channel access in packet radio networks. The major novel element of our work is that we assume first that radios have unlimited power and can adjust their transmitting power every transmission and second that each radio has perfect information about the location of all other radios in the network, as well as its own location. Thus when one radio wishes to transmit to another, it adjusts its transmitting radius so as to just include the intended receiver. Figure 1.4.1 illustrates this idea by showing the transmitting radii that radio A would use to transmit to radio B and to radio C. Note that other radios in the network may hear these transmissions, depending on whether or not they are located

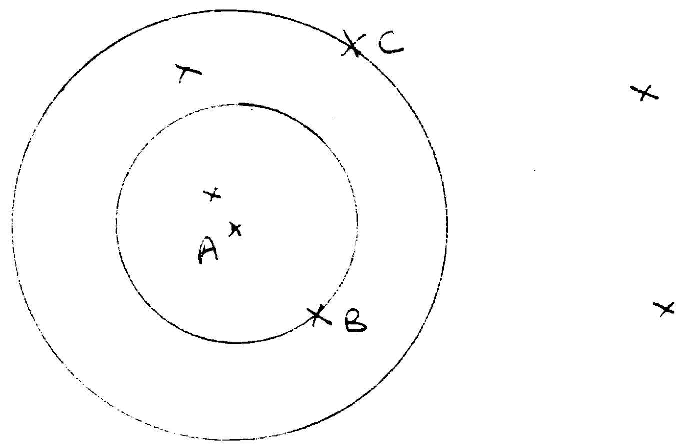

Fig. 1.4.1 Transmitter with Adjustable Power

within the radius of transmission.

While these assumptions are clearly simplifications of reality, they are reasonable within limits. It is indeed technically possible to adjust power (as some commercial radio stations do at night). The assumption of unlimited power clearly depends on the size of the network and the nature of the transmitters but, in any case, we investigate several strategies which do not depend on unlimited power. Finally, while perfect information about radio location is impossible due to noise and radio mobility, for slow-moving radios a small amount of bandwidth could be devoted to updating position information throughout the network.

Our motivation for studying dynamically-powered networks is that from a technical viewpoint, such networks are desirable in that they circumvent the problem of an unconnected network, since a radio can adjust its power so as to heard by at least one other radio. Moreover, a radio can adjust its power so as to be heard by a repeater in the direction of the destination radio. Thus, the forward progress made during a hop is controllable and is not a random variable dependent only upon network topology, as was assumed in [Klei78] and [Taka84]. The ability to control power thus leads to a richer set of routing and scheduling strategies and the goal of our work is to develop insight into these.

We first consider regular line networks. In Chapter 2, we evaluate slotted-Aloha under multi-hop and one-hop routing for both very small and very large regular line networks and show that one-hop routing is superior. Our work differs from that of [Silv83a] in that we assume the case of dynamic transmitter power and consider the effect of a radio's location on its operation while Silvester looks at operation only for radios near the middle of the network.

The remainder of the paper is concerned with scheduled channel access schemes. In Chapter 3, we derive an upper bound on the throughput in any line network using any routing and channel access strategies and propose Distance-based Time Division Multiple Access (DTDMA), a scheduling algorithm which approaches this upper bound. We show how this scheme can be applied to regular planar networks.

In Chapter 4, we look at random line networks and propose Cellular Distance-based Time Division Multiple Access (CDTDMA), which is based upon DTDMA but can be used in random networks. We show that under certain conditions, the upper bound on throughput derived in Chapter 3 can be attained in random line networks with this algorithm. We then analyze the performance of this algorithm in random planar networks.

Finally, Chapter 5 contains some half-baked ideas which can be used as bases for further research.

2. SLOTTED-ALOHA IN UNIFORM LINE NETWORKS

2.1 Assumptions

In this chapter, we determine throughputs obtainable in regular line networks which use slotted-Aloha channel access. We assume that the PRNET operates in the manner described in Chapter 1: namely, each transmission radius is adjusted so as to barely include the intended receiver and no radio within this radius can simultaneously receive a packet from another transmitter.

In determining the throughput, we impose a uniform traffic matrix, so that flows between all source-destination pairs are equal.

Under these assumptions, we compare throughputs obtainable in networks which use multihop routing, those which use one-hop routing where transmission radii are adjusted, and fully connected networks where each transmission is heard by all radios, regardless of the intended receiver's location.

In multihop routing, we consider the case where each packet is sent a distance of N radios in the direction of the destination on each hop, until it reaches the destination. We assume that $N \ll n$, the number of radios in the network, so that it typically takes many hops to transmit a packet from source to destination. In the one-hop case, all packets are transmitted in one hop, regardless of distance.

For the remainder of this chapter, $r(i,j)$ will denote flow between each source-destination pair (i,j), $t(i,j)$ the expected number of successful transmissions (also referred to as "traffic") per slot on link (i,j) and p(i,j) the probability of transmission on this link during a slot. The network throughput will be denoted R and is the sum of the flows on each link in the network. Thus,

$$R=\sum_{i,j\ ;i\neq j} r(i,j) \qquad (2.1.1)$$

It is also convenient to define an interference factor q(k) as the probability that radio k hears no transmissions on a given slot, including those intended for it. Assuming independence among transmissions from different radios

$$q(k)=\prod_{i=1}^{n}(1-\sum p(i,h)) \qquad (2.1.2)$$

h such that transmissions from i to h heard by k.

Each of the factors in q(k) represents the probability that radio i does not interfere with a transmission to k. For a successful transmission from radio j to radio k, radio j must transmit to k while k hears no other transmissions. It follows that

$$t(j,k)=p(j,k)q(k)/(1-\sum p(j,h)) \qquad (2.1.3)$$

h such that transmissions from j to h are heard by k

The denominator in (2.1.3) is factored out of q(k) because j cannot interfere with its own transmission to k. It may seem awkward to include this factor in the numerator of (2.1.3) and to then cancel it out. However, this tactic simplifies the calculations regarding large networks which occur in proceeding sections.

2.2 Three-Radio Networks

To illustrate the calculations needed to determine maximum throughput, we will first consider the simplest non-trivial network, one comprising three radios, labelled 1,2, and 3. Using multihop routing with N=1, link (1,2) carries all traffic between nodes 1 and 2 and between nodes 1 and 3 while node (2,3) carries all traffic between nodes 1 and 3 and between nodes 2 and 3. Similar relations hold in the

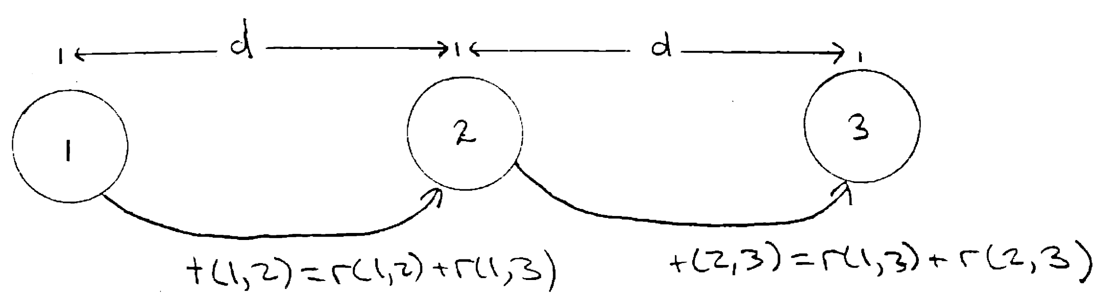

Fig. 2.2.1 A Three Radio Network

opposite direction. This is illustrated in Figure 2.2.1. Formally,

$$t(1,2)=r(1,2)+r(1,3)=2r \quad (2.2.1a)$$

$$t(2,3)=r(2,3)+r(1,3)=2r \quad (2.2.1b)$$

We have made the uniform traffic requirement explicit in the above equations by substituting a constant r for all r(i,j).

To relate the traffic on each link to the transmission probabilities, the following facts are needed:

(1) Radio 2 hears all of its own transmissions and all those from 3 and 1.

(2) Radio 3 hears all of its own transmissions and all of 2's transmissions but does not hear 1's transmissions to 2.

Thus,

$$q(2)=(1-p(2,1)-p(2,3))(1-p(1,2))(1-p(3,2)) \quad (2.2.2a)$$

$$q(3)=(1-p(2,1)-p(2,3))(1-p(3,2)) \quad (2.2.2b)$$

Substituting (2.2.3) and (2.2.4) into (2.1.3) yields:

$$t(1,2)=p(1,2)(1-p(3,2))(1-p(2,1)-p(2,3))=2r \quad (2.2.3a)$$

$$t(2,3)=p(2,3)(1-p(3,2))=2r \quad (2.2.3b)$$

By symmetry, p(1,2)=p(3,2) and p(2,3)=p(2,1) and the resulting equations are:

$$t(1,2)=p(1,2)(1-p(1,2))(1-2p(2,3)) \quad (2.2.4a)$$

$$t(2,3)=p(2,3)(1-p(3,2)) \quad (2.2.4b)$$

Hence,

$$p(2,3)=p(1,2)/(1+2p(1,2)) \quad (2.2.5)$$

To determine maximum throughput, we maximize t(1,2) over all p(1,2) by substituting (2.2.5) into (2.2.3), setting the derivative to 0 and finding the root of the resulting polynomial. It turns out that p(1,2)=.366 and p(2,3)=.211. Using these values, r=.067. To determine R, we solve (2.1.1) under the assumption of a uniform traffic matrix. Since there are n(n-1) transmitter-receiver pairs in the network, this yields

$$R=\sum_{i,j;\ i\neq j} r(i,j)=n(n-1)r. \qquad (2.2.6)$$

For n=3, R=.402 using short transmissions. When n=2, a similar calculation yields R=.5 and for n=4, R=.370. For very small networks, it thus appears that throughput falls with n when multihop routing is used.

A similar calculation for one-hop routing with adjustable power yields p(1,2)=p(1,3)=.1890, p(2,3)=.1590 and R=.481 for n=3. Note that in this case, traffic over pairs (1,2), (2,3) and (1,3) must be accounted for and that t(i,j)=r(i,j) for all pairs (i,j) since traffic between two nodes is carried solely on the link between them. This calculation becomes tedious for larger networks and we have not attempted it for n>3.

Finally, we can make this calculation for networks which are fully connected so that each transmission is heard by all radios. This is just the classical slotted-Aloha network and the maximum throughput is given in [Robe72] to be

$$R=(1-1/n)^{n-1} \qquad (2.2.7)$$

The optimal probability P of transmission to any receiver is shown in this paper to be 1/n so that the probability of transmission between a particular transmitter and receiver must be

$$p(i,j)=P/(n-1)=1/[n(n-1)] \qquad (2.2.8)$$

since all receivers are transmitted to equally frequently. Note that p(i,j) is independent of i and j because all transmissions are heard by every radio, so that radio location is irrelevant. Thus, for n=2, R=.5 and p(i,j)=.5; for n=3, R=.444 and p(i,j)=.167; and for n=4, R=.422 and p(i,j)=.083.

It is worthwhile to make two observations in comparing the results for three-radio networks.

(1) Transmission probabilities are lowest in the cases where one hop is required for each delivery of a packet from origin to destination.

(2) One-hop routing yields a higher maximum throughput than that attained in a fully connected network which in turn attains a higher throughput than that attained in multi-hop networks. This will be shown to be true for very large networks as well under a slotted-Aloha access scheme. We have not investigated intermediate-sized networks but see no reason why this relationship should not hold for these.

2.3 Large Networks Using Multihop Routing

In this section, we calculate maximum throughputs for networks of n radios which use multihop routing in the limit where n becomes infinite. We first consider a network in which all transmissions are to adjacent radios (N=1) and then generalize our results to networks in which all transmissions are of the same distance and this distance is small compared to the length of the network. The conclusion we draw is that short transmission multi-hop strategies yield throughputs which can never exceed that of a fully-connected network, which has been shown in [Robe72] to be 1/e.

2.3.1 The Short Transmission Case (N=1)

In a network of n nodes, with nodes being labelled from left to right with the numbers 1 to n, the throughput was defined to be n(n-1)r in the last section, where r is the number of packets sent between each source-destination pair every slot. To determine the maximum throughput, relations must be derived, as they were in the last section, among each link's traffic and the network throughput. Towards this end, we note that for N=1, link (1,2) must carry all the traffic from link 1 to each of the other n-1 nodes in the network. Therefore,

$$t(1,2)=(n-1)r \qquad (2.3.1.1)$$

Link (2,3) must carry traffic from nodes 1 and 2 to the other n-2 nodes in the network. Thus,

$$t(2,3)=2(n-2)r \qquad (2.3.1.2)$$

In general,

$$t(i,i+1)=i(n-i)r \qquad (2.3.1.3)$$

Note that this is simply a generalization of (2.2.1) for arbitrary networks. In the same manner, (2.2.3a) can be generalized to yield

$$t(i,i+1)=p(i,i+1)[1-p(i+1,i)-p(i+1,i+2)] \times[1-p(i+2,i+1)-p(i+2,i+3)] \qquad (2.3.1.4)$$

This is true because the only transmissions which can interfere with a transmission to node i+1 are those which are sent from nodes i+1, i+2. No other transmission includes node i+1 in its radius. This holds for all links except those at the ends of the network, whose traffic-probability relations are derived by generalizing (2.2.3b).

In principle, we could solve this system to determine the maximum obtainable throughput as we did for the three-radio network. This would involve the reduction of the system to a single equation relating some link traffic to some link probability followed by the maximization of this traffic with respect to the link probability. The throughput would then be obtained by substituting this traffic into (2.3.1.3) to determine r, which would in turn be substituted into (2.3.6) to determine R. However, for large networks, the reduction is tedious and computationally intensive. We therefore determine an approximation to the maximum network throughput which converges to the actual throughput as the number of nodes approaches infinity.

The idea is to work backwards by starting with a single equation relating a link traffic to a link probability. Towards this end, we define a quantity P(i,i+1) and substitute this quantity for each link occurring in (2.3.1.4). Thus, (2.3.1.4) becomes

$$t(i,i+1)=P(i,i+1)[1-2P(i,i+1)]^2=i(n-i)r \qquad (2.3.1.5)$$

We use this equation to derive initial estimates of the set of p(i,i+1).

To maximize throughput, we must choose P(j,j+1) so that traffic on the link (j,j+1) which has the greatest traffic is maximized. Then, the set of P(i,i+1) for all other links can be determined in terms of this traffic using (2.3.1.3) and (2.3.1.5). This set of P(i,i+1) can then be used as initial estimates of p(i,i+1) in (2.3.1.4) to obtain a better estimate of the maximum traffic which can flow on link (j,j+1), which can in turn be used to obtain a better estimate of the network throughput.

From (2.3.1.5) it can be seen that the maximum t(i,i+1) over all i occurs when i=n/2. Thus, assuming a network with an even number of nodes, the most congested link is the middle one, (n/2,n/2+1). Maximizing t(n/2,n/2+1) with respect to

$P(n/2,n/2+1)$ yields

$$P(n/2,n/2+1)=1/6$$
$$t(n/2,n/2+1)=n^2r/4=2/27 \qquad (2.3.1.6)$$

To obtain a better estimate of $t(n/2,n/2+1)$, we must determine $P(n/2+1,n/2)$, $P(n/2+1,n/2+2)$, $P(n/2+2,n/2+1)$, and $P(n/2+2,n/2+3)$. From (2.3.1.3), it can be seen that $t(i,i+1)=t(i+1,i)$ so that $P(i+1,i)=P(i,i+1)$. Thus,

$$P(n/2+1,n/2)=P(n/2,n/2+1)=1/6.$$

Similarly,

$$P(n/2+2,n/2+1)=P(n/2+1,n/2+2).$$

We determine $P(n/2+1,n/2+2)$ by using (2.3.1.5) to obtain a Taylor series expansion about $t(n/2,n/2+1)$. Letting $t'=dt/dp$, we have

$$t' =12p^2-8p+1$$
$$t'' =24p-8$$
$$t'''=24,$$
$$t^{(n)} =0, \; n>3 \qquad (2.3.1.7)$$

Letting $\Delta P(1)=P(n/2+1,n/2+2)-P(n/2,n/2+1)$ the expansion is

$$t(n/2+1,n/2+2)=t(n/2,n/2+1)+t'(n/2,n/2+1)\Delta P(1)$$
$$+t''(n/2,n/2+1)[\Delta P(1)]^2/2$$
$$+t'''(n/2,n/2+1)[\Delta P(1)]^3/6 \qquad (2.3.1.8)$$

Thus,

$$[n^2/4-1-n^2/4]r=0-4[\Delta P(1)]^2+4[\Delta P(1)]^3 \qquad (2.3.1.9)$$

But our initial estimate of r from (2.3.1.6) is $8/(27n^2)$. Thus, manipulation of (2.3.1.9) yields

$$n^2[\Delta P(1)]^2[1-\Delta P(1)]=2/27 \qquad (2.3.1.10)$$

As n grows large $\Delta P(1)$ becomes proportional to 1/n and our estimate of P(n/2+1,n/2+2) approaches 1/6. It can similarly be shown that our estimate of P(n/2+2,n/2+3) approaches 1/6 as n grows large so that the approximation (2.3.1.4) converges to the true expression (2.3.1.5). Thus, the new estimate of t(n/2,n/2+1) converges to 2/27. For large n, $R \cong n^2 r$ and

$$R \cong n^2 r=4(n^2/4)r=8/27=.296 \qquad (2.3.1.11)$$

Thus, we have shown that for very large networks, a throughput of .296 can be obtained with multihop routing in which all transmissions are to adjacent radios.

We believe that the method proposed here can be used as a basis for an iterative algorithm that can approximate the throughput in a finite network with arbitrary accuracy. However, the development and analysis of convergence of such an algorithm is beyond the scope of this paper.

Our results for multihop routing, with N=1, for 2,3, 4 and an infinite number of radios lead us to believe that the throughput falls monotonically from .5 to .296 as n increases due to the increasing number of hops required to deliver a packet from source to destination but we have been unable to prove this.

2.3.2 The General Case

The above analysis can be extended to multihop strategies where each packet travels to its destination in hops of a fixed distance of N radios until it reaches a radio within N radios of the destination and is delivered over the required

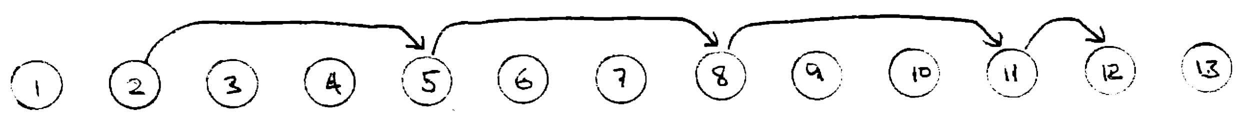

Fig. 2.3.2.1 Multi-hop Routing: N=3

remaining distance (Fig. 2.3.2.1). For networks of n radios, where N/n is small, these "delivery hops" represent a small fraction of the total number of transmissions. More precisely, as n gets large, the distance between radios is typically O(n) and it takes O(n/N) hops to route a typical packet from origin to destination. There is one delivery hop for each packet delivered between an origin-destination pair (including delivery hops of 0 radios which occur when the distance between origin and destination is a multiple of N).

the delivery hops represent a fraction of $O(N/n) \ll 1$ of all transmissions and we will ignore them as we look at networks for which $N \ll n$.

