CHANNEL ACCESS PROTOCOLS FOR MULTIHOP PACKET RADIO NETWORKS

Si-Taur Wang

COMMUNICATIONS & SIGNAL PROCESSING LABORATORY
Department of Electrical Engineering and Computer Science
The University of Michigan
Ann Arbor, Michigan 48109

November 1986

Technical Report No. 242
Approved for public release; distribution unlimited.

Prepared for
OFFICE OF NAVAL RESEARCH
Department of the Navy
Arlington, Virginia 22217

ACKNOWLEDGEMENTS

I would like to express my appreciation to my doctor committee members: Professors Daniel Atkins, Douglas Smith, Marlin Ristenbatt, Tony Woo, and Wayne Stark. Special thanks go to Professor Douglas Smith, my dissertation chairman, for his guidance and patient good humor in helping me with this research and writing this thesis. To Professors Wayne Stark and Douglas Smith, I offer my sincere thanks for stimulating discussions during the seminar in which I presented the outline of this thesis. Also, I would like to thank Bryan Aupperle for his help with the simulation programs.

I would also like to acknowledge the financial support that I received from the Office of Naval Research under contract number "N00014 85 K 0545" for a portion of this research.

To my parents, brother, and sister, to whom I dedicate this dissertation, I owe them everything. Without their help, I could not finish my college degree, let alone the Ph.D. degree.

Finally, I would like to thank Dr. Chin-Chan Chu of Ford Motor for helping me in everything during the years I stayed in Michigan.

TABLE OF CONTENTS

LIST OF FIGURES

FIGURE

LIST OF TABLES

LIST OF APPENDICES

APPENDIX

LIST OF SYMBOLS

$\alpha_\bullet$ Neighbor constants for regular networks equal to the numbers of nodes adjacent to a node divided by two -- for triangular networks $\alpha_t = 3$, for square networks $\alpha_s = 2$, and for hexagonal networks $\alpha_h = 1.5$

$\beta_\bullet$ Covering constants for strong spatial TDMA -- for triangular networks $\beta_t = 2$, for square networks $\beta_s = 1$, and for hexagonal networks $\beta_h = 1$

$\gamma_\bullet$ Constants for strong spatial TDMA indicating the numbers of neighbors to which nodes can directly transmit -- for triangular networks $\gamma_t = 6$, for square networks $\gamma_s = 4$, and for hexagonal networks $\gamma_h = 3$

$\delta_\bullet$ Numbers of slots in a TREE/WSTDMA frame if transmission ranges of nodes are fixed such that transmissions just reach nodes that are 1-distant for triangular (t), square (s), and hexagonal (h) networks

$\varsigma_{x,y}$ One-hop nodal capacities (supremum if $\dot{\varsigma}$ and time average if $\bar{\varsigma}$) for weak spatial TDMA ($x = w$), strong spatial TDMA ($x = s$), and a variant of strong spatial TDMA ($x = s'$) operating on triangular ($y = t$), square ($y = s$), and hexagonal ($y = h$) networks

λ Mean rate at which packets arrive at nodes

$\nu_\bullet$ Node densities for triangular (t), square (h), and hexagonal (h) networks

ρ Traffic intensity for a queue

τ Maximum propagation delay between two nodes

$\phi_\bullet(k)$ Numbers of slots in a strong spatial TDMA frame if transmission ranges of nodes are fixed such that a transmission just reaches nodes that are k-distant for triangular (t), square (s), and hexagonal (h) networks

$A_\bullet(k,m)$ Numbers of time slots required for k-distant transmissions in the SSTDMA' protocol

$h_{\bullet}$ Expected numbers of hops in triangular (t), square (h), and hexagonal (h) networks

n Number of nodes surrounding the source of data which may be the ultimate destination of the data

N Average number of terminals associated with a repeater in the TREE/TDMA protocol

$n_{\bullet}(m)$ Numbers of nodes within m, k-distant hops of the source of data which may be the ultimate destination of the data for triangular (t), square (s), and hexagonal (h) networks