We concentrate our attention on the radios in the middle of the network and assume that they each have the same transmission probability, P_c. The validity of this assumption for large networks was established in the previous section. Whereas, for N=1, the center link's traffic in one direction represented 1/4 of the network throughput, the work is now split up among N middle links. This becomes clear if one subdivides the network into N subnetworks, as shown in Figure 2.3.2.2, where the leftmost node of subnetwork j is the jth node of the original and each adjacent node of subnetwork j

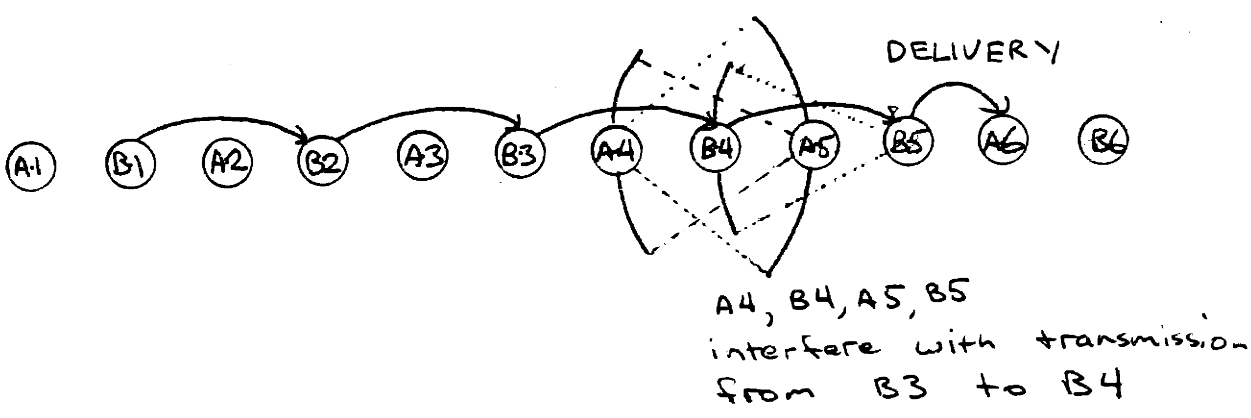

Fig. 2.3.2.2 Subdivision of Network (N=2)

corresponds to the next N-hop. Virtually all transmissions occur within these subnetworks, since the only communication between subnetworks is due to the final delivery of packets, which is negligible, as explained above. Thus, each subnetwork carries 1/N of the traffic on the original network and the center link of the original carries 1/(4N) of the network throughput. A successful transmission on this link requires that each of the 2N nodes to the right do not transmit in either direction, as illustrated in Figure 2.3.2.2. The expression for the throughput in such a network is thus

$$R=4NP_c(1-2P_c)^{2N} \qquad (2.3.2.1)$$

Keeping n fixed and maximizing over P, we have

$$P_c=1/(4N+2) \qquad (2.3.2.2)$$

and substitution of (2.3.2.2) into (2.3.2.1) yields

$$R=(4N/(4N+2))(1-1/(2N+1))^{2N} \qquad (2.3.2.3)$$

As N gets large, R increases asymptotically to 1/e=.368. Thus, the throughput varies from .296 when N=1 to .368 when N gets large as shown in Figure 2.3.2.3.

The significance of this result is that the throughput obtained using the multihop strategies outlined above, coupled with a slotted-Aloha access scheme, is relatively insensitive to the length of the hops and can never exceed the 1/e throughput of a fully-connected network. In the next section, we show how this performance can be roughly doubled by using one-hop routing.

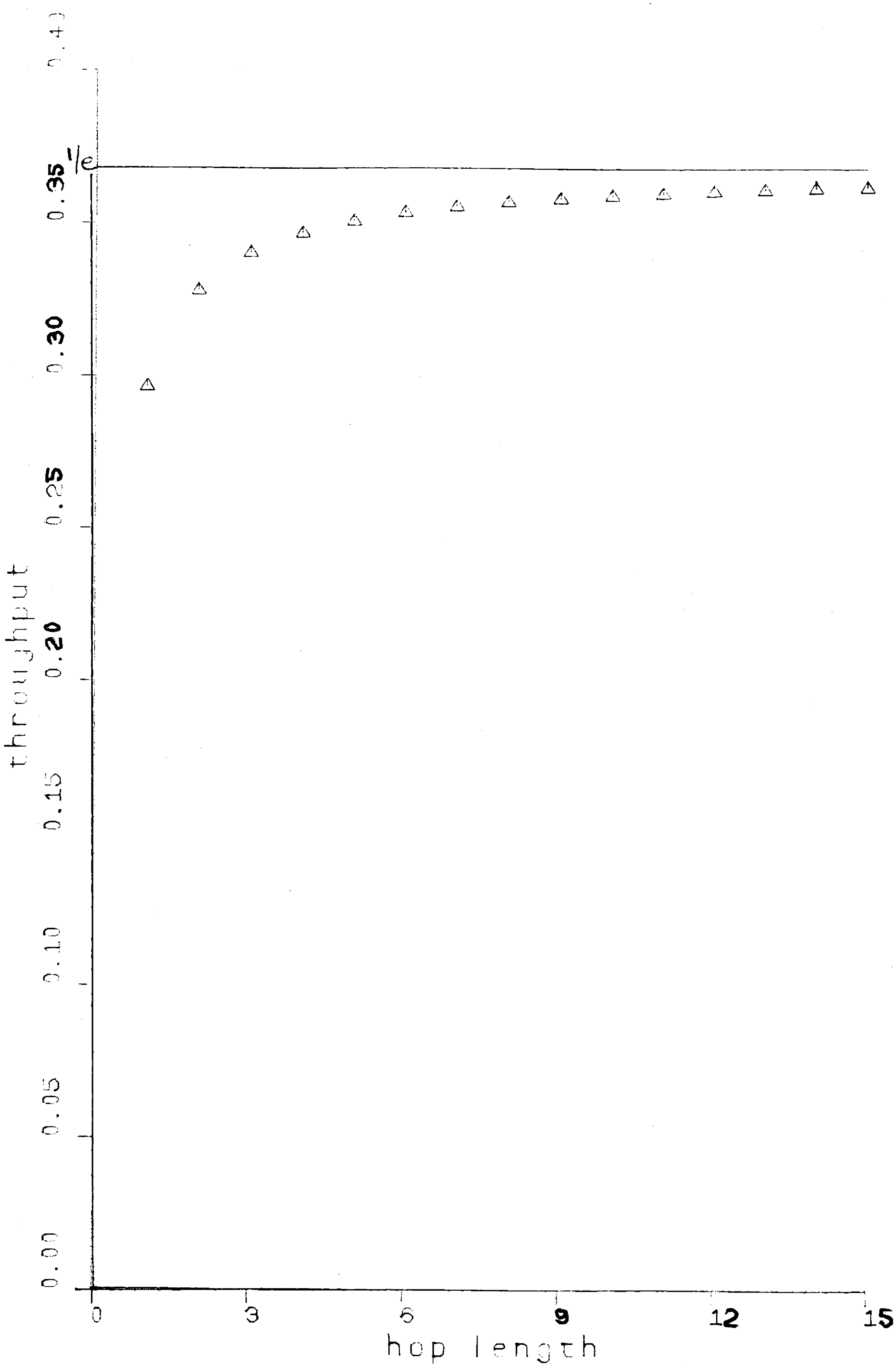

Fig. 2.3.2.3   Throughput vs. Hop Length

2.4 Large Networks Using One-Hop Routing

Given packet radios which can adjust their power so as to be heard by all radios within a radius of the destination and no more and which have complete information about the locations of all other radios in the network, the simplest strategy for routing is to avoid it entirely by transmitting to the destination radio directly. Here, we calculate the throughput obtained by such a strategy. First, a ring network will be considered, simplifying analysis by avoiding the end effects which occur in a line network. Following this analysis, we will study the line network of the last section under one-hop routing.

2.4.1 Ring Networks

Consider the ring network of Figure 2.4.1. For the purposes of the following argument, we will assume that there is an odd number of radios in the ring and these will numbered clockwise from 0 to n-1. This assumption is not critical to the argument but is convenient. The network throughput is n(n-1)r. Because all packets are delivered in one hop, the traffic on any link (i,j) equals the number of packets from i destined for j per slot. Thus,

$$R=n(n-1)r=n(n-1)t \qquad (2.4.1.1)$$

Determining t requires evaluating each of the factors in q(0) which appear on the right side of (2.1.1). This calculation

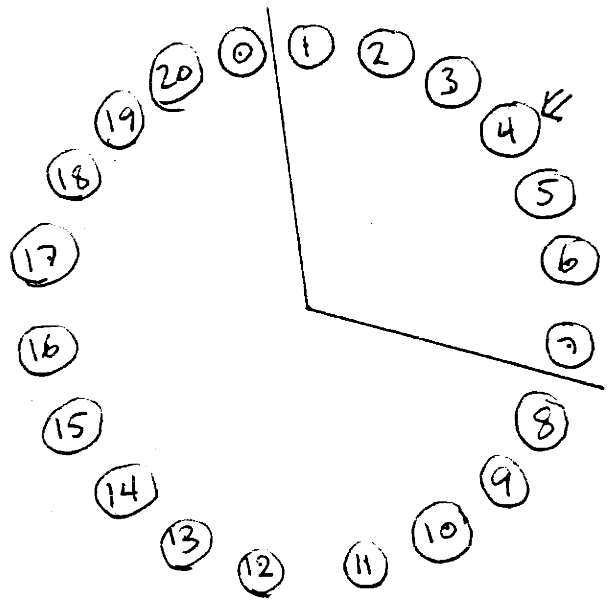

Fig. 2.4.1 A Ring Network

is simplified by assuming that to maintain a uniform traffic matrix, all link probabilities p(i,h) are equal. The assumption is justified as follows: For a large network with n radios, a transmitter's optimal probability of transmitting during a slot is on the order of 1/n, as will shown below for the network considered here. Thus, for large n, the denominator in (2.1.2) can be approximated by unity and (2.1.2) becomes

$$t(j,k)=p(j,k)q(k) \qquad (2.4.1.2)$$

Due to symmetry, q(k) is constant for all k and due to traffic uniformity, t(j,k) is constant for all pairs (j,k). Thus, p(j,k) must be constant for all links and will be denoted p. Applying (2.1.1) to the radio arbitrarily denoted radio 0, we have

$$q(0)=\prod_{i=1}^{n}(1 - f(i)p) \qquad (2.4.1.3)$$

where f(i) denotes the number of destinations for which transmissions from radio i are heard by radio 0. Consider a radio i such that $1\leq i\leq(n-1)/2$. As shown in Figure 2.4.1, with i=4, radio 0 hears transmissions from i to all radios except those within a radius of i-1 from i. There are 2(i-1) such radios. Thus,

$$f(i)=(n-1)-2(i-1)=n-2i+1 \qquad (2.4.1.4)$$

Exploiting symmetry to account for the radios i such that $(n+1)/2\leq i\leq n$, substituting (2.4.1.4) into (2.4.1.3) and making a change in the summation variable,

$$q(0) =[\prod_{i=1}^{(n-1)/2} (1-2ip)]^{2} \qquad (2.4.1.5)$$

Assuming p<<1, approximating ln(1-kp) by -kp and carrying out the summation

$$\ln q(0)=-[(n^{2}-1)/2]p \qquad (2.4.1.6a)$$

$$q(0)=\exp[-(n^{2}-1)p/2] \qquad (2.4.1.6b)$$

Applying equation (2.1.2) and retaining only first-order terms yields

$$t=p\ \exp(-n^2 p/2) \qquad (2.4.1.7)$$

Maximizing t yields

$$\hat{p}=2/n^2$$

$$R=n(n-1)(2/n^2)/e \cong 2/e \qquad (2.4.1.8)$$

Note that the probability of transmission to any other radio is $(n-1)p=2/n$ and the approximation made in (2.4.1.2) is justified.

This result corresponds to that in [Silv83a] for the optimal throughput in large ring networks where each radio is heard by a constant, non-adjustable number of receivers, showing that one-hop routing with adjustable transmission radii performs as well as any strategy based on non-adjustable transmission radii. In the next section, we show that for line networks, one-hop routing yields roughly twice the throughput in a line network as the throughput obtained for the multihop strategies of section 2.3.

2.4.2 Line Networks

Line networks differ from ring networks in that the probability of transmission to a given radio is dependent on its position. In particular, there is more contention near the center of the network and more attempted transmissions are needed here than at the ends to obtain the same number of successfully received packets. However, since we are still dealing with a large network, the approximation made in

(2.4.1.2), i.e, that the probability of transmission is independent of the transmitting radio, is still valid. Thus p(j,k) is dependent on k but not on j. We will implicitly use this fact in our notation by renaming p(j,k) to p(k). It follows from the above argument that q(k) is also dependent on k. Approximating the denominator of (2.1.2) by unity (since n is large) the expression for traffic in this network becomes:

$$t(j,k)=p(k)q(k) \qquad (2.4.2.1)$$

Although t(j,k) is actually a constant independent of j and k because of the assumption of a uniform traffic matrix, we will refer to it for now as a function of j and k. This will

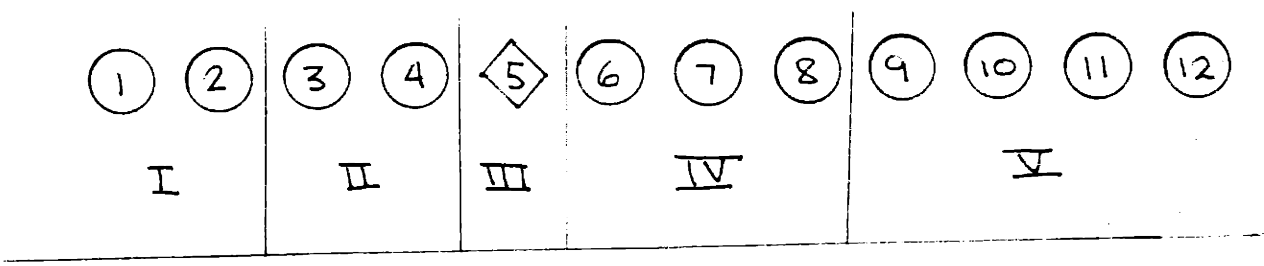

Fig. 2.4.2.1 Groups Contributing to Interference Factor at Radio k (k=5)

clarify our subsequent calculations.

As illustrated in Figure 2.4.2.1, the interference factor q(k) is the product of five terms corresponding to the following sets of radios:

(1) Radios i such that $1 \le i \le (k-1)/2$. These radios are to the

left of k and are closer to the left edge of the network than they are to k. Thus, any transmission from such a radio to its left is not heard by k. Radio k is interfered with only if i is transmitting to j such that $j \geq k$.

(2) Radios i such that $(k-1)/2<i<k-1$. These radios are to the left of k but k hears transmissions from such a radio to a radio j which satisfies either $i-j \geq k-i$ or $j \geq k$.

(3) Radio j itself. If it transmits to any other radio, it cannot receive a packet successfully.

(4) Radios to the right of j which are the images of those in group 2.

(5) Radios towards the right end of the network which are the images of those in group 1.

Hence,

$$q(k) = \prod_{i=1}^{(k-1)/2} \left[1 - \sum_{j=k}^{n} p(j)\right] \prod_{i=(k-1)/2+1}^{k-1} \left[1-\left(\sum_{j=1}^{2i-k} p(j)+\sum_{j=k}^{n} p(j)\right)\right]$$

$$\times \left[1 - \sum_{i=1}^{n} p(j)\right]$$

$$\times \prod_{i=k+1}^{(k+n+1)/2} \left[1-\left(\sum_{j=1}^{k} p(j)+\sum_{j=2i-k}^{n} p(j)\right)\right] \prod_{i=(k+1+n)/2+1}^{n} \left[1 - \sum_{j=1}^{k} p(j)\right]$$

(2.4.2.2)

Regrouping terms and taking the logarithm of both sides using the approximation $\ln(1-x)=-x$ for $x<<1$ yields

$$\ln q(k)=-[\sum_{j=1}^{k}(n-k/2-j/2)p(j) + \sum_{j=k+1}^{n}(k/2+j/2)p(j)] \quad (2.4.2.3)$$

Taking the logarithm of both sides of (2.4.2.1) and substituting (2.4.2.3), we have

$$\ln t(j,k)=\ln p(k)-[\sum_{j=1}^{k}(n-k/2-j/2)p(j) + \sum_{j=k+1}^{n}(k/2+j/2)p(j)] \quad (2.4.2.4)$$

Now the task at hand is to determine a set of p(k) which yields a constant t(j,k) and which maximizes this quantity. For n approaching infinity, this is impractical. Instead, we convert (2.4.2.4) into a non-linear, non-constant coefficient, differential equation and solve it numerically. We convert radio position from a discrete variable, k, to a continuous variable, x, and define f(x) as a "transmission probability density" such that $f(x)/n^2$ at x=k/n is p(k). We include the n^2 scaling factor so that the limits of integration which represent the summation of (2.4.2.4) are 0 and 1, regardless of n. Likewise, we define T(x) so that $T(x)/n^2$ at x=k/n is t(j,k). Thus, (2.4.2.4) becomes

$$\ln T(x)=\ln f(x) - [\int_{0}^{x}(1-x/2-x'/2)f(x')dx' + \int_{x}^{1}(x/2+x'/2)f(x')dx'] \quad (2.4.2.5)$$

Because T(x) is constant for all x, differentiating (2.4.5) with respect to x yields

$$f'/f - \left[\int_x^1 f(x')dx' - \int_0^x f(x')dx'\right]/2 + (2x-1)f=0 \qquad (2.4.2.6a)$$

Differentiating again and multiplying through by f yields

$$f''-(f')^2/f+(2x-1)ff'+3f=0 \qquad (2.4.2.6b)$$

To solve this numerically, boundary conditions for f' and f must be specified. These two degrees of freedom in (2.4.2.6b) correspond to the two desired properties for our solution f(x): that it yield a constant t and that it maximize this t.

We have specified the first boundary condition by noting that the network is symmetrical about x=n/2. Therefore, the two integrals of (2.4.2.6a) are equal and negate each other. Furthermore, the final term is 0. Thus, to satisfy (2.4.2.6a), we must have

$$f'(n/2)=0 \qquad (2.4.2.7)$$

This makes sense intuitively because x=n/2 is the point of highest congestion, meaning that more attempted transmissions to this point are required to obtain traffic equal to that nearer the extreme radios. Thus, f(x) should be maximized here, leading us again to (2.4.2.7).

Specifying (2.4.2.7) leads to a family of solutions to (2.4.2.6b) which achieve a constant t(x), which we rename as t. We have determined the second condition by trial and error;

running the program which solves (2.4.2.6b) with different values of f(n/2) until a maximum t was reached. Figure 2.4.2.2 shows f(x) for different values of f(n/2), along with the corresponding throughputs obtained, including the optimal solution in which

$$f(n/2)=2.7 \qquad (2.4.2.8a)$$

This translates to

$$p(n/2)=2.7/n^2 \qquad (2.4.2.8b)$$

For this value,

$$R \cong n^2 t=2.05/e=.754 \qquad (2.4.2.9)$$

Thus, maximum throughput is more than doubled for line networks in changing from the multihop routing of section 2.3 to one-hop routing.

Note that f(x) falls off to less than 1/2 its value at the center at the edges of the network, indicating that attempted transmissions are more than twice as frequent to radios in the center of the network as to those on the edges. The assumption made for ring networks -- that transmissions are sent equally frequently towards all radios in the network -- is not justified for line networks.

2.5 Interpretation of Results

The maximum throughputs which we have determined in this chapter for networks of various sizes under various routing

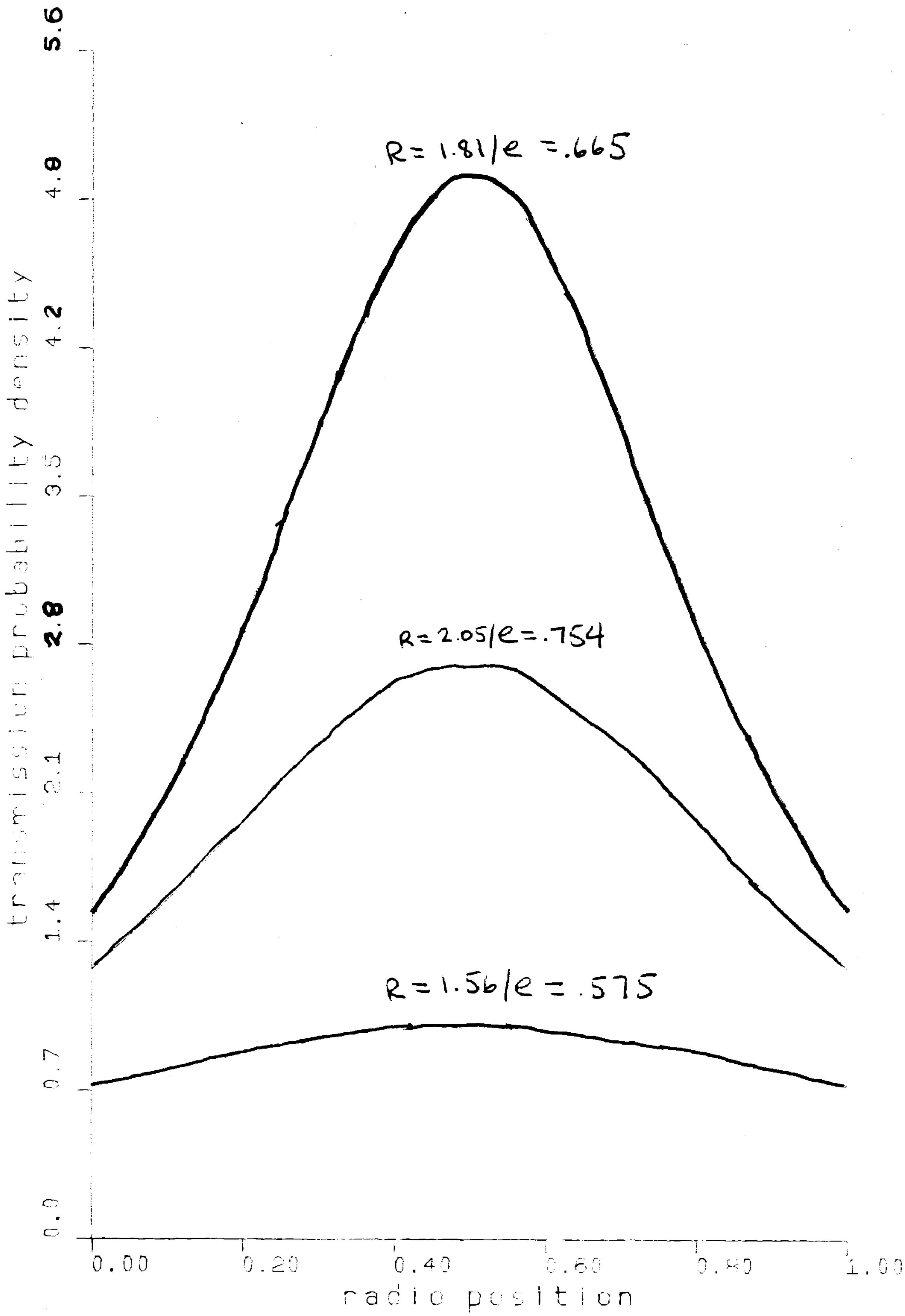

Fig. 2.4.2.2 Transmission Probability Density vs. Radio Position

strategies are summarized below in Table 2.5.1

	Number of nodes			
	2	3	4	
Multihop(N=1)	.500	.402	.370	.296
Multihop(1<<N<<n)	N/A	N/A	N/A	.368
One-hop	.500	.481	?	.754
Fully connected	.500	.444	.422	.368

Table 2.5.1

These results indicate that it is beneficial to endow transmitters in a line PRNET with adjustable and unlimited power, which are the properties assumed for one-hop routing. Note that if power is unadjustable and limited, a multihop routing strategy with constant transmission radius must be used, yielding throughputs less than 1/e for large networks. The throughput in such a network is limited because of the large number of transmissions typically required to deliver a packet from origin to destination.

On the other hand, if each transmitter has enough power to be heard by all radios in the network but cannot adjust this power, the network is fully connected and the throughput for a large network is 1/e. Here the throughput is limited because at most one successful transmission can occur per slot.

A network whose transmitters have both unlimited and adjustable power can complete each packet delivery in one hop while allowing multiple simultaneous transmissions when the transmissions are short, thus yielding a throughput more than twice as great as the other strategies in large networks and at least as great as the other strategies for the smaller networks we have investigated.

One ramification of this is that adjustability appears to be a beneficial property with or without unlimited power. For a network whose transmitters have a maximum transmission radius of N radios so that multihop routing is required, greater performance should be possible if this radius is decreased for transmissions shorter than N radios so as to allow a greater number of expected successful transmissions per slot.

In fact, we have not shown here that one-hop routing is optimal, only that it is better than the other schemes considered. A combination of adjustabilty and multi-hop routing for long transmissions may yield a throughput greater than .754 in large networks.

3. UNIFORM NETWORKS USING CONTENTION-FREE CHANNEL ACCESS

3.1 An Upper Bound on Throughput

Consider a cut between nodes i and i+1 in a packet radio network in which all nodes occupy a line and are numbered from 1 to n. Such a cut is shown in Figure 3.1.1. If a uniform traffic matrix is imposed, a necessary condition for completing transmissions between all source-destination pairs in the network is the transmission of a packet from each of the i nodes to the left of the cut to each of the n-i nodes to the right of the cut. Thus, i(n-i) packets must cross this cut, independent of how they are actually routed from source to destination. In particular, the left half of the network must send (n/2)(n/2) packets to the right half and vice-versa.

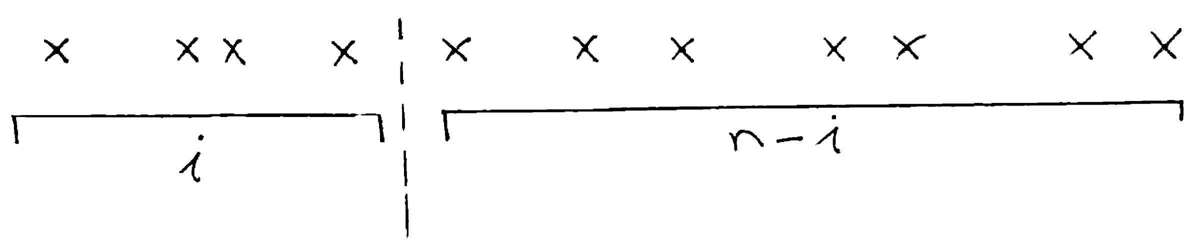

Fig. 3.1.1 A Cut in a Line Network

Therefore, it is necessary that $n^2/2$ packets cross this cut to complete all transmissions. The completion of all such transmissions will be termed a "cycle" for the remainder of this paper. We show here that, under the model we have imposed in the last chapter, only one packet can be successfully transmitted across this cut during a slot. Therefore, at least $n^2/2$ slots are required to complete a cycle.

In the lemmas that follow, we suspend the assumption that nodes are uniformly spaced and refer to radios by their co-ordinates in space, not their cardinal number within the network. The lemmas' results are illustrated in Figures 3.1.2a and 3.1.2b.

<u>Lemma</u> <u>3.1.1</u> If more than one radio on the same side of a cut attempts to transmit across this cut, only one transmission will be successfully received.

<u>Proof</u> Consider a cut made at a point x. Let nodes at t_1 and t_2 attempt to transmit to nodes at r_1 and r_2 respectively. We assume that $t_1, t_2 \leq x$ and $r_1, r_2 \geq x$. With no loss of generality, assume $t_2 > t_1$. Note that $t_2 \neq t_1$ because no radio can transmit simultaneously to more than one receiver. There are two cases to consider, as shown in Figures 3.1.2a:

(1) $r_1 \leq r_2$

In this case, $t_1 < t_2 < x \leq r_1 \leq r_2$. Thus, $t_2 < r_1 \leq r_2$. But our model assumes that r_1 thus hears the tranmission from t_2 to r_2 and cannot successfully receive the transmission intended for it. Thus, only one of the two transmissions is successfully received.

(2) $r_1 > r_2$

In this case, $t_1 < t_2 < x < r_2 < r_1$. Thus $t_1 < r_2 < r_1$ and only one

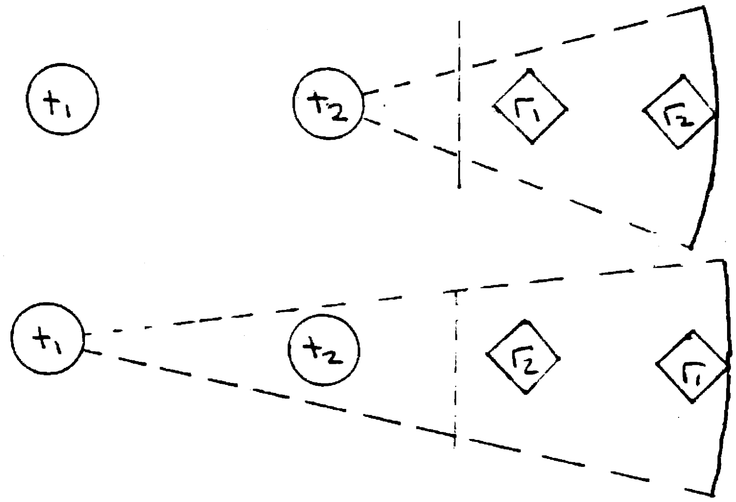

Fig. 3.1.2a Illustration of Lemma 3.1.1

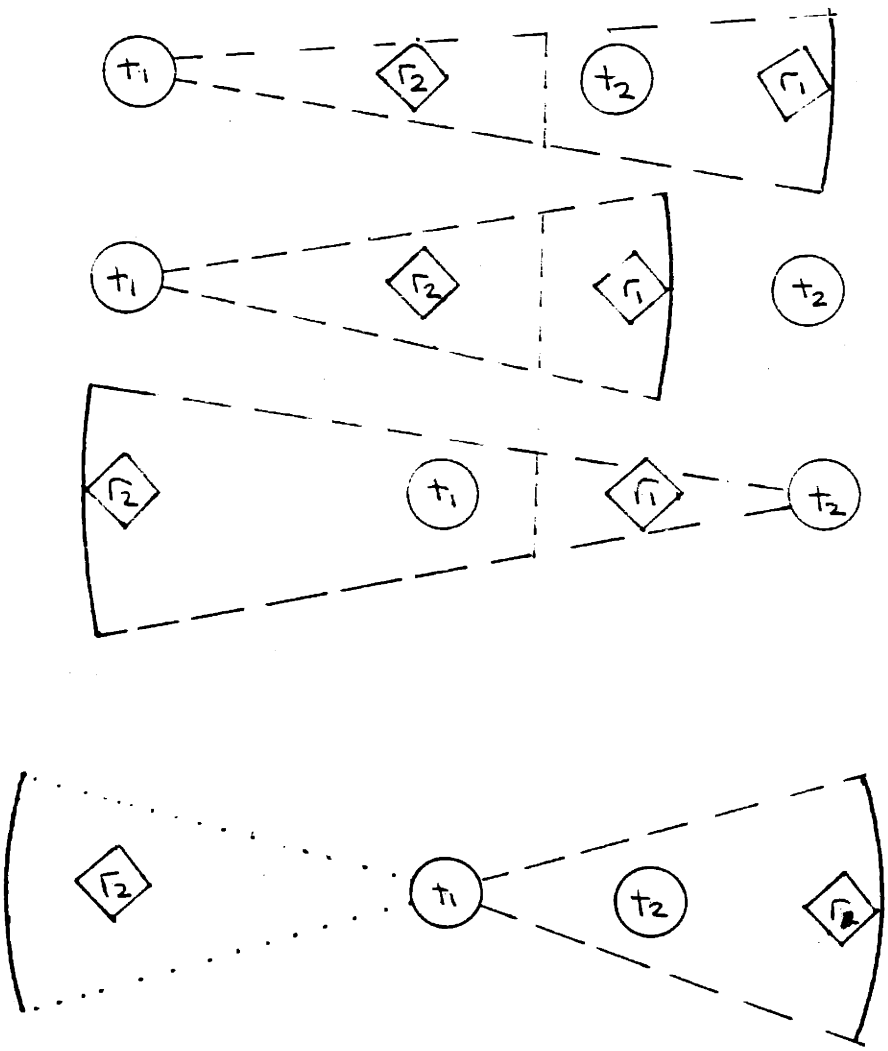

Fig. 3.1.2b Illustration of Lemma 3.1.2

transmission is successful.

Lemma 3.1.2 If radios on opposite sides of a cut attempt to simultaneously transmit across the cut, at most one transmission will be successful.

Proof Consider a cut made at x. Let node $t_1<x$ attempt to transmit to $r_1>x$ while $t_2>x$ attempts to simultaneously transmit to $r_2<x$. According to our model, the first of these transmissions is heard by all radios within the radius of transmission. These include all radios located at positions c such that

$$t_1-(r_1-t_1)\le c\le r_1,$$

Thus,

$$2t_1-r_1\le c\le r_1 \qquad (3.1.1)$$

A successful transmission from t_2 to r_2 requires that r_2 not satisfy (3.1.1). Now, under our assumptions, $r_2<x<r_1$. Thus, a successful transmission to r_2 from t_2 requires that

$$r_2<2t_1-r_1 \qquad (3.1.2)$$

Similar reasoning leads to the conclusion that a successful transmission from t_1 to r_1 is only possible if

$$r_1>2t_2-r_2 \qquad (3.1.3)$$

Manipulating (3.1.2) and (3.1.3) yields

$$t_2<(r_1+r_2)/2<t_1 \qquad (3.1.4)$$

However, we have stated that $t_1<x<t_2$. Thus, (3.1.4) cannot be satisfied and the two transmissions cannot be simultaneously successful. The combined results of lemmas 3.1.1 and 3.1.2 substantiate the contention made above: at most one packet can be successfully transmitted across any cut in the network, regardless of the location of transmitter and receiver.

We can use this fact to determine an upper bound on the throughput in any line network with uniform traffic requirements, independent of routing and regardless of whether or not the radios are spaced uniformly. The throughput is defined as the average number of packets delivered to their destinations per slot. There are n(n-1) packets which must be delivered to their destinations per cycle. Letting S define the number of slots required for a cycle, we have

$$R=n(n-1)/S\leq n(n-1)/(n^2/2)<2 \qquad (3.1.5)$$

In the next two sections, we investigate routing strategies analgous to those discussed in chapter 2, but which are based on contention-free channel access. We show that, for uniformly spaced radios, these yield throughputs approaching the bound of (3.1.5). In section 3.4, we extend the results of this chapter to regular grid networks.

3.2 One-Hop Routing

At the expense of long delays, transmissions in packet radio networks can be scheduled so as to prevent interference. In this case, radios do not transmit packets upon receiving them but must wait until they are scheduled to transmit. An extreme version of this is pure TDMA, where each slot is reserved for exactly one transmission. Clearly, this strategy yields a throughput of 1. However, it is suboptimal for packet radio networks because it fails to take advantage of the fact that more than one transmission can be successfully completed during a slot, providing that the transmissions do not interfere. We propose here a scheduling algorithm for line networks whose radios are uniformly spaced, which we refer to as Distance-based Time Division Multiple Access (DTDMA).

As shown in Figure 3.2.1a, the scheme is based on allocating each slot to transmissions of a given length, N. Radios are labelled modulo 2N+2 and all transmissions of length N on each cycle are accomplished in 2N+2 slots for $N<n/2$. To calculate the throughput using this scheme, we count the number of slots required per cycle.

As shown in Figure 3.2.1b, only one successful transmission can occur per slot allocated to transmissions where $N \geq n/2$. For a given N, one slot must be dedicated to each left-to-right transmission, in which the transmitters are numbered from 1 to n-N and the receivers from N+1 to n, and for each right-to-left transmission, in which the transmitters and receivers are reversed. Thus, 2(n-N) slots are required to complete transmissions for each such N.