P Probability that a terminal wishes to transmit to a repeater during a frame

r Distance between adjacent nodes in a regular network, e.g., between repeaters in the TREE/TDMA protocol

$r_{\bullet}$ Density normalized distances between adjacent nodes in regular networks

$R_{\bullet}$ Transmission ranges of repeaters in the TREE/TDMA protocol for triangular ($y=t$), square ($y=s$), and hexagonal ($y=h$) networks

$S_{x,y}$ Network nodal capacities (supremum if $\dot{S}$ and time average if $\bar{S}$) for weak spatial TDMA ($x=w$), strong spatial TDMA ($x=s$), a variant of strong spatial TDMA ($x=s'$), and TREE/TDMA ($x=t$) operating on triangular ($y=t$), square ($y=s$), and hexagonal ($y=h$) networks

CHAPTER 1

INTRODUCTION

1.1. Preliminaries

A packet radio netw rk consists of a collection of geographically distributed and possibly mobile nodes which wish to communicate with one another via one or more broadcast radio channels.

Associated with each node is a network interface unit (NIU) consisting of a communications section, which transmits and receives units of information called packets, and a logic section, which controls the communications section and processes packets. If an NIU serves only as the source or sink of packets, then it is called a terminal; if it serves only to forward packets, then it is called a repeater; and if it serves both as a terminal and a repeater, then it is called a station.

Because of the limitations of NIUs, e.g., insufficient transmission power, packets may be sent to one or more intermediary NIUs, stations or repeaters, before reaching their final destinations and each link in this chain is called a hop. A single-hop network is one in which all NIUs are terminals and each packet is transmitted directly from source to destination. All other networks are multihop networks.

At a specified time and with respect to a certain NIU, a communication channel is idle if no transmission is detected by the NIU on the channel; otherwise, the channel is busy. It is assumed that channels are noiseless and that transmissions are received by all NIUs within specified regions surrounding the transmitters unless two or more transmissions on the same channel "collide" at an NIU, i.e., the transmissions are coincident at an NIU within overlapping transmission regions. Thus power capture and time capture are modelled as follows: an NIU can receive but one transmission on any channel at any time and any other transmission on the same channel within a specified region will interfere. Note, however, that there can be several successful coincident transmissions in different parts of a multihop network, i.e., spatial reuse of the communication channel, if the transmissions do not collide at the intended recipients' NIUs.

Access to a shared communication channel is governed by a set of rules called a channel access protocol. These rules specify the procedure that must be followed by NIUs which wish to use the channel. They dictate when, which, and to whom NIUs may transmit. Protocols may be centralized or distributed: in a centralized protocol a single NIU controls access of all NIUs to the channel and in a distributed protocol each NIU controls its own access to the channel. Protocols may dedicate a portion of the capacity of the communication channel to each NIU or may permit each NIU to randomly access the entire capacity of channel. If decisions at an NIU are based on information about the state of the network which is not available locally, the protocol must specify the manner in which information is exchanged among NIUs.

In multihop networks, a routing algorithm is used to determine the path that a packet follows between its source and ultimate sink. The forward progress of any transmission is defined to be the distance achieved along the line connecting the transmitting NIU to the NIU which is the ultimate sink of the packet, and a routing algorithm which maximizes the forward progress of any transmission is called a most forward progress (MFP) routing algorithm.

The criteria used to evaluate channel access protocols in this dissertation include measures of channel utilization, packet delay time, packet throughput, stability, fairness, and capacity. The definitions of some of these criteria depend on a statistical characterization of the network or the class of networks to which they are applied, i.e., probabilistic specifications of node locations, nodal packet arrival processes, packet length distributions, traffic patterns and routing algorithms.