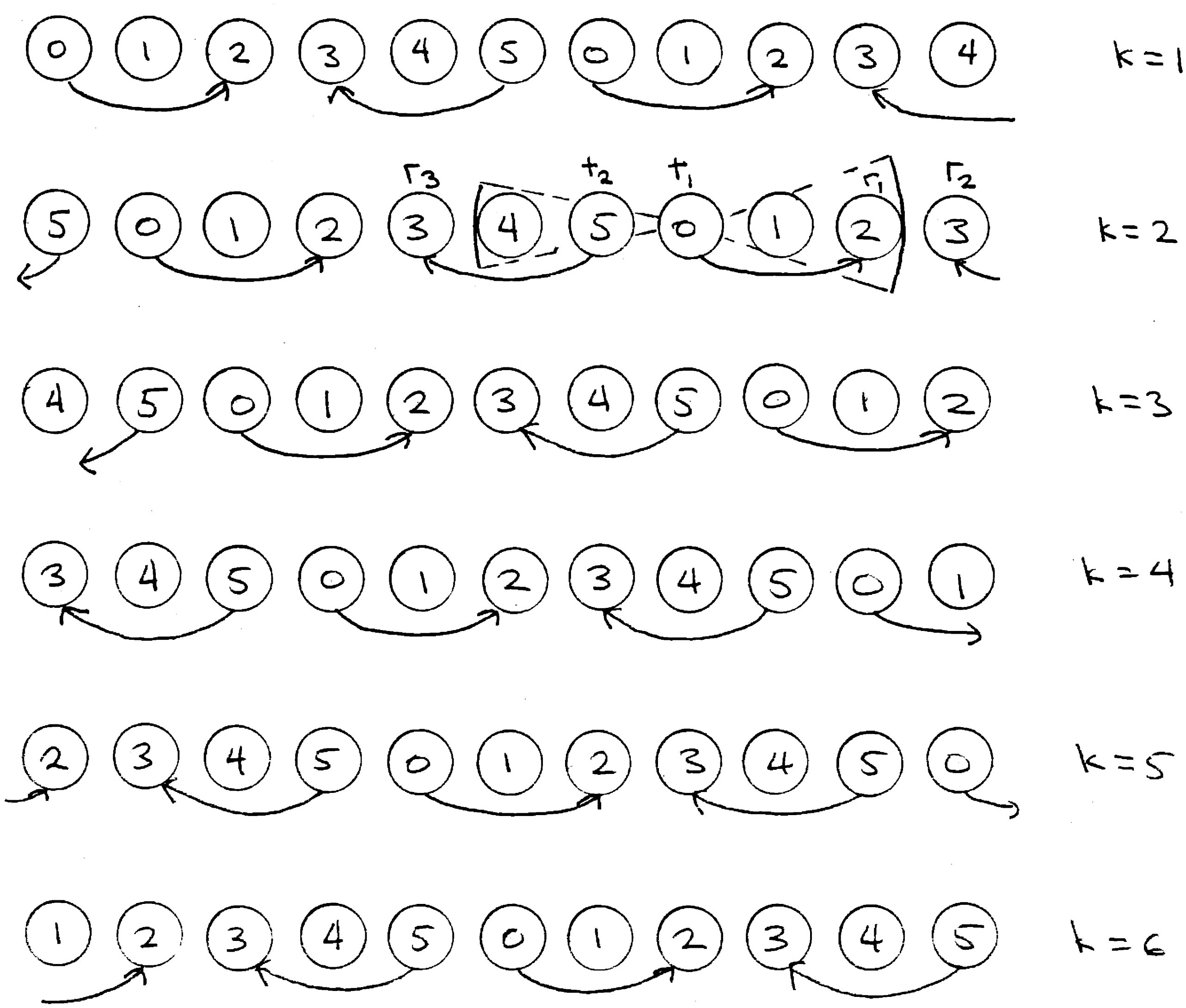

Fig. 3.2.1a   DTDMA with N=2

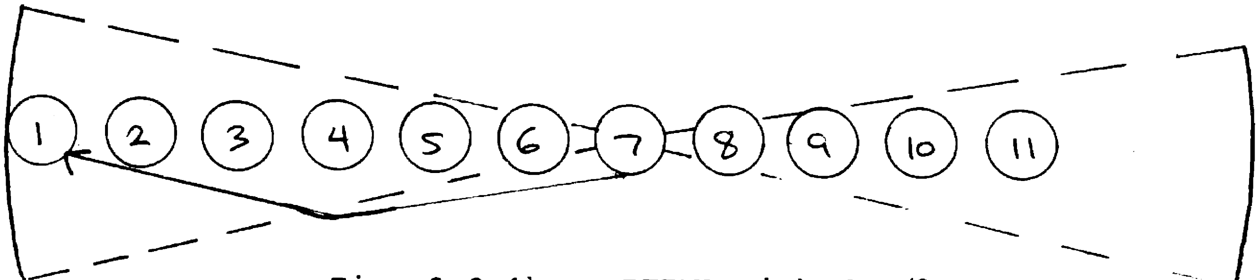

Fig. 3.2.1b   DTDMA with N>n/2

For $N<n/2$, we describe below the scheduling of all slots allocated to transmissions of a given length. In the following description, S is a counter of the slots used, k is an index determining the set of radios which transmit on a given slot, N is the distance in radios from transmitter to receiver, and j is the index of each radio labelled from left to right from 1 to n. I(j) is a function of j,k and N. It is used to label radios and identify those which transmit. The algorithm is:

```
         S=1, k=1
START:   for j=1 to n, I(j)=(j-k) mod (2N+2)
L_TO_R:  for all radios j with I(j)=0, transmit to radios j+N
R_TO_L:  for all radios j with I(j)=2N+1, transmit to radios
         j-N
         k=k+1
         if k>2N+2 stop
         S=S+1
         go to start
         end
```

We can show formally that, as illustrated in Figure 3.2.1a, the left-to-right transmissions in step L_TO_R and the right-to-left transmissions of step R_TO_L do not interfere with each other. Consider a radio j. With no loss of generality let $k=2N+2$. Relating j to I(j), we get

$$j=m(2N+2)+I(j), \text{ m integer} \qquad (3.2.1)$$

When a radio t_1 with $I(j)=0$ transmits from left-to-right, it is heard by radios c to the right of it for which

$$t_1 < c \leq t_1 + N,$$
$$1 < I(c) \leq N \qquad (3.2.2)$$

Thus, it does not interfere with the right-to-left transmission to radio r_2 with $I(r_2)=N+1$ and clearly does not interfere with transmissions whose receivers lie anywhere to the right of r_2.

The transmission from t_1 is also heard by all radios d to the left of it for which

$$t_1 - N \leq d < t_1 ,$$
$$(-N) \bmod (2N+2) \leq I(d) \leq 2N+1 \bmod (2N+2)$$
$$N+2 \leq I(d) \leq 2N+1 \qquad (3.2.3)$$

Thus, this transmission does not interfere with the right-to-left transmission from transmitter $t_2=t_1-1$ with $I(t_2)=2N+1$ to receiver $r_3=t_1-(N+1)$ with $I(r_3)=N+1$ and does not interfere with any transmissions to receivers which lie to the left of r_3. A similar argument can be made to show that the right-to-left transmission from t_2 does not interfere with any others.

To see that all transmissions of length N are completed in 2N+2 steps, note that each radio must transmit to two receivers, located N radios to the right and to the left. Consider any radio

$$j=m(2N+2)+p, \quad m, p \text{ integers} \tag{3.2.4}$$

Equation (3.2.4) can be satisfied for all j with $1 \le p \le 2N+2$. Now, referring to the algorithm, when k=p, radio j transmits from left to right and when $k = p \bmod (2N+2)+1 \le 2N+2$, j transmits from right to left. This is true for all radios in the network.

Applying the results derived above to a network with an even number of nodes (for convenience) yields

$$S = \sum_{N=1}^{n/2-1} (2N+2) + \sum_{N=n/2}^{n-1} (n-N) = n^2/2+n-2 \tag{3.2.5}$$

Thus, for large n, we can approach a throughput of 2 by using one-hop routing.

3.3 Multi-hop Routing

Another strategy for delivering packets is to use multihop routing, in which packets are sent N radios in the direction of the destination. This type of routing was analyzed for the slotted-Aloha channel access in Chapter 2. We study it here in conjunction with a scheduling algorithm similar to the one outlined in Section 3.2.

As we pointed out in Chapter 2, 1/(4N) of the network traffic is carried on each of the N center links in one direction. Thus, n(n-1)/(4N) packets must be sent in each direction over these links before all transmissions are finished. Using last section's scheduling algorithm with a fixed N, each link can carry a packet every 2(N+1) slots in a given direction.

Furthermore, as we showed in Chapter 2, a fraction o(N/n) of the slots must be dedicated to delivery transmissions. Because S is on the order of n^2, O(nN) slots must be added for these transmissions. Thus, for n large, the number of slots required per cycle using this algorithm, to a first-order approximation is given by the expression

$$S=2(N+1)(n^2/[4N])+O(nN)$$
$$=(n^2/2)(1+1/N)+O(nN) \qquad (3.3.1)$$

For moderate values of N (i.e., 1<<N<<n) a throughput approaching the upper bound of 2 can be attained with this strategy. Note that for N=n, this strategy is identical to pure TDMA, which yields a throughput of only 1, indicating that the second term of 3.3.1 cannot be ignored.

It is interesting that in DTDMA, multihop throughput is almost as great as one-hop throughput while for slotted-Aloha, one-hop throughput is almost twice as large. We have not come up with a good intuitive explanation for this phenomenon.

3.4 Regular Grid Networks

The scheduling algorithm of Section 3.2 can be extended for use in the regular grid network of Figures 3.4.1a,3.4.1b, and 3.4.1c. This network consists of n nodes placed in a square of n nodes on each edge. We label radios in this network by their x and y co-ordinates so that radio (x_1, y_1) refers to the radio of the x_1th column from the left and the

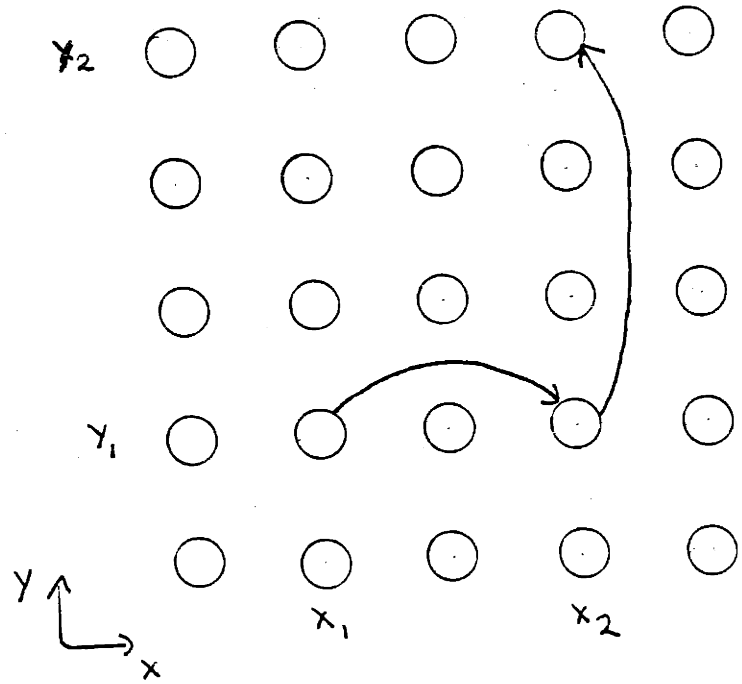

Fig. 3.4.1a Routing in Regular Planar Network

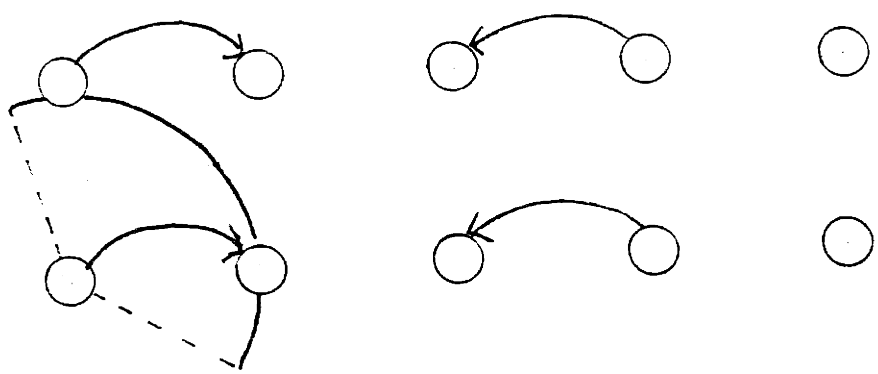

Fig. 3.4.1b Non-interfering Parallel Transmissions

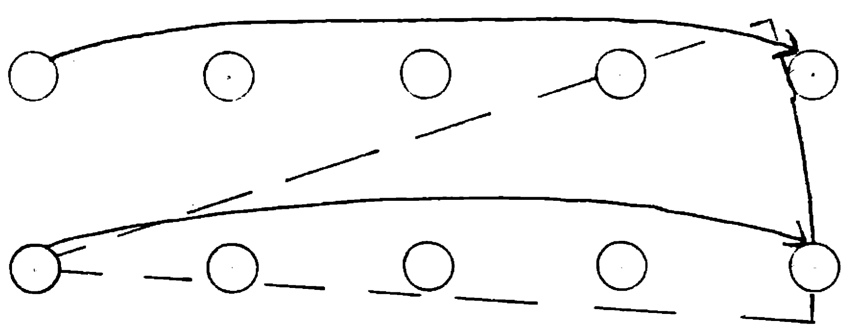

Fig. 3.4.1c Almost-interfering Parallel Transmissions

y_1th row from the bottom.

Our proposed scheme is to route packets from (x_1, y_1) to (x_2, y_2) through (x_2, y_1). Thus, packets are first routed along rows and then along columns. The index k in the algorithm is designated to be the same for each row so that when (x_1, y_1) transmits to (x_2, y_1) all radios in column x_1 transmit along their rows to radios in column x_2. To see that there are no collisions using this scheme, we must consider the effect of a transmission from (x_1, y_1) to (x_2, y_1) upon transmissions between columns x_1 and x_2 in other rows and upon other transmissions which occur during the slot. Figure 3.4.1b shows that there is no interference among these transmissions.

The set of radios potentially interfered with by such a transmission lie within a radius of $|x_2 - x_1|$ of (x_1, y_1). All radios (x_2, y_0) in other rows are a distance

$$\left[(x_2 - x_1)^2 + (y_1 - y_0)^2\right]^{.5} > |x_2 - x_1|$$

from (x_1, y_1) and all transmissions to such radios are successfully received.

Consider radio (x_3, y_1) which is also receiving a transmission during this slot. Our scheduling algorithm ensures that this radio is outside of the transmission radius about which we are concerned. By a similar argument to the one made above, all radios in column x_3 must be outside of this radius. Since the set of columns being transmitted to within each row is the same by design, these radios also are supposed to receive transmissions on the slot being considered here and these transmissions are successful.

Before calculating the throughput attainable using this strategy, we should point out that our model is not a good representation of reality in cases where there are long transmissions in each row (or column). To see this, let the distance between adjacent radios be 1 unit. A transmission from radio (x_1, y_1) to (x_1, y_1+N) thus has a radius of N. The distance from (x_1, y_1) to (x_1+1, y_1+N) is $\sqrt{N^2+1} \cong N(1+(1/2)/N^2)$ for $N \gg 1$. Thus, radio (x_1+1, y_1+N) is virtually within the radius of radio (x_1, y_1). This is shown in Figure 3.4.1c. However, strict application of our model leads one to believe that (x_1+1, y_1+N) successfully receives a transmission from (x_1+1, y_1) during this slot.

To avoid this limitation in the model, we will determine throughput only for multihop routing with N small. In this case, the assumption that adjacent rows' transmissions do not interfere corresponds more closely to physical reality.

Consider any radio in the network. During the period in which only transmissions within rows take place, it must send $\sqrt{n}$ packets to each radio in its row. We showed in Section 3.3 that it takes $(\sqrt{n}^2/2)(1+1/N)+O(N\sqrt{n})$ slots to complete one transmission for each source-destination pair in a row of $\sqrt{n}$ radios. Thus, letting S_R be the number of slots needed to complete all transmissions within rows,

$$S_R = (n\sqrt{n}/2)(1+1/N)+O(Nn) \qquad (3.4.1)$$

When this part of the algorithm has been completed, transmissions between rows must take place. Each radio now has received $\sqrt{n}$ packets from each of the $\sqrt{n}$ radios in its row and must now send $\sqrt{n}$ packets to each of the $\sqrt{n}$ radios in its column. (Actually, there are $\sqrt{n}$ packets at each radio which have already reached their destinations at this point and which do not have to be re-transmitted. However, for large n, this effect is negligible.) The second part of the algorithm is therefore identical to the first and also takes S_R slots. Thus, for this scheme,

$$S = (n\sqrt{n})(1+1/N)+O(Nn) \qquad (3.4.2a)$$

$$R=n(n-1)/S \cong \sqrt{n}/(1+1/N), \text{ if } N<<n \qquad (3.4.2b)$$

The analysis used in deriving this result, as well as the results from preceding sections regarding line networks, relies greatly on the regular structure of the network. However, the analysis yields insights into the more general case of randomly placed radios to be considered in later chapters and the results serve as benchmarks with which to evaluate those obtained in the analysis of such networks.

4. AN ALGORITHM FOR RANDOM LINE NETWORKS

4.1 Description of CDTDMA

We propose here a modified form of the one-hop DTDMA algorithm for use in networks whose nodes are randomly distributed along a line. In this new scheme, which we refer to as CDTDMA for Cellular Distance-based Time Division Multiple Access, the line is divided into a number of cells of equal length. Due to random placement, cells may be populated by 0,1,or more radios each. The algorithm is performed over a number of "rounds". On each round, slots are allotted to cell pairs according to a schedule similar to that used in DTDMA so that for each pair of cells, one radio in each cell communicates in both directions with one in the second cell. Additional rounds continue until each pair of radios has conducted communication in both directions. Radios in the same cell never communicate simultaneously except with each other. Figures 4.1.1a through 4.1.1d illustrate the algorithm schematically.

We have chosen this method for communicating among randomly placed radios after fruitlessly trying to devise a method based upon the radios' cardinal numbers in the network rather then their positions in space. CDTDMA is superior for two reasons:

(1) Dividing the network into cells of uniform length imposes a regular structure in the network similar to that present when nodes are spaced uniformly, allowing

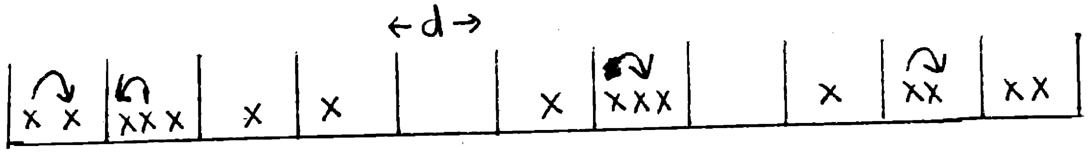

Fig. 4.1.1a CDTDMA: Round=1, N=0, k=1

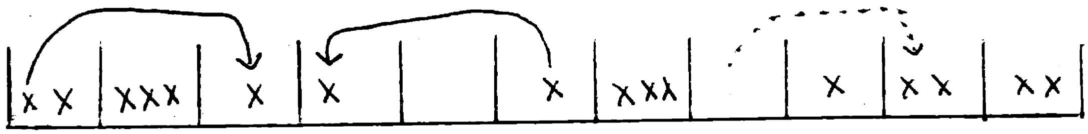

Fig. 4.1.1b CDTDMA: Round=1, N=2, k=1

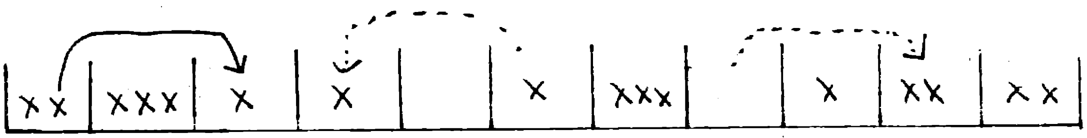

Fig. 4.1.1c CDTDMA: Round=2, N=2, k=1

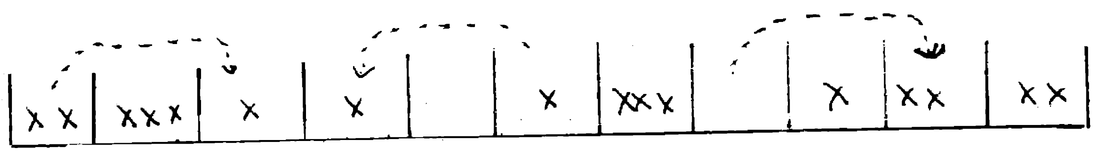

Fig. 4.1.1d CDTDMA: Round=3, N=2, k=1 (Unallotted)

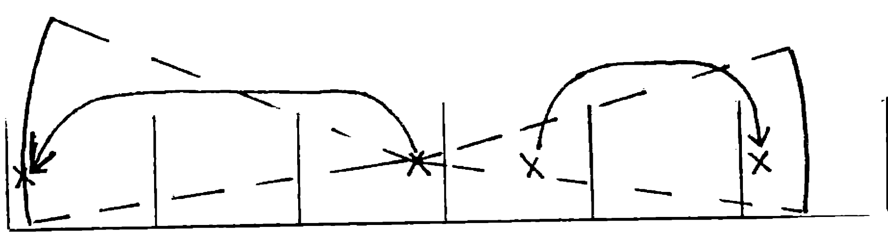

Fig. 4.1.2 Interference Between Transmissions from Adjacent Cells

for tractable analysis using similar methods to those of Chapter 3.

(2) From the point of view of implementation, this scheme is relatively simple in that radios need not have perfect information about exact positions. Instead, each must know positions to the precision of a cell width. Thus, in a mobile network, positions only need be updated by radios when they cross cell boundaries rather then continuously. Likewise, transmission radii need be controlled only to the precision of a cell width.

We describe the algorithm here using the formalism that was introduced in the DTDMA description of Chapter 3. In the following description,

c = cells in network,

S = slots used,

k = index determining set of transmitting cells,

j = cell index (from 1 to c)

N = distance between transmitting and receiving cells (in cells)

r = round number

M(r) = slots used on round r

Additionally, we define T(A,B) to be the statement: All radios in cell A have transmitted to all those in cell B and we define A->B to be the action of a particular radio in cell A transmitting to a particular radio in cell B for the first time. The algorithm is:

```
          r=1, S=0, N=0, k=1
START:    for j=1 to c, I(j)=(j-k)mod(2N+3)
TEST:     if T(j,j+N), all j s.t. I(j)=0
          and T(j,j-N), all j s.t. I(j)=2N+1
          then
              goto NEWSLOT
L_TO_R:   for all j s.t. I(j)=0,
             if not T(j,j+N), j->j+N
R_TO_L:   for all j s.t. I(j)=2N+1
             if not T(j,j-N), j->j-N
          S=S+1, M(r)=M(r)+1
NEWSLOT:  k=k+1
          if k>2N+3, then goto NEWDIST
          else
             goto START
NEWDIST:  k=1, N=N+1
          if N>c-1,  then goto NEWROUND
          else
             goto START
NEWROUND: if M(r)=0, STOP
          else r=r+1, N=0, M(r)=0, goto START
          END
```

Note that N=0 corresponds to intracell communication. To prevent interference among cells performing such communication, L_TO_R must involve a transmission from left to right within a cell while R_TO_L must involve transmission in the opposite direction.

There are several other points worth noting about CDTDMA. First of all, the algorithm must continue until all radios have communicated in both directions. Two-way communication between each pair of cells occurs on each round. Therefore, if some cell holds U_1 radios and another holds U_2 radios, it takes $U_1 U_2$ rounds until all radios in the two cells have communicated with each other. Likewise, if some cell contains U_1 radios, it takes $U_1(U_1-1)/2$ rounds for it to complete intracell communication. Thus, the number of rounds in the algorithm is given by the expression

$$\max_{U_1, U_2,\ U_1 \neq U_2}((U_1 U_2),(U_1(U_1-1)/2))$$

Additionally, CDTDMA possesses several subtleties that were absent in DTDMA. First of all, a "scheduled cell pair" must be distinguished from an actual transmission. This distinction is due to the "if" clauses of the lines labelled L_TO_R and R_TO_L. Each of these lines schedules a transmission between a pair of cells. However, the "if" clause in each of these lines specifies that an actual transmission only occurs if a scheduled cell pair results in a new transmitter-receiver pair. Examples of unused scheduled cell pairs appear in Figures 4.1.1b through 4.1.1d as dotted lines between the cells.

A second distinction is between "allotted" and "unallotted" slots. A slot is unallotted if it satisfies the condition in the line labelled "TEST", i.e, only if no scheduled cell pair results in a transmission during the slot. This is crucial for making the scheme efficient. If this test were not done, each round of the algorithm would use the same number of slots and the complete number of transmissions would be proportional to the number of rounds needed to connect the two most populous slots. Since only slots during which new transmitter-receiver pairs communicate are allotted, the number of slots used in each round must decrease since there will be an increasing number of unallotted slots each round. Moreover, a throughput of 1 packet per slot is ensured in this manner, guaranteeing that CDTDMA performs at least as well as pure TDMA. An example of an unallotted slot is illustrated in Figure 4.1.1d.

One minor difference between this scheme and DTDMA is that radios are labelled modulo 2N+3 rather then 2N+2. The reason for this is illustrated in Figure 4.1.2, which shows that adjacent cells which transmit simultaneously may interfere. Hence, when cell j with I(j)=0 transmits to the right, the cell with I(j)=2N+3 directly to the left does not. This implies that it requires 2N+3 slots to complete all scheduled cell pairs of length N for N<c/2.

4.2 Analysis of CDTDMA

4.2.1 Notation

The symbols to be used in this section for the analysis of CDTDMA are listed below:

λ = radio density (per unit length)

L = length of network

n = no. of radios in network

d = cell width

c = no. of cells in network

u = no. of radios in cell j

R = network throughput

S = no. of slots to complete cycle

P = algorithm performance measure

r = round number

s(d) = no. of slots to complete cell cycle as function of d

w(r) = Prob[slot unallotted on round r]

q(r) = expected fraction of scheduled cell pairs on round r which do not result in a transmission

N = length of transmission in cells

Y = no. of scheduled cell pairs on some slot

h(Y) = Prob[no actual transmissions occur | Y cell pairs scheduled]

G(Y) = no. of slots over cell cycle during which Y cell pairs are scheduled

[x] = the integer part of x

4.2.2 A Performance Measure

In our analysis of CDTDMA, we assume that the radios are placed independently on a line of length L. The density of radios is λ. Thus, n, the number of radios in the network is a random variable with distribution

$$p_n(n_0)=(\lambda L)^{n_0}\exp(-\lambda L)/n_0! \qquad (4.2.2.1)$$

Note that allowing L to be deterministic and n to be random instead of vice-versa allows the radio placement to be truly memoryless since there are no constraints on the exact number of radios in the network. We will be concerned here with first-order results for large networks satisfying the condition that L>>1. Due to the law of large numbers, n/E(n) stochastically converges to 1 and the decision to make n random has a vanishingly small effect on the results, and no effect on first-order results.

The network is divided into c cells of length d so that c=L/d. Because the placement of nodes is a memoryless renewal process, and because d is the same for each cell, the number of radios in any cell can be described by a random variable whose distribution is independent of cell boundary placement and of the number of radios in other cells. Hence we denote this random variable as an unsubscripted u and it is clear that

$$p_u(U)=(\lambda d)^{u}\exp(-\lambda d)/U! \qquad (4.2.2.2)$$

Before outlining the method used to analyze this algorithm's performance, we remind the reader that the throughput must satisfy $1\leq R<2$, where the lower bound is due to the constraint on only allotting slots in which at least one packet is delivered and the upper bound was derived in Chapter

3. We show in this chapter that for a very large network of randomly distributed radios, the upper bound can be approached.

As was the case in the last chapter, it is more natural to deal with the quantity S, the number of slots needed to complete a cycle, than with R. The relation between these variables, from equation (3.1.5), is

$$R=n(n-1)/S \tag{4.2.2.3}$$

The expected throughput, E(R), is at first glance, a reasonable-looking measure of performance. However, its computation would require an unwieldy derived distribution for R based on the random variables n and S. Instead, we will determine

$$P=E[n(n-1)]/E(S) \tag{4.2.2.4}$$

For large networks, the ratios n(n-1)/E(n(n-1)) and S/E(S) both stochastically converge to 1 so that

$$R/P=[n(n-1)/S]/[E(n(n-1))]/E(S)]$$

stochastically converges to 1. Thus, P is a good measure of performance in large networks.

To complete this analysis, we must:

(1) Determine the numerator of P, E(n(n-1)).

(2) Determine s(d), the number of slots needed to complete a "cell cycle". Over a cell cycle, one transmission

between each cell pair is scheduled. Thus, one cell cycle occurs per round. In determining s(d), we include both allotted slots, during which actual transmissions occur, and unallotted ones, during which no actual transmissions occur. It follows that s(d) is independent of the round number and is only a function of c, the number of cells, which is in turn a function of d, the cell width.

(3) Determine w(r), the expected fraction of slots in round r which are unallotted due to no transmission actually occurring during these slots. Hence the expression:

$$E(S) = s(d)\sum_{r=1}^{\infty}(1-w(r)) \qquad (4.2.2.5)$$

This expression states that the expected number of slots is the sum of the expected number on each round. The expected number on each round is the number required for a cell cycle (which is a constant) multiplied by the expected fraction of these slots actually allotted.

(4) Compute q(r), the probability that a cell pair scheduled on round r does not result in a transmission. This step is needed for the determination of w(r).

4.2.3 Determination of E(n(n-1))

Because n is a Poisson random variable with expectation L,

$$E[n(n-1)]=(E(n))^2+Var(n)-E(n)=(\lambda L)^2 \qquad (4.2.3.1)$$

4.2.4 Determination of s(d)

Assuming that there are c cells in the network, we can solve for s, the number of slots needed to complete a cell cycle. Using reasoning similar to that used in analyzing DTDMA,

$$s=\sum_{N=0}^{c/2-1}(2N+3)+2\left(\sum_{N=c/2}^{c-1}(c-N)\right)$$

$$=c^2/2+3c/2 \qquad (4.2.4.1)$$

Since c=L/d,

$$s(d)=(L/d)^2/2+3(L/d)/2 \qquad (4.2.4.2)$$

We will simplify subsequent calculations by assuming that c=O(n)>>1 and only considering first order effects. This is reasonable because for a large network we would want many cells so as to increase the maximum number of potential transmissions during a slot. As an extreme counterexample, suppose there were only two cells. Then there could only be one potential transmission per slot. Clearly, we would like to exceed this since the algorithm has been designed to attain at least this performance under any conditions. For a network with many cells, we use the first-order approximation

$$s(d)=(L/d)^2/2 \tag{4.2.4.3}$$

Consistent with this approximation, we will not concern ourselves from this point onwards with intracell transmissions. All such transmissions can be completed in $O(v^2)$ slots where v is the number of radios in the most populous slot. But since v^2 must be less than the sum of the squares of u^2, the number of radios in each slot, it is true that

$$\begin{aligned} E(v^2) &\leq E[\sum_{j=1}^{c} u_j^2]=cE(u^2)=c[(\lambda d)^2+\lambda d] \\ &=c(\lambda L/c)^2+c\lambda L/c=(\lambda L)^2/c+\lambda L \\ &<<(\lambda L)^2=E[n(n-1)],\ c>>1 \end{aligned}$$

Thus, the number of slots required for intracell transmissions is negligible compared to that required to complete a cycle and can be ignored.

Combining (4.2.2.4), (4.2.3.2) and (4.2.4.3) yields an expression for performance:

$$P=2(\lambda d)^2/\sum_{r=1}^{\infty} 1-w(r) \tag{4.2.4.4}$$

All that remains in the analysis is the determination of w(r).

4.2.5 Determination of w(r)

Because the number of radios in each cell is independent and identically distributed among all cells, q(a,b,r), the probability that a scheduled transmission between cells a and b does not result in a transmission, is independent of a and b and can be denoted q(r). Suppose there are Y scheduled cell pairs in some slot. Let h(Y) be the probability that none result in transmissions. Then

$$h(Y)=q(r)^{Y} \qquad (4.2.5.1)$$

because failures to establish new transmitter-receiver pairs are independent events. Thus, letting G(Y) be the number of slots during a cell cycle in which Y cell pairs are scheduled and Ymax be maximum number scheduled (which occurs for N=1, since these are slots with shortest transmissions.)

$$w(r)= \sum_{Y=1}^{Ymax} G(Y)h(Y)/s= \sum_{Y=1}^{Ymax} (G(Y)/s)q(r)^{Y} \qquad (4.2.5.2)$$

Determining G(Y) involves counting the number of slots in which Y cell pairs are scheduled for each transmission distance N. For instance, if $N \geq c/2$, (very long transmissions) Y=1 while for N=1, Y=O(c). In general, Y is inversely proportional to N.

To precisely specify the relationship between Y and N, we consider slots which are allocated to transmissions of length N for the case $N<c/2$. For convenience, we assume that c is even. As we showed in Chapter 3, there are 2(c-N) cell pairs scheduled between all pairs of cells located N cells apart per round. In CDTDMA, these transmissions require 2N+3 scheduled slots. Figure 4.2.5.1 illustrates that in some fraction f(N)

of these slots, an integer quantity, K(N), of these transmissions are scheduled and in the remaining slots, K(N)-1

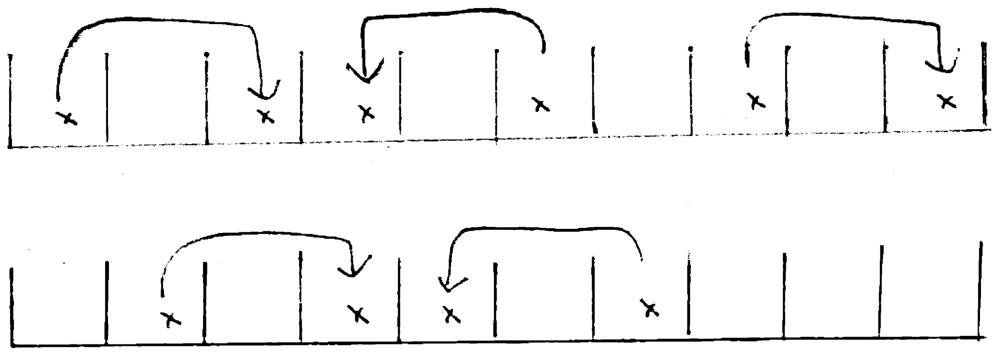

Fig. 4.2.5.1 Number of Scheduled Transmissions for Given Hop Length

are scheduled.

Thus,

$$2(c-N)=[f(N)(2N+3)]K(N)+[(1-f(N))(2N+3)](K(N)-1)$$
$$=(2N+3)(K(N)-1+f(N))$$

Therefore,

$$K(N)=[2(c-N)/(2N+3)]+1-f(N) \qquad (4.2.5.3)$$

We will eventually show that slots in which N>>1 dominate the quantity being computed, w(r). Thus, we approximate (4.2.5.3) as

$$K(N)=2(c-N)/2N+1-f(N)=c/N-f(N) \qquad (4.2.5.4a)$$

$$K(N)-1=c/N-f(N)-1 \qquad (4.2.5.4b)$$

The second of these equations is derived from a simple manipulation of the first and is included because it will be convenient to refer to in the following discussion.

The above equations are intuitively reasonable. It takes roughly 2N slots to complete 2(c-N) transmissions between cell pairs scheduled transmissions, assuming $N>>1$. Thus, the average number per slot is $2(c-N)/2N=c/N-1$. This average can be represented as the convex combination of two adjacent integers, which we have represented as K(N) and K(N)-1.

The slots during which Y cell pairs are scheduled can thus be separated into 2 components:

(1) Those corresponding to $Y=K(N)$. In this case, a fraction f(N) of slots allocated to transmissions of length N have Y cell pairs scheduled.

(2) Those corresponding to $Y=K(N)-1$. In this case, a fraction $1-f(N)$ of the slots have Y cell pairs scheduled.

To determine the first component, we re-arrange (4.2.5.4a) and set $Y=K(N)$ to obtain

$$N=c/(Y+f(N)) \qquad (4.2.5.5a)$$

We want to find the range of N for which a given Y occurs. Since f(N) ranges from 0 to 1, a fraction f(N) of slots are scheduled for Y cell pairs when

$$c/(Y+1)<N\leq c/Y$$

To compute the contribution from the second component, we re-arrange (4.2.5.4b) and set Y=K(N)-1 to obtain

$$N=c/(Y+f(N)+1) \qquad (4.2.5.5b)$$

so that a 1-f(N) of the slots are scheduled for Y cell pairs when

$$c/(Y+2)< N\leq c/(Y+1)$$

The above relations hold for 0<N<c/2. To complete the determination of G(Y), we must account for the case c/2<N<c-1. In this case, only one transmission per slot is possible. Let G'(1) be the number of slots potentially allotted for these. Then,

$$G'(1) = 2\sum_{N=c/2}^{c-1}(c-N) = c^2/4 \qquad (4.2.5.6)$$

since there are c-N pairs of cells separated by N cells and each pair must complete 2-way communication during a cell cycle. Since the slots for which c/3<N<c/2 are also included in G(1),

$$G(1)=G'(1) + \sum_{[c/3]+1}^{[c/2]} 2N(1-f(N)) \qquad (4.2.5.7)$$

Rearranging 4.2.5.4b yields

$$1-f(N)=2+Y-c/N \qquad (4.2.5.8)$$

For large N, the approximations

$$\sum_{N=[i]+1}^{[j]} 1=j-i,$$

$$\sum_{N=[i]+1}^{[j]} N=j^2/2-i^2/2$$

hold. Using these expressions and substituting (4.2.5.8) into (4.2.5.7) yields

$$G(1)=c^2/4+c^2/12=c^2/3 \qquad (4.2.5.9)$$

Similar reasoning and large amounts of algebra yield

$$G(Y)=2c^2/(y(y+1)(y+2)), \quad Y>1 \qquad (4.2.5.10)$$

Comparison of (4.2.5.9) and (4.2.5.10) shows that the latter expression holds for Y=1 aswell as for Y>1. Through partial fraction expansion of (4.2.4.10), it can be shown that, as expected,

$$\sum_{Y=1}^{\infty} G(Y)=c^2/2=s$$

Substituting (4.2.5.10) into (4.2.5.2) yields

$$w(r)=\sum_{Y=1}^{Ymax} 4/(Y(Y+1)(Y+2))q(r)^y \qquad (4.2.5.11)$$

$$=\sum_{Y=1}^{Y_{max}} 4q(r)^Y/[Y(Y+1)(Y+2)]$$

The summand in (4.2.5.11) varies with Y as q^Y/Y^3. Thus, the contribution to w(r) from slots on which many cell pairs are scheduled is negligible. This justifies the approximation we made in (4.2.5.3) and also allows us to let the summation's upper limit tend towards infinity.

We determine w(r) in terms of q(r) by expanding the summand into partial fractions and using the fact that

$$\sum_{N=1}^{\infty} x^N/N=\ln(1/(1-x))$$

After much algebra, this leads to the desired result

$$w(r)=2\ln(1/(1-q(r))(1-(1/q(r))^2)+3-2/q(r) \qquad (4.2.5.12)$$

Figure 4.2.5.2 shows a plot of w(r) vs. q(r).