- The *(effective) channel utilization* of a protocol for a specific network is defined to be the ratio of the expected rate at which data are (successfully) transmitted on the communication channel to the bandwidth of the channel, excluding potential spatial reuse, for a specified statistical characterization of the network. Unless qualified by the term "nodal", channel utilization refers to the global utilization of the channel in the network. (For a homogeneous network with isotropic traffic patterns the global utilizations are just the products of the nodal channel utilizations and the number of nodes in the network.) The average (effective) channel utilization of a protocol for a specific statistical characterization of a network is the (effective) channel utilization averaged over time. Note that for multihop networks the global channel utilization may exceed one because of spatial reuse and may also be normalized to that per unit area.

  For large, homogeneous networks with isotropic traffic patterns in which edge effects can be ignored, having a slotted channel communication channel in which packets are the length of one slot, and having nodes with independent Poisson exogenous packet arrival processes with mean rates λ per slot, then the nodal channel utilization for any node in the t-th slot is

$$E(I_{s,t}(\lambda) + I_{u,t}(\lambda))$$

where $I_{s,t}(\lambda)$ and $I_{u,t}(\lambda)$ are indicator random variables assuming values 1 if and only if there are successful or unsuccessful transmissions, respectively, from the node in the t-th slot; and the effective nodal channel utilization for any node in the t-th slot is

$$E(I_{s,t}(\lambda)).$$

The average nodal channel utilization and the average effective nodal channel utilization are

$$\limsup_{T\to\infty} \frac{1}{T} \sum_{t=0}^{T} E(I_{s,t}(\lambda) + I_{u,t}(\lambda)) \text{ and } \limsup_{T\to\infty} \frac{1}{T} \sum_{t=0}^{T} E(I_{s,t}(\lambda)),$$

respectively. The global channel utilizations are simply the nodal utilizations multiplied by the number of nodes in the network.

- *Packet delay* is defined to be the difference between the time at which a packet appears at its source and the time at which the last bit of a packet is successfully received at its ultimate destination unless qualified by the term "one-hop", in which case the delay is measured from arrival to reception at the next destination. The mean or expected packet delay for a specific protocol and statistical characterization of a network is the expected value of the packet delay.
- The *throughput* for a specific protocol and statistical characterization of a network is the expected rate at which data, typically packets, arrive at their destinations. The average throughput for a specific protocol and statistical characterization of

a network is the rate at which data arrive at their destinations averaged over time. Throughput is usually specified as the rate at which data reach their ultimate destinations unless qualified, e.g., by the term "one-hop". Like utilization, throughput may be normalized to the unit area.

- A protocol is *stable* for some statistical characterization of a network if the average packet delay in the network is finite.
- A protocol is called *fair* if for all homogeneous networks with isotropic traffic patterns, the average packet delays at all nodes in a specific network are equal.
- The *(average) capacity* of a protocol for a homogeneous network with an uniform, isotropic traffic pattern in which the exogenous data arrival processes at nodes are independent Poisson processes with mean rates λ is the maximum for $0 \leq \lambda < \infty$ of the (average) effective channel utilization for a statistical characterization of the network. Unless qualified as "nodal", capacity refers to the global capacity of the network. Unless qualified as "one-hop", capacity refers to the rate at which data are delivered to their ultimate destinations (normalized by the bandwidth of the communication channel).

For large, homogeneous networks with isotropic traffic patterns in which edge effects can be ignored, having a slotted channel communication channel in which packets are the length of one slot, and having nodes with independent Poisson exogenous packet arrival processes with mean rates λ per slot, then the one-hop nodal capacity for any node in the t-th slot is

$$\max_{0 \leq \lambda < \infty} E(U_{i,t}(\lambda))$$

FDMA protocols for the single-hop, finite source, network model is 1.0 and both protocols are stable in this context.

Nelson investigated the use of TDMA in static, multihop, packet radio networks with known topologies and packet flow patterns [Nel82]. He attempted to find a method to assign the time slots in TDMA frames to sets, or cliques, of transmitter-receiver pairs which could converse without interfering with one another in such a way that the average packet delay in the network would be minimized. While he was unable to find an optimal assignment of slots to cliques he was able to find an upper bound to the minimum delay by computing the delay for the case in which cliques are randomly assigned to slots.

Dedicated protocols like FDMA and TDMA can be effective in packet radio networks if the traffic is not bursty in nature. (An information source is called bursty if its burst factor β [Lam78], the ratio of the acceptable average message delay time to the mean message interarrival time, is small.) Dedicated protocols become inefficient if the traffic is bursty because the portion of the entire bandwidth of the channel allocated to each NIU must be large enough to meet message delay constraints; however, the capacity is unused most of the time. It is this inefficiency which motivates the use of variants of strict TDMA protocols like statistical TDMA and of random access protocols like ALOHA and CSMA in which the entire capacity of the communication channel is allocated on packet-by-packet basis when required.

1.2.2. ALOHA protocols

A simple distributed random access protocol for packet radio networks was developed by Abramson and is called ALOHA [Abr70]. (Note that a distinction is and the network nodal capacity for the i-th node is

$$\max_{0 \le \lambda < \infty} \mathrm{E} \sum_{k=1}^{n} I_{i,t,k}(\lambda)$$

where $I_{i,t,k}(\lambda)$ is an indicator random variable assuming value 1 if and only if node k is successfully receiving a transmission that originated at node i, i.e., node i was the original source of the data, and is also the ultimate destination of the data in the t-th slot. (Without loss of generality it is assumed that nodes $1,2,...,n$ are the potential ultimate destinations for data originating at node i.) The average capacities and the global capacities are defined in manners analogous to the average and global channel utilizations, respectively.

1.2. Channel access protocols

In this section, we review two classes of dedicated access protocols, the TDMA protocols and the FDMA protocols; review four classes of random access protocols, the ALOHA protocols, the CSMA protocols, the BTMA protocols, and the tree protocols; and review one class of hybrid protocol, the ALOHA/FDMA protocol. All of these protocols can be used in either single-hop or in multihop networks.

1.2.1. TDMA and FDMA protocols

In time division multiple access (TDMA) and frequency division multiple access (FDMA) a portion of the capacity of the communication channel allocated to each NIU. In TDMA the channel is temporally multiplexed among NIUs and an NIU is assigned one or more time "slots" in a repeating sequence of "frames" during which it can use the channel. In FDMA the channel is frequency multiplexed among NIUs and an NIU is allocated a portion of the frequency spectrum. The capacity of the TDMA and drawn between the ALOHA network and the ALOHA protocol.) In this protocol, whenever an NIU is not busy and a new packet arrives, the packet is transmitted. NIUs make no attempt to ascertain if the communication channel is idle before transmitting. Also, NIUs make no attempt to detect interference or to abort unsuccessful transmissions. In the event that two or more packets collide, none is acknowledged and the sources retransmit the packets after a random delay. Lam found that the capacity of ALOHA for the single-hop, infinite source, network model[1] is $1/2e$ or approximately 0.184. He also showed that ALOHA is unstable in this context [Lam75].

The characteristics of slotted ALOHA , ALOHA with a temporally slotted channel in which NIUs can initiate slot length transmissions only on slot boundaries, were first described by Roberts [Rob72]. Like ALOHA, the protocol is unstable in the context of a single-hop, infinite source, network model; however, its capacity is $1/e$ or approximately 0.368 [Lam75].

In multihop packet radio networks the capacity of the ALOHA protocols may increase since several conversations can coexist without interfering with one another. Silvester coined the term "spatial reuse" to describe this phenomenon and he and Akavia investigated the use of slotted-ALOHA in multihop networks [Aka79, Sil80].

For multidimensional, homogeneous networks, Akavia derived a relationship among the optimal transmission range for an NIU (the range that minimizes the channel bandwidth required to satisfy constraints on average packet delay), the average distance traveled by messages, the total traffic emerging from a unit area of the network, and the desired average packet delay. He showed that ALOHA performed

[1] See Appendix A.

well when the traffic is bursty and that networks in which the transmission range of NIUs could be controlled suffered less from steady traffic than single-hop ALOHA networks. While in the latter case, ALOHA is only $1/e$ times as efficient as the best $M/D/1$ queue; in the former case, it is $1/\sqrt{e}$ times as efficient. Of course, such a queue is an ideal model and cannot be constructed in a distributed environment.