4.2.6 Determination of q(r)

As we stated in our discussion of CDTDMA, U_1U_2 rounds are required to complete all transmissions between two cells, one of which contains U_1 radios and the other, which contains U_2. Thus, on the rth round, the only pairs of radios still to transmit to each other are those occupying cells whose radio populations satisfy the expression

$$U_1U_2 \geq r$$

This implies that the rth round is the final round in which actual transmissions occur between any cells in which $U_1U_2=r$. For example, r=6 corresponds to the final round in which any

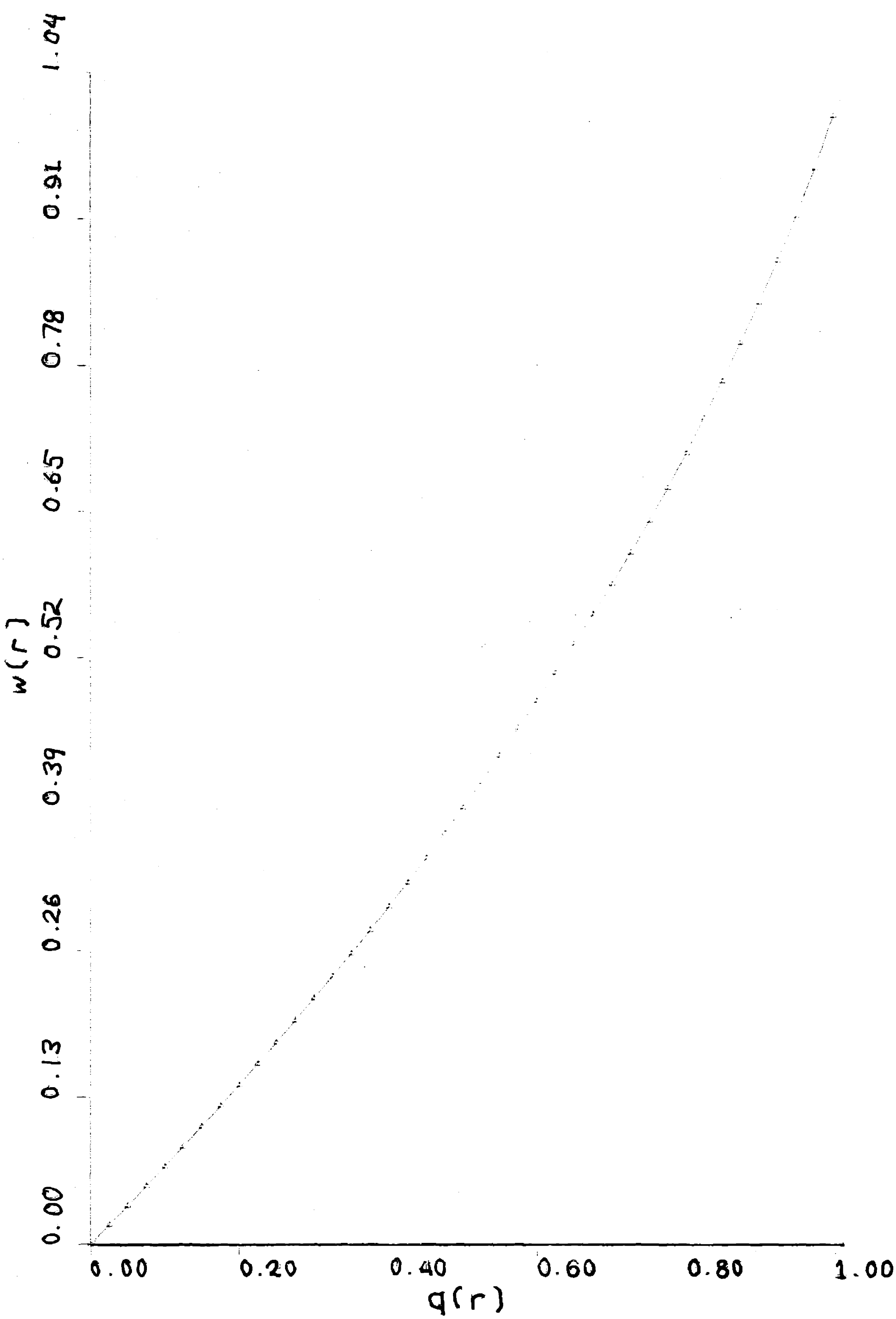

Fig. 4.2.5.2
Probability of Unallotted Slot vs. Expected
Fraction of Non-communicating Cell Pairs on Some Round

cell with 2 radios communicates to one with 3 or in which any cell with 1 radio communicates to one with 6.

Thus, if we define u as a random variable defining the population of some cell, q(r) is described by the recurrence relation:

$$q(1)=1-(1-p_u(0))^2$$

$$q(r)=q(r-1)+\sum_{U_1,U_2 \text{ s.t. } U_1U_2=r} p_u(U_1)p_u(U_2) \qquad (4.2.6.1)$$

We determine q(r) by substituting $p_u(U)$ from (4.2.3.2) into the above expression. Note that $p_u(U)$ is a function only of the parameter λd, and therefore q(r) and w(r) are also functions only of this parameter. A plot of q(r) as a function of r for three values of λd is shown in Figure 4.2.6.1. A plot of w(r) as a function of r for these cases is shown in Figure 4.2.6.2. Note that as λd grows large, q(r) and w(r) becomes smaller at each round since for large numbers of radios per cell, the probability that a scheduled cell pair does not result in a new pair of communicating radios is small and, therefore, only a small fraction of slots are unallotted.

4.2.7 Performance Results

We showed in the previous sections that w(r) is a function of λd, the expected number of radios in a cell. Thus, P is also a function of λd and can be renamed $P(\lambda d)$. However, because they are related in a complicated manner, we have obtained performance results numerically, solving (4.2.4.4) by

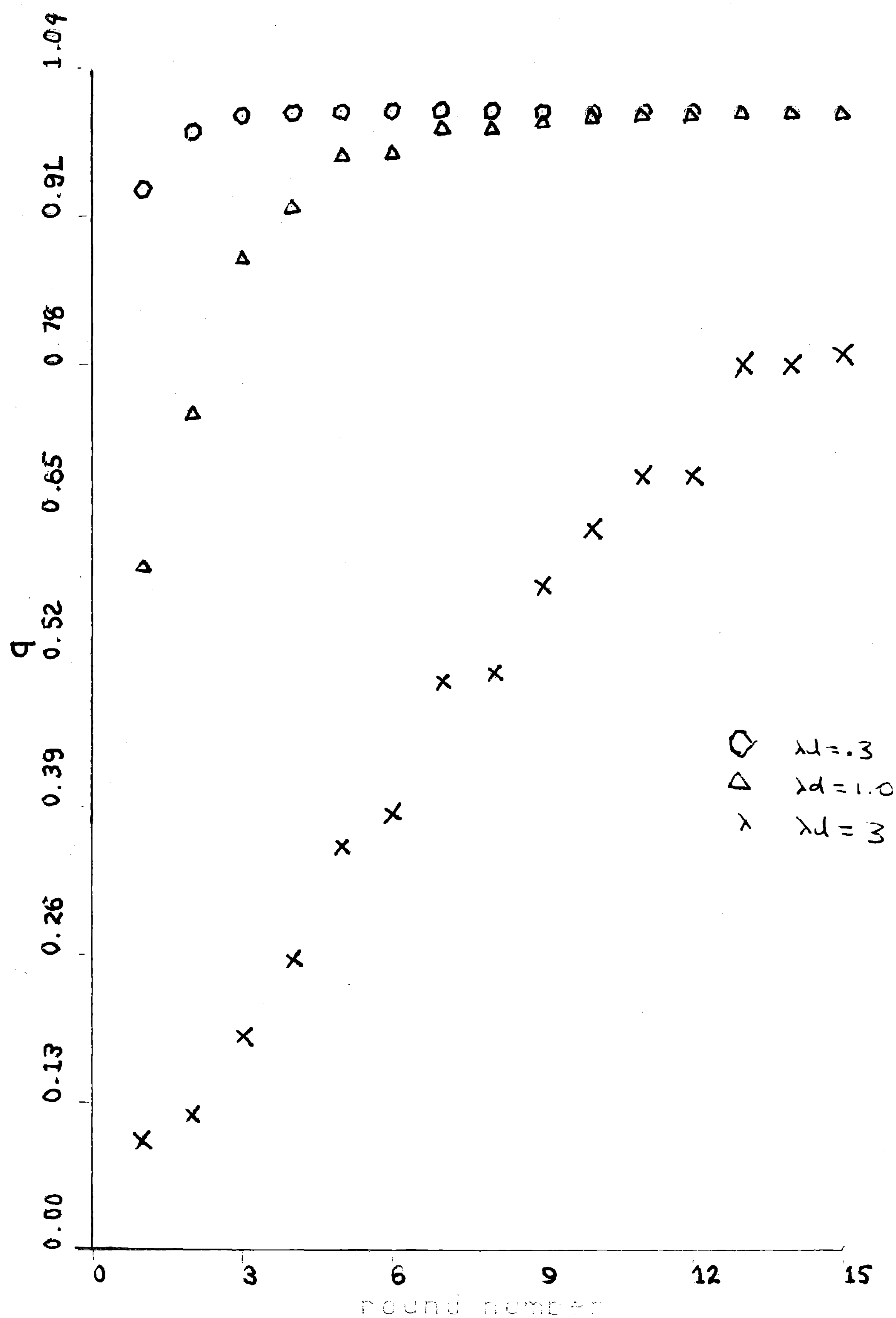

Fig. 4.2.6.1 Expected Fraction of Non-communicating Cell Pairs vs. Round

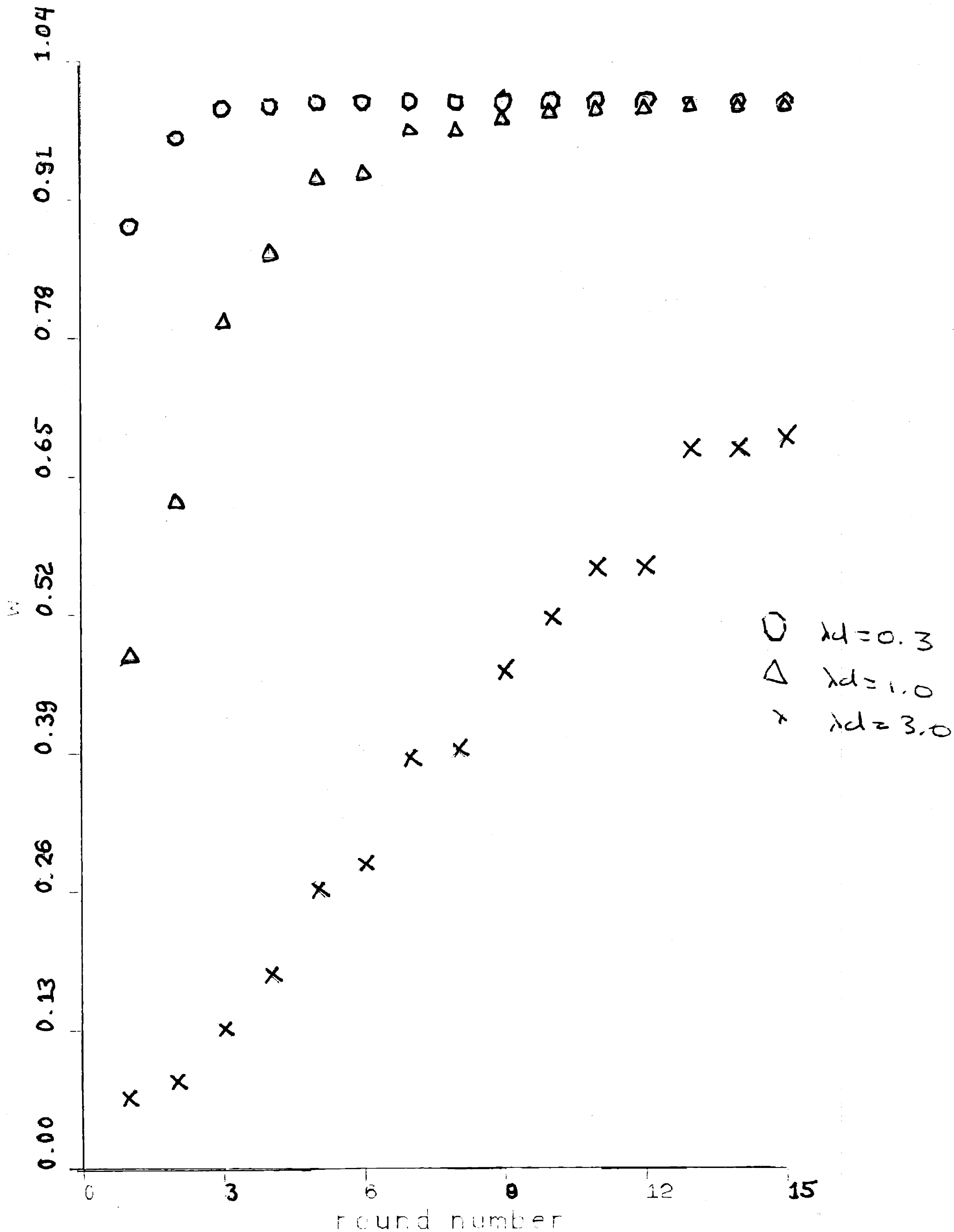

Fig. 4.2.6.2 Probability of Unallotted Slot vs. Round

summing enough terms in the denominator for for $P(\lambda d)$ to converge. A plot of P vs. log (λd) appears in Figure 4.2.7.1. We have extrapolated our numerical results for large values of λd because of difficulties encountered in our attempts to numerically determine them. In the next section, we validate the extrapolation by showing that $P(\lambda d)$ does approach 2 as λd gets large.

From this plot, it can be observed that, $P(\lambda d) \to 1$ as $\lambda d \to 0$. This is intuitively reasonable. In this case, the probability of a cell holding more than one radio approaches 0. Thus, virtually all transmissions occur on the first round. Furthermore, multiple transmissions are only scheduled for the same slot when transmitter and receiver are separated by the same number of cells for each transmission. As $\lambda d \to 0$, the cell width, d, becomes infinitesimal, assuming a constant λ. Thus, for one transmitter-receiver pair to be separated by the same number of cells as another set, each must be separated by virtually the same distance. Since the radio positions are continuous random variables, the probability of this event's occurrence approaches 0 as $\lambda d \to 0$.

Thus, for this extreme case, CDTDMA schedules one transmission per slot until all radios have communicated. This is identical to pure TDMA, which has a throughput of 1 packet per slot so this result makes sense.

The other extreme case, with $\lambda d \gg 1$, is far more interesting. The plot shows that an expected throughput approaching the upper bound of 2 packets per slot is

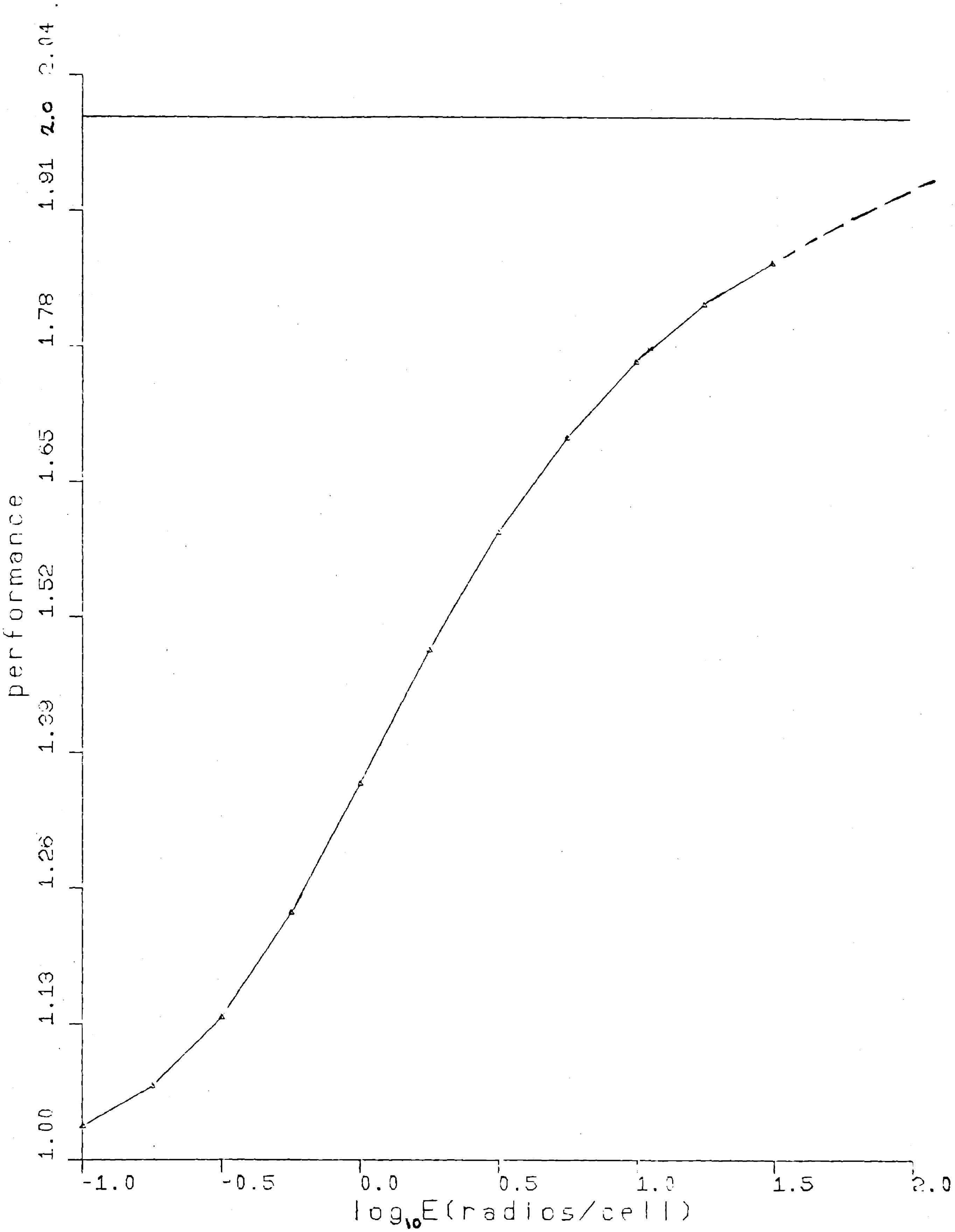

Fig. 4.2.7.1 CDTDMA Performance vs. Expected Cell Population

attainable when the number of radios in each cell is large. Thus, CDTDMA performs as well as any other algorithm for large, randomly spaced line networks. In the next section, we provide an interpretation of this key result.

4.3 Convergence of CDTDMA Throughput to Upper Bound

Suppose that CDTDMA is used in a regular line network. In this case, cells of equal widths hold equal numbers of radios. Let there be n radios in the network and let each cell hold u radios. Note that in this case, u is deterministic while in the random network, it is a random variable. This implies that there are n/u cells in the network. We assume that $1 \ll u \ll n$.

In the last section, we showed that a cell cycle for a network with c cells requires, to a first-order approximation, $c^2/2$ slots. Thus, the regular line network considered here requires $(n/u)^2/2$ slots to complete a cell cycle. During each cell cycle, each cell communicates once to each other cell. Thus, it takes u^2 cell cycles for each of the u radios in some cell to complete communication with each of the u radios in all other cells. So, for this case,

$$S=u^2(n/u)^2/2=n^2/2$$

to a first order approximation. Thus, CDTDMA used in a regular line network attains the upper bound on throughput derived in Chapter 3.

Now consider the random line network proposed in the previous section. The expected number of radios in each cell, $E(u)$, is λd. The radios can be divided into two groups, the first one encompassing all radios in cells with populations u such that $u \leq \lambda d+b$ and $\lambda d+b$ radios in each of the other cells. The second group consists of the "excess" radios in each cell which holds more than $\lambda d+b$ radios.

In $(\lambda d+b)^2$ rounds of CDTDMA, all radios in the first group can complete intercell communication with each other. Each round requires at most $c^2/2$ slots since each cell can communicate once to each other cell in a cell cycle, which requires $c^2/2$ slots. In fact, since some slots are unallotted each round, this is an upper bound. Thus, these radios require at most $(c^2/2)(\lambda d+b)^2$ slots to complete intercell communications with each other.