Silvester looked at one- and two-dimensional homogeneous networks in which each NIU is assumed to have the same, limited transmission range and to be equally likely to transmit to any other NIU. He showed that there exists an optimal value for this transmission range for which the expected forward progress of a packet per attempted transmission is maximized. If the transmission range is increased beyond this value, then the expected forward progress decreases due to the increased probability of interference from other transmissions. For two-dimensional networks having regular or random topologies which are neither fully connected nor disconnected, he found that the network nodal capacity is proportional to the reciprocal of the square root of the number of nodes in the network. He also showed that for two-dimensional networks with random topologies the optimal transmission range is such that an average of 5.89 other NIUs can "hear" a transmission.

Nelson generalized Silvester's work to include capture [Nel82] for random planar networks using slotted-ALOHA. He derived a formula for the network capacity as a function of the number of nodes in the network, the average number of nodes within the hearing range of a randomly selected node, the capture parameter of the receivers in the network, and the probability that a node in the network transmits in any slot. He showed that when the capture parameter is 0.7, typical for good FM, the network nodal capacity is $0.074/\sqrt{n}$ and occurs when an average of 4.99 other nodes can hear a transmission.

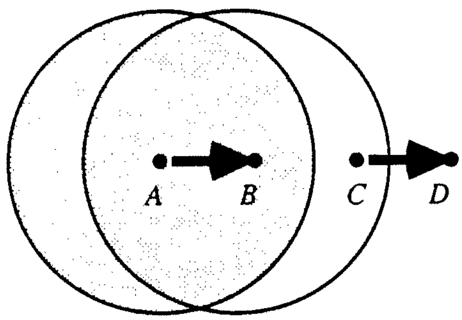
Akavia investigated the performance of hierarchical slotted-ALOHA networks in which many terminals are linked to repeaters and repeaters to one central station [Aka79]. Each repeater is assumed to receive packets from a unique collection of terminals, each repeater is associated with a different collection of terminals, and each collection of terminals is assumed to generate traffic as a Poisson source with common rate. All terminals transmit to local repeaters using the slotted-ALOHA protocol and similarly, each repeater transmits to the central station using slotted-ALOHA. Akavia suggested that such hierarchical networks could be established by either employing capture and assigning different transmission power levels to different hierarchies or by employing FDMA and assigning different frequencies to different hierarchies. Akavia showed that multilevel ALOHA networks of the type described have better performance when heavily loaded than comparable single level ALOHA networks.

1.2.3. CSMA protocols

Tobagi analyzed the performance of a class of carrier sense multiple access (CSMA) protocols in a single-hop packet radio environment [Tob74]. Unlike the ALOHA protocols, the CSMA protocols require that an NIU test the channel for activity before transmitting and defer from transmitting if the channel appears to be busy. In a p-persistent CSMA protocol an NIU (a) defers until the channel is idle and then it either (b) transmits immediately with probability p $(0 < p \leq 1)$, or (c) defers for a short period of time with probability $1-p$ and then reverts to (a). In a non-persistent CSMA protocol, an NIU repeatedly defers for a randomly chosen interval of time until it finds the channel idle, whence it transmits immediately.

In a single-hop network, collisions occur if a transmission is initiated by an NIU before it realizes that another NIU has also started to transmit. (A transmission is vulnerable for a period of at most τ units of time from its start, where τ is the maximum signal propagation time between NIUs in the network.) In the event that two or more packets collide in either a p-persistent or a non-persistent CSMA protocol, no packet is acknowledged and the sources retransmit the packets after a random delay. The capacity of a non-persistent CSMA protocol for the single-hop, infinite source, network model is about 0.815 when the ratio of propagation delay to packet transmission time a is .01 [Kle75a].