Using pure TDMA, the transmissions involving the excess radios can be completed at the rate of one per slot. In the worst case, each radio in this second group must communicate in both directions with each of the other radios in the network to complete a cycle. If there are $g(b)$ radios in the second group, the expected number of slots required is at most $2E(g(b)(n-1))$. The factor of 2 must be included because communication between each pair of radios must proceed in both directions. Thus,

$$E(S) \leq (c^2/2)(\lambda d+b)^2 + 2E(g(b)(n-1)) \qquad (4.3.1)$$

We show here that for large λd, $b \ll \lambda d$ can be chosen so that virtually all pairs of radios communicate in $(\lambda d+b)^2$ rounds.

Thus, we show that CDTDMA performance in this case converges to its performance in the regular line network discussed above in the sense that the expected number of rounds required to complete a cycle converges to $[E(u)]^2$ while in the regular network, the number of rounds needed was shown to be u^2. We use this to show that $P(\lambda d) \to 2$ as λd approaches infinity, as was claimed in the last section.

Our goal is to determine the right-hand term on the right side of (4.3.1) and show that for the proper choice of b, it vanishes as λd approaches infinity. First of all, we make the approximation

$$E[g(b)(n-1)]=E[g(b)]E[n-1]<E[g(b)](\lambda L) \qquad (4.3.2)$$

This is valid because for a large network, $n/E(n)$ stochastically converges to 1 and is virtually independent of $g(b)$.

To determine $g(b)$, we note that each cell holding $\lambda d+b+i$ radios contributes i radios to $g(b)$. Thus,

$$E[g(b)]=c\sum_{i=1}^{\infty} i\Pr[u=\lambda d+b+i]$$

$$=c\sum_{i=1}^{\infty} i(\lambda d)^{\lambda d+b+i}\exp(-\lambda d)/(\lambda d+b+i)! \qquad (4.3.3)$$

In general, $j!/(j+i)!<j^{-i}$. We use this fact and factor (4.3.2) to obtain

$$E[g(b)]<c(\lambda d)^{\lambda d+b}\exp(-\lambda d)/(\lambda d+b)!\sum_{i=1}^{\infty} i(1+b/\lambda d)^{-i}$$

$$=c(\lambda d)^{\lambda d+b}\exp(-\lambda d)(1+b/\lambda d)(\lambda d/b)^{2}/(\lambda d+b)! \quad (4.3.4)$$

Because $\lambda d+b>>1$, Stirling's approximation,

$$(\lambda d+b)! \cong (\lambda d+b)^{\lambda d+b+.5}\exp(-\lambda d+b)\sqrt{2\pi}$$

is accurate. Thus, after some algebra

$$E[g(b)]<c\ \exp(b)(1+b/\lambda d)^{\lambda d+b}\ (\lambda d/b^{2})\sqrt{\lambda d+b}/\sqrt{2\pi}$$

$$=c\ \exp[b-(\lambda d+b)\ln\ (1+b/\lambda d)](\lambda d/b^{2})\sqrt{\lambda d+b}/\sqrt{2\pi} \quad (4.3.5)$$

The expression $\ln(1+\lambda d/b)$ can be expanded so that the right side of (4.3.7) becomes

$$\exp[b-b+b^{2}/(2\lambda d)-b^{3}/(3(\lambda d)^{2})+b^{4}/(4(\lambda d)^{3})\ \ldots$$

$$-b^{2}/(\lambda d)\ +b^{3}/(2(\lambda d)^{2})-b^{4}/(3(\lambda d)^{3})\ \ldots\ ]$$

$$=\exp[b-b+b(-b/(2\lambda d)+[b/(\lambda d)]^{2}/6-[b/(\lambda d)]^{3}/12+\ \ldots)]$$

$$\leq\exp[-(b^{2}/\lambda d)/2+b^{3}/(\lambda d)^{2}/6]$$

$$\leq\exp[-b^{2}/(3\lambda d)], \qquad b\leq\lambda d \quad (4.3.6)$$

Since $c=L/d$, combination of (4.3.2),(4.3.5), and (4.3.6) yields

$$E[g(b)(n-1)]<(\lambda L)(L/d)(\lambda d/b^{2})\sqrt{(\lambda d+b)/(2\pi)}\exp[-b^{2}/(3\lambda d)]$$

$$<(\lambda L)^{2}(\lambda d)^{.5}\ b^{-2}\exp[-b^{2}/(3\lambda d)] \quad (4.3.7)$$

Note that we have simplified here by using the fact that $(\lambda d+b)/(2\pi)<\lambda d$ for $b<\lambda d$. If b is chosen so $b=(\lambda d)^{0.5(1+\epsilon)}$, $0<\epsilon<1$, this becomes

$$E[g(b)(n-1)]<(\lambda L)^{2}(\lambda d)^{-(0.5+\epsilon)}\exp(-(\lambda d)^{\epsilon}/3) \quad (4.3.8)$$

We claim that this quantity vanishes when the assumption that $1<<\lambda d<<\lambda L$ is satisfied. To show this, we take the logarithm of both sides of (4.3.8), using the fact that $(\lambda d)^{-(1.5+\epsilon)}<1$. Hence,

$$\ln E[g(b)(n-1)]<2\ln(\lambda L)-(\lambda d)^{\epsilon}/3 \quad (4.3.9)$$

Thus, if

$$(\lambda d)^{\epsilon}>6\ln(\lambda L)<<\lambda L,\ \lambda L>>1,$$

the expected number of slots required to complete the TDMA phase of CDTDMA vanishes as λd becomes large while remaining much smaller than λL.

To complete this discussion, we must show that the first term in (4.3.1), corresponding to the number of slots required to complete intercell communication among the "non-excess" radios, approaches $(\lambda L)^2/2$ as λd grows large. Thus, we show that

$$P(\lambda d)=E[n(n-1)]/E(S)\ ->(\lambda L)^2/[(\lambda L)^2/2]=2$$

which is the desired result.

The number of rounds required to complete this phase of CDTDMA is given in (4.3.1) to be $(\lambda d+b)^2$. Since

$b=(\lambda d)^{0.5(1+\epsilon)}$,

$$(\lambda d+b)^2=(\lambda d)^2+2(\lambda d)^{1.5+.5\epsilon}+(\lambda d)^{1+\epsilon} \qquad (4.3.9)$$

For $\lambda d>>1$ and $0<\epsilon<1$, only the first term on the right side of (4.3.9) is significant. Thus, approximating E(g(b)(n-1)) by 0, we have

$$E(S)\cong(c^2/2)(\lambda d)^2=(\lambda L)^2/2 \qquad (4.3.10)$$

and $P(\lambda d)\to 2$ as d approaches infinity, as we claimed in the last section.

It should be kept in mind that this result is based on the assumption that $1<<\lambda d<<\lambda L$. In a finite-sized network, this is an idealization since the number of radios per cell and the number of cells cannot both be made arbitrarily large. In the analysis of a small network, the first-order approximations that we have made to obtain our results would be inaccurate. Thus, determining CDTDMA performance as a function of the number of radios per cell for a small network would require the retention of the second-order terms that were dropped in our analysis.

4.4 CDTDMA in Planar Networks

Consider a random planar network which occupies a square of dimension L on which radios are placed with density λ per square unit. To analyze CDTDMA in such a network, we assume that each cell is a square of dimension d with d<<L. Let there be c cells in this network. Since the network area is L^2 and

each cell occupies an area of d^2, $c=(L/d)^2$. Figure 4.4.1 illustrates the network and the routing used. Packets are routed along rows of cells and then along columns, similar to the manner discussed in Chapter 3 for regular planar networks.

Ideally, a packet originating in cell (x_1,y_1) destined for cell (x_2,y_2) is routed through (x_2,y_1). However, this routing is impossible if cell (x_2,y_1) is empty, an occurrence not encountered in regular networks. We address this problem in the ensuing analysis by assuming that the expected number of radios per cell, $E(u)=\lambda d^2 >> 1$. Thus, the probability of an empty cell is $\exp(-\lambda d^2)<<1$ and we will show that empty cells have neglible effect on the results we obtain.

The transmissions along each row and column are scheduled as in CDTDMA for line networks except that a multihop version of CDTDMA is used, as shown in Figure 4.4.1. Thus, each packet is transmitted N cells per hop from its origin (x_1,y_1) to cell (x_2,y_1) and then in the same manner to cell (x_2,y_2).

We refer to the process of completing communication between all pairs of radios in the network which occupy the same row as the row phase of the algorithm. This process is followed by the column phase of the algorithm.

Furthermore, we define a round of CDTDMA to be the period over which one packet is delivered from each cell in the network to each other cell. In a multihop environment, a distinction must be made between an origin-destination cell pair and a transmitter-receiver cell pair. Over the course of a round, a transmitter may communicate many times with a

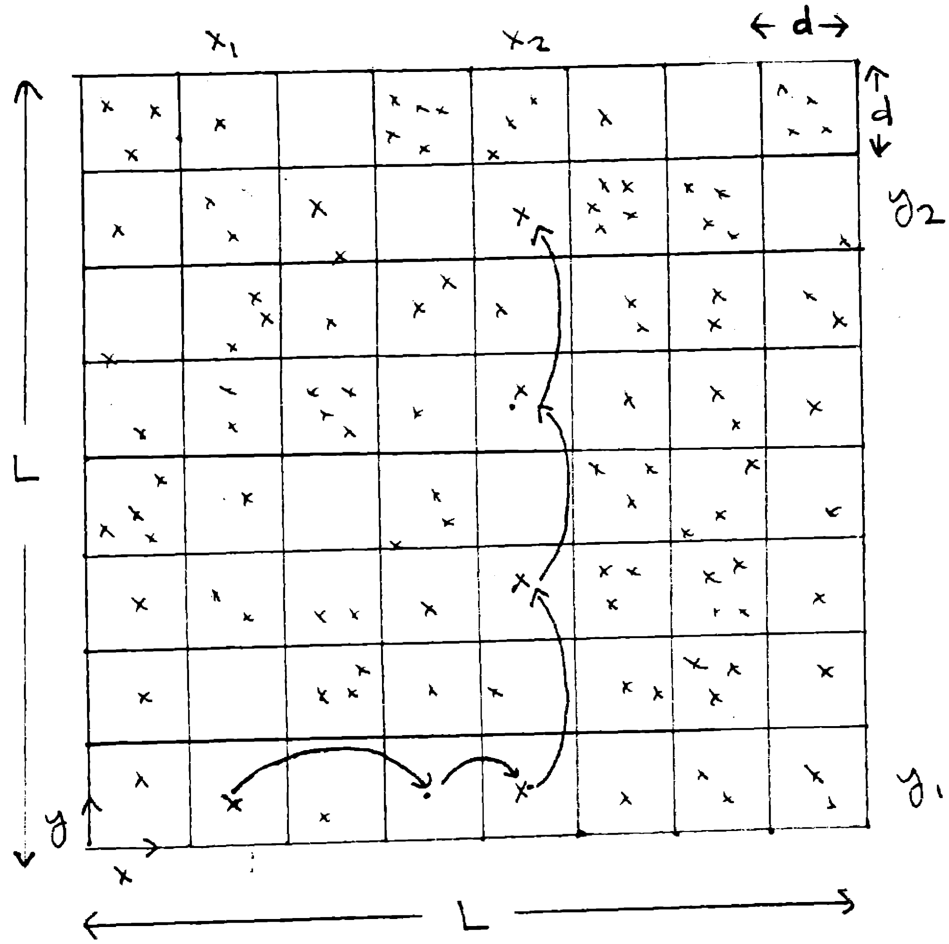

Fig. 4.4.1 Routing in Planar CDTDMA

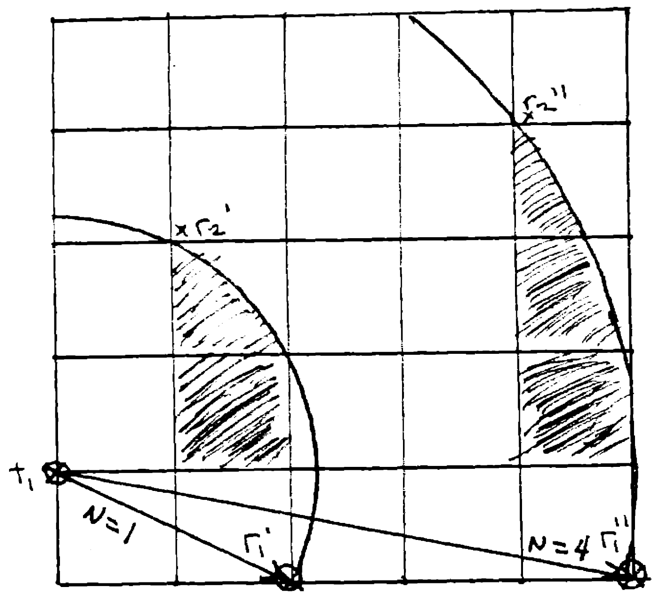

Fig. 4.4.2 Hop Length and Interference Between Rows

particular receiver since the transmitter and receiver must act as intermediate repeaters for many source-destination pairs of cells. This is untrue of one-hop routing, which makes no use of repeaters for communication within a row or within a column.

We use this multihop strategy to increase the number of non-interfering simultaneous transmissions among parallel rows (or columns). We shall now show that short hop lengths (transmission radii) are most efficient for CDTDMA in planar networks and will use this fact to justify the use of multihop routing.

Figure 4.4.2 illustrates the relation between hop length and interference. The radio in the bottom left-hand corner is shown attempting left-to-right transmissions of length N, for N=1 and N=4. The cross-hatched areas in this figure represent the areas in cells which hear these transmissions so that they can not simultaneously receive parallel transmissions. Note that for N=1, two rows are shown in which parallel transmissions cannot occur and for N=4, three such rows are shown.

To analyze this relationship, let the co-ordinates of one transmitter t_1 be $(d(x_1+v_1), d(y_1+w_1))$, with $0 \le v,w<1$. Thus, this radio is located in cell (x_1, y_1). Its intended receiver r_1 is located in cell (x_1+N, y_1). At the same time, a parallel transmission is attempted between transmitter t_2, located in cell (x_1, y_1+M), and r_2, located in cell (x_1+N, y_1+M), $M>0$. To guarantee that both transmissions are successful under the assumption of a well-defined transmission radius, M must be chosen so that r_2 is located outside t_1's radius and vice-versa. We only look at the first of these requirements here because symmetry implies that by choosing M to always satisfy the first condition, we can satisfy the second. Formally, the problem is to choose M so that $|t_1 - r_2| > |t_1 - r_1|$, where we use radio labels interchangeably with their co-ordinates.

For a given value of u_1, $|t_1 - r_2|$ is smallest when r_2 is located near the bottom left-hand corner of its cell while t_1 is located near the top edge of its cell ($w_1 \to 0$). Thus,

$$|t_1 - r_2|^2 > [(N - v_1)^2 + (M-1)^2]d^2 \qquad (4.4.1)$$

Note that there is a strict inequality because radios cannot be located exactly on an edge or corner of a cell. At the same time, $|t_1 - r_1|$ is largest for this position of t_1 when r_1 is located near the bottom right-hand corner of its cell so

$$|t_1 - r_1|^2 < [((N+1) - v_1)^2 + 1^2]d^2 \qquad (4.4.2)$$

Thus, to avoid interference,

$$\frac{|t_1 - r_2|^2}{d^2} > (N - v_1)^2 + (M-1)^2 \geq (N - v_1 + 1)^2 + 1 > \frac{|t_1 - r_1|^2}{d^2}$$

so

$$(M-1)^2 \geq 2(N - v_1) + 2, \quad (M-1)^2 \geq 2(N+1) \text{ since } v \geq 0$$

Thus,

$$M \geq \sqrt{2(N+1)}+1 \qquad (4.4.3)$$

For example, transmission between adjacent cells ($N=1$) requires that $M \geq 3$ so that at least three slots are required to complete a set of parallel transmissions in each row. Note that satisfying (4.4.3) only ensures that r_2 is marginally outside t_1's radius. In real networks, where transmission radius is a fuzzier concept, we would have to be more conservative. This phenomenon was discussed in Chapter 3 and was our rationale for using multi-hop DTDMA to avoid long transmissions in regular planar networks.

Equation (4.4.3) suggests an essential difference between the performance of DTDMA in regular planar networks and CDTDMA in random planar networks. In DTDMA, the regular geometric structure of the network allows all rows to simultaneously conduct parallel transmissions, if a distinct transmission radius is assumed. Thus, the number of slots required for all rows to complete the row phase of planar DTDMA is the same as the number required for a single row. We showed in Chapter 3 that this number is $(1+1/N)(n\sqrt{n})/2$.

The analgous quantity to consider in the analysis of CDTDMA performance is the number of slots required to complete one round of the row phase. Using reasoning similar to that employed in the analysis of DTDMA, it can be shown that if there are c cells in the network and, hence, $\sqrt{c}$ cells per row, a single row requires $(1+1/N)c\sqrt{c}/2$ slots to accomplish this. However, on any slot, radios in only 1 out of every M adjacent rows can transmit so as to avoid interference. Thus, the number of slots required for one round of the row phase is a factor M greater than that required in a single row, and is $M(1+1/N)c\sqrt{c}/2$. The same number is required for a round of the column phase so a round of planar TDMA requires $M(1+1/N)c\sqrt{c}$ slots, assuming that all slots are allotted. In other words, this number represents the length of a planar CDTDMA cell cycle.

The number of slots required for a complete cycle of CDTDMA is, at most, the product of the number required in a cell cycle and the number of rounds required. The second of these quantities is dependent only on the distribution of radios in each cell and not on the hop length, N. However, the choice of N determines the first quantity and by minimizing M(1+1/N), the number of slots can be minimized given some distribution of cell populations. From (4.4.3), we see that the quantity to be minimized is

$$M(1+1/N) \geq ([\sqrt{2(N+1)}+1]+1)(1+1/N), \text{ N integer}$$

where we are using [x] to represent the integer part of x. This is necessary because M must be an integer at least as great as $\sqrt{2(N+1)}+1$. Since this is a simple function of N and the minimization appears to require a small N, we do this by trial and error. Table 4.4.1 shows M(1+1/N) as a function of increasing values of N.

Dependence of Cell Cycle on Hop Length

N	M	M(1+1/N)
1	3	6
2	4	6
3	4	5.33
4	5	6.25
5	5	6
6	5	5.83
7	6	6.86

Table 4.1.1

Since

$$M(1+1/N)>M>6, \qquad N>7,$$

N=3 is the optimal hop length in cells, although there is not much sensitivity to the choice for N<7.

Were one-hop routing being used, N would not be constant and the analysis would be more difficult. However, it is true that a typical hop length in one-hop routing is on the order of the number of cells in a row, L/d. For L>>d, a typical hop length is much greater than 3 cells. It is these long transmissions which makes one-hop CDTDMA among rows and columns inefficient. This justifies the strategy of multi-hop routing among rows and columns.

The use of short transmission radii is intuitively reasonable because in planar networks, the number of radios interfered with increases as the square of the transmission radius while the number of hops per typical origin-to-destination packet delivery only increases linearly. Therefore, small radii are desirable in this case, unlike the situation in a line network.

Our goal is to make a rough analysis of the behavior of CDTDMA in planar networks under the model proposed above. Our reasoning is similar to that used in section 4.3 so we will provide a minimum amount of explanation regarding concepts covered there. In our analysis, we initially assume that there are no empty cells and then show that the effect of these cells is negligible, even under worst-case assumptions about how the algorithm handles the occurrence of empty cells.

We have already shown that $M(1+1/N)c\sqrt{c}$ slots are required for a cell cycle. If all slots are allotted, it requires this many slots to complete each round. Thus, to determine the number of slots required for a cycle, the number of required rounds must be determined.

To compute the number of rounds required to complete a cycle, we divide the network into two groups, the first consisting of all radios in cells with λd^2+b or less per cell and the second consisting of the "excess" radios. In $(\lambda d^2+b)^2$ rounds, all communication among the first group can be completed. An upper bound on the number of slots required per round is the number required in a cell cycle. Thus, S_1, the number of slots required to complete communication among all radios in the first group is bounded by:

$$S_1 \leq M(1+1/N)c\sqrt{c}(\lambda d^2+b)^2$$

For b chosen small enough, substitution of $c=(L/d)^2$ yields

$$S_1 \le M(1+1/N)(L/d)^2 (L/d)(\lambda d^2)^2 (1+b/(\lambda d^2))^2$$
$$= M(1+1/N)\lambda L^2 \sqrt{\lambda L^2} \sqrt{\lambda d^2} (1+b/(\lambda d^2))^2$$
$$= M(1+1/N)E(n) \sqrt{E(n)} \sqrt{E(u)} (1+b/(\lambda d^2))^2 \qquad (4.4.4)$$

We show below that for a proper choice of b, this is the dominant term in S, the number of slots required to complete a cycle.

By transmitting directly between origin and destination, each transmission involving the excess radios can be accomplished in one slot. Thus, as in the last section, S_2, the number of slots required here under this strategy is given by

$$S_2 \cong E[g(b)]E[n-1] \cong \lambda L^2 E[g(b)] \qquad (4.4.5)$$

To determine E[g(b)], we use the results of section 4.3 but modify them to account for the fact that the network is planar rather then linear. Equations (4.3.3)-(4.3.5) show that E[g(b)] can be expressed as the number of cells multiplied by a function of λd and b. Here, we replace λd by λd^2 because this is the expected number of radios per cell in the planar case. Making this substitution and applying the simplifications used in (4.3.6) and (4.3.7), we get

$$E[g(b)] < c(\lambda d^2/b^2)\sqrt{\lambda d} \exp[-b^2/(3\lambda d^2)] \qquad (4.4.6)$$

Letting $b=(\lambda d^2)^{.5(1+\epsilon)}$, $0<\epsilon<1$ and substituting $c=(L/d)^2$, combination of (4.4.5) and (4.4.6) yields

$$S_2 < (\lambda L^2)^2 (\lambda d^2)^{-(.5+\epsilon)} \exp(-(\lambda d^2)^{\epsilon}/3) \qquad (4.4.7)$$

As we showed in section 4.3, this quantity can be made arbitrarily small providing that $O[(\lambda d^2)^{\epsilon}] > O(\ln \lambda L^2)$. Letting $\epsilon \to 1$, this condition can be simplified to

$$O(\lambda d^2) > O(\ln \lambda L^2) \qquad (4.4.8)$$

Recall that as long as this condition is barely met, S_2 will vanish as λL^2 approaches infinity. This is especially true because the bound (4.4.7) is weak due to the series of approximations and worst-case assumptions made in section 4.3.

Our final task is to calculate S_3, the expected number of slots required to deliver packets to destinations which cannot be reached through normal CDTDMA routing because of the presence of empty cells. We assume that all packets scheduled to be routed through empty cells are delivered using TDMA, one packet per slot. The number of packets forwarded by any cell cannot exceed the sum of the number of packets scheduled to be routed through the cell's row and through its column. Since each radio in the cell's row must communicate in both directions with every other radio in the network, the expected number of packets scheduled to be routed through some row is double the product of the number of cells per row, the expected number of radios per cell, and the expected number of radios in the network, under the assumption that the last two random variables are nearly independent. This quantity is about $2(\lambda L^2)(\lambda d^2)(L/d)$. The expected number of packets scheduled to be routed through some column is the same, by symmetry. Thus,

$$S_3 \le 4(\lambda L^2)(\lambda d^2)(L/d)c \; \Pr(\text{cell is empty})$$
$$= 4\lambda L^2 \lambda d^2 (L/d)^3 \exp(-\lambda d^2) \qquad (4.4.9)$$

This expression is a conservative upper bound on S_3 because it assumes that if there are multiple empty cells in a row, each packet routed through the row must be transmitted to each of its destinations once per empty cell, even though only one such transmission would in fact be necessary. An inspection of (4.4.9) reveals that S_3 can be made to vanish as L approaches infinity if condition (4.4.8) is met. Thus, it too can be neglected.

Approximating S_2 and S_3 by 0 and approximating the factor $(1+b/(\lambda d^2))^2$ by 1 since b can be chosen so that $b<<(\lambda d^2)$, we have

$$P(\lambda d^2)=E[n(n-1)]/E(S)=(\lambda L^2)^2/[M(1+1/N)(\lambda L^2)^{1.5}(\lambda d^2)^{.5}],$$
$$O(\ln \lambda L^2)<O(\lambda d^2)<<O(\lambda L^2) \qquad (4.4.10)$$

We could simplify this to

$$P(\lambda d^2)=(L/d)/[M(1+1/N)]$$

implying that throughput can be made arbitrarily large independent of network size by using lots of cells. However, this interpretation ignores the condition that $O(\lambda d^2)>O(\ln \lambda L^2)$ which we have determined to be sufficient and which is probably necessary to ensure that S_2 and S_3 can be ignored.

If we let $O(\lambda d^2)$ barely exceed $O(\ln \lambda L^2)$, we can interpret (4.4.10) as

$$P(\lambda d^2)=O[\lambda L^2/\ln \lambda L^2]^{.5} /[M(1+1/N)]$$
$$=O[E(n)/\ln E(n)]^{.5} /[M(1+1/N)] \qquad (4.4.11)$$

Note that the expected throughput is on the order of $\ln \sqrt{E(n)}$ worse than the throughput of $O(\sqrt{n})$ obtained using DTDMA in a regular planar network. As can be seen in (4.4.10), the culprit here is the factor $\lambda d^2 = E(u)$. This suggests that better results may be possible if cells are made smaller, so that, typically, a smaller number of radios occupies each cell. In this case the analysis of CDTDMA would have to be totally different from that performed in this section since condition (4.4.8) would not be met, meaning that a significant number of slots would be required for communication involving excess radios, assuming that TDMA is used for such communication. Furthermore, the number of empty cells in the network would be significant. For example, if we let $\lambda d^2=1$, the probability of a cell being empty is 1/e. Thus, it would be folly to use TDMA to route packets through rows and columns containing empty cells.

Optimal use of CDTDMA in such a network would involve clever routing so as to avoid empty cells and to compensate for the relatively large variance in the number of radios contained in each cell. To avoid empty cells, adjustable transmission power (which is unneeded for multihop routing with constant N) could be used to "leapfrog" empty cells by increasing transmission radius so as to reach a cell which contains at least one radio. Likewise, the effect of variance in cell population can be reduced by routing packets around cells containing large numbers of radios and through cells which have few radios. This would have the effect of equalizing the number of rounds required to complete communication between different pairs of cells, since a cell with few radios would have proportionately more transmissions per radio during a cycle than a cell with many radios. We believe that this would be beneficial.

This analysis of CDTDMA would also require the determination of the optimal cell size given a routing algorithm. Making each cell as small as possible so that $\lambda d^2 \to 0$ is not a good idea because the effect is the same as making $\lambda d \to 0$ in a line network. We discussed in section 4.3 that in this case, CDTDMA converges to pure TDMA because the probability that more than one transmission could occur simultaneously would approach 0. Since pure TDMA, which yields a throughput of 1, is clearly non-optimal, the optimization of cell size in a planar network is a non-trivial problem.

We believe that good routing algorithms along with optimal cell size would enable CDTDMA to attain a throughput proportional to the square root of the expected number of radios in the network. It is beyond the scope of this work to make a more thorough analysis of CDTDMA in random planar networks, and we believe that such an analysis would constitute a most logical extension of our work. In the next chapter, we suggest further areas for research related to our work.

5. SUGGESTIONS FOR FURTHER RESEARCH

There are several subjects that have been introduced in our work which we believe should be investigated more thoroughly. One such area of investigation, the problem of optimizing planar CDTDMA, was introduced at the end of Chapter 4. We now present several other possible extensions of our work and some suggestions on how to go about looking at them.

5.1 An Upper Bound on Throughput in Planar Networks

In order to have a benchmark against which to measure different routing strategies in planar networks, it would be desirable to determine an upper bound on network throughput analgous to the one determined for a line network in Chapter 3. One way to obtain such a bound is illustrated in Figure 5.1.1. A cut is made in the network so that n_1 radios are on one side and $n-n_1$ are on the other. To complete a cycle, $n_1(n-n_1)$ packets must cross this cut in either direction. If the maximum number of packets which can cross this cut per slot in both directions is B, then under any routing strategy, it must be true that

$$S \geq 2Bn_1(n-n_1) \qquad (5.1.1)$$

This is just a generalization of the procedure we used for line networks in which we showed that B=1 for any cut and chose $n_1=n/2$ to maximize S.

However, in the case of a planar network, this problem is

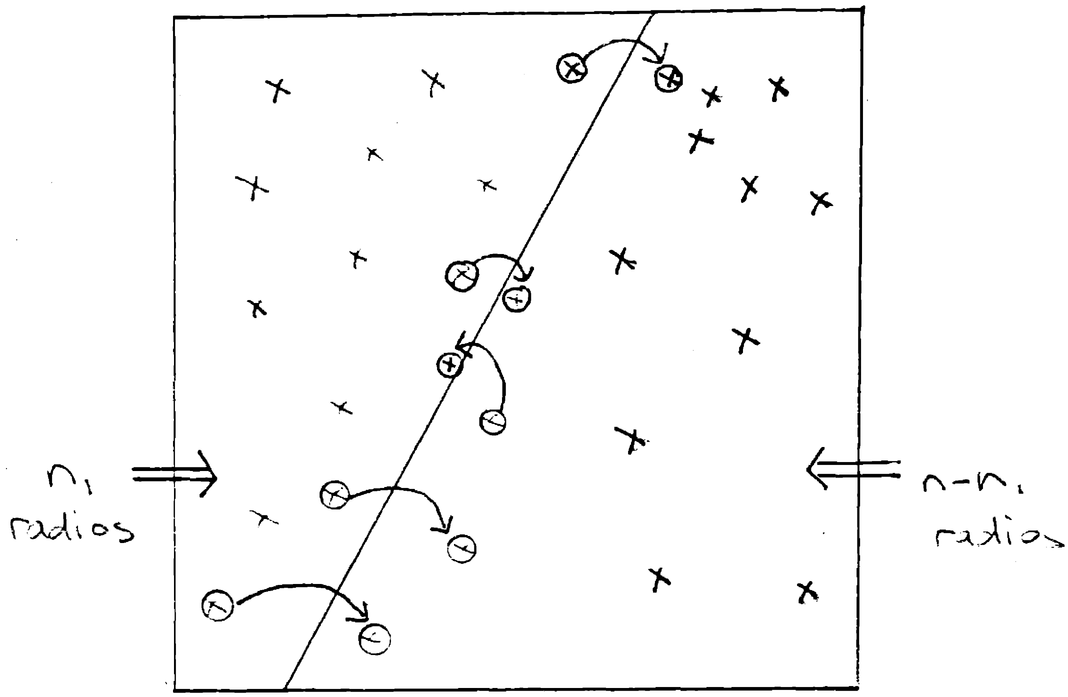

Fig. 5.1.1 A Cut in a Planar Network

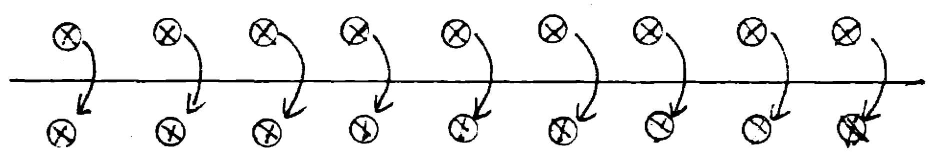

Fig. 5.1.2a Inappropriate Cut for Determining Performance Bound

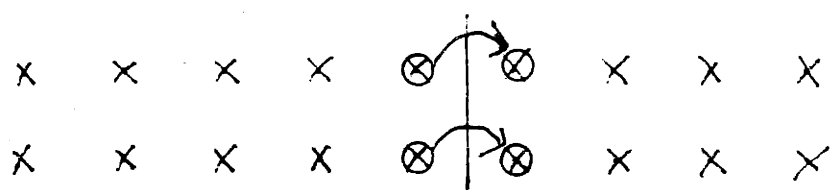

Fig. 5.1.2b Appropriate Cut for Determining Performance Bound

more complicated. First of all, since the above expression holds for any cut, determining an upper bound on throughput (which is equivalent to determining a lower bound on S) requires that we choose the particular cut which maximizes S. While it is unclear how to do this, it seems reasonable that the cut should be made near the middle of the network so that approximately n/2 nodes are located on either side of the cut. This would at least maximize the expression $n_l(n-n_l)$. Were we to do this, we would still have the problem of choosing the orientation of the cut, since some cuts would allow greater numbers of packets to cross than others. Figures 5.1.2a and 5.1.2b show an extreme example of this phenomenon. Note that if the cut is made as in 5.1.2a, n/2 packets can cross per slot from the top half of the network to the bottom half. However, this avoids the issue of communication between the left and right halves of the network. The proper determination of B requires that the cut be made as in Figure 5.1.2b. Note that only 2 packets can cross this cut per slot.

On the other hand, in a random network with a large number of independently located nodes, the probability that the network resembles that shown in Figures 5.1.2a and 5.1.2b is very small. In fact, as n is made large, asymmetries due to the randomness of node location should disappear due to the laws of large numbers and the orientation of the cut should have little effect on B as long as the cut passes through the middle of the network.

The remaining problem is to determine B. This involves choosing transmitter-receiver pairs in such a way that the number of non-interfering pairs is maximized. We are unsure how to do this. Reasonable approaches include:

(1) Choose non-interfering pairs in order of transmitter-receiver distance.

(2) Choose each receiver in order of its distance from the cut and match it with its closest potential transmitter.

These approaches are reasonable in that they lead to transmission radii which typically encompass only a small number of receivers, thus allowing a large set of receivers to be paired with transmitters.

Note that the choice of a non-optimum B does not provide a lower bound on throughput unless it can be shown that a routing strategy exists that enables 5.1.1 to be satisfied with equality. Thus, the presence of problems in determining any formal bound suggest that it may be more realistic to determine a benchmark which is not a bound but nevertheless provides insight into expected throughput.

Instead of using probabalistic methods, which only yield information about the expected throughput over all network topologies, the approach mentioned above could be used as the basis of an algorithm to determine attainable throughputs in networks whose topologies are known. Such an algorithm would use rules similar to those discussed above to choose the position of a cut and to pair up transmitters and receivers. In this way, simulation could be used to determine bounds on performance in networks with arbitrary topologies.

5.2 Slotted-Aloha Networks

Besides Chapter 2, all of our work has been concerned with TDMA-like channel access schemes. However, in networks which are lightly loaded or in which large delays are undesirable, contention resolution channel access schemes are more efficient in terms of delay because no unused slots need be allotted. We have looked at one such scheme, -- slotted-Aloha -- in regular line networks and have discovered that even here, the analysis is complicated. There is much to be done concerning the use of slotted-Aloha in networks which are random and/or planar. Such analyses have been made by [Silv83b], [Klei78] and [Taka84] but not under the assumption of dynamically controllable transmitter power. We believe that the methods employed in Chapter 2 could be generalized to cover such situations but that such analyses would be considerably more complex.

Another interesting extension of our work would be the development of a "cellular slotted-Aloha" scheme in which routes would be established among cells rather then specific radios. In such a scheme, good routing algorithms could be implemented based only on knowlege of the cell which each radio occupies rather then its exact position. For example, dynamic routing algorithms based on updated cellular traffic statistics could be developed so as to equalize traffic over all regions of the network, lessening channel contention in densely populated regions. Collecting statistics by cell rather then by individual radio is advantageous because each radio tends to hear not only packets intended for it but those intended for its closest neighbors; i.e, those occupying the same cell (its cellmates).

In fact, it appears that routes which are good under CDTDMA should also be good under slotted-Aloha since the former should also be designed to smooth out traffic over the network, as we discussed at the end of Chapter 4. Thus, it may be possible to design a good routing algorithm for networks with slotted-Aloha channel access without worrying about the complexities of analyzing slotted-Aloha, which we believe are much greater than those involved in the analysis of CDTDMA.

5.3 More Sophisticated Models

The network model we have used here could be made more sophisticated so as to yield more realistic quantitative results. The most basic change that could be made would be to model the capture effect more realistically. Capture has been considered in [Taka83a] but not in conjunction with either cellular strategies or dynamically-powered transmitters.

Finally, our work could be generalized to include networks which possess non-uniform traffic matrices. An interesting example would be a network whose radios communicate most frequently with their nearest neighbors. In such a network, higher throughputs than those determined here could be attained due to the decrease in the frequency of very long transmissions which tend to interfere with many radios.

References

[Abra70] N. Abramson, "The ALOHA System - Another Alternative for Computer Communications," AFIPS Conference Proceedings, FJCC 37, pp.281-285, (1970).

[Bake81] D. J. Baker and A. Ephremides, "The Architectural Organization of a Mobile Radio Network via a Distributed Algorithm," IEEE Transactions on Communications, Vol. COM-29, No. 10, pp. 1694-1701 (November 1981).

[Kahn78] R. E. Kahn, S. A. Gronemeyer, J. Burchfiel, R. C. Kunzelman, "Advances in Packet Radio Technology," Proceedings of the IEEE, Vol. 66, No. 10, pp. 1468-1496 (November 1978).

[Klei78] L. Kleinrock and J. Silvester, "Optimum Transmission Radii for Packet Radio Networks, or Why Six is a Magic Number," Proceedings of Network Telecommunications Conference, Birmingham, Alabama (December 1978).

[Robe72] L. G. Roberts, "ALOHA Packet System With and Without Slots and Capture," ASS Note 8 (NIC 10290), ARPA Network Information Center, Stanford Res. Inst., Menlo Park, Ca. (June 1972). Reprinted in Computer Communication Review, Vol. 5, pp. 28-42 (April 1975).

[Silv83a] J. A. Silvester and L. Kleinrock, "On the Capacity of Multihop Slotted ALOHA Networks with Regular Structure," IEEE Transactions on Communications, Vol. COM-31, No. 8, pp. 974-982 (August 1983)

[Silv83b] J. A. Silvester and L. Kleinrock, "On the Capacity of Single-Hop Slotted ALOHA Netwoks for Various Traffic Matrices and Transmission Strategies," IEEE Transactions on Communications, Vol. COM-31, No. 8, pp. 983-991 (August 1983).

[Taka84] H. Takagi and L. Kleinrock, "Optimal Transmission Ranges for Randomly Distributed Packet Radio Terminals," IEEE Transactions on Communications, Vol. COM-32, No. 3, pp. 246-257 (March 1984).