The effect of using any CSMA protocol in a multihop environment differs from that of using the same protocol in a single-hop environment. In a single-hop CSMA network, an idle channel guarantees that a transmission will be successful if the effect of the propagation time τ is neglected and a busy channel guarantees that a collision will occur. However, in a multihop CSMA network, neither of the aforementioned is true. These phenomena, depicted in Figures 1.1 and 1.2, are called "hidden area effects" and are inherent in any multihop communication network in which NIUs have limited transmission power. In Figure 1.1, the channel appears idle to node C even though node A is transmitting to node B and any transmission from C will interfere with the transmission from A to B. In Figure 1.2, the channel appears busy to node C; however, were node C to transmit to D, the transmission would be successful and would not interfere with the transmission from node A to B.



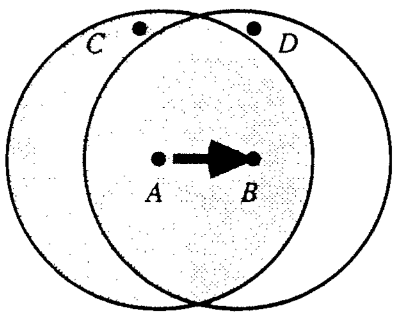
* Node A is transmitting to node B.

* Node C sees an idle channel and transmits to node D.

* The transmission from C to D interferes with that from A to B.

Figure 1.1

Hidden area effect



* Node A is transmitting to node B.

* Node C sees a busy channel and does not transmit to node D.

* The transmission from C to D would not interfere with that from A to B.

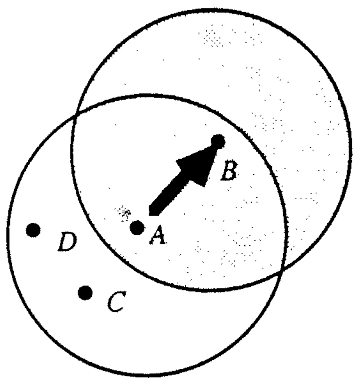
Figure 1.2

CSMA inefficiency

Tobagi showed that the capacity of a non-persistent CSMA protocol decreases from 0.815 to a value less than 0.30 when one-half or more of all NIUs are hidden from each NIU ($a = 0.01$) and he introduced the busy tone multiple access (BTMA) protocol as a modification of the CSMA scheme to eliminate the effect of hidden areas [Tob75].

In his work, Tobagi considered networks with a unique topology: he assumed that the networks were composed of terminals which wish to communicate with a central station, e.g., a centralized computer system. All terminals are within the transmission range of the central station but not necessarily within the transmission range of one another; thus, transmissions from one terminal to the central station may not be apparent to another terminal. To eliminate the hidden area problem, he suggested that the central station transmit a "busy tone" on a subchannel whenever the main channel is busy and that terminals check the status of the subchannel rather than the main channel before transmitting. The capacity of the non-persistent BTMA protocol for a network with the topology described above is approximately 0.70. This is significantly better than the performance of non-persistent CSMA in the same environment but inferior to non-persistent CSMA in an environment in which all nodes can hear one another.

While the BTMA protocol solves the problems that any CSMA protocol experiences when used in a multihop environment, it introduces another problem which is depicted in Figure 1.3. Here the channel appears idle to node C; however, were node C to transmit to node D, the transmission would experience interference from node A and would be unsuccessful. As we show later, this seemingly innocuous problem limits the performance of BTMA multihop packet radio networks. (We also described a method to correct the problem.)



* Node A is transmitting to node B.

* Node C sees an idle channel and transmits to node D.

* The transmission from A to B interferes with that from C to D.

Figure 1.3

BTMA inefficiency

Nelson suggested another approach to improve the performance of CSMA in the multihop networks [Nel82]. He observed that if NIUs were able to estimate the number of nodes that are transmitting within their hearing range, then this information might be used to bias a strict CSMA protocol. In particular, he suggested that a node may want to transmit if few NIUs within its range are transmitting on the chance that its transmission would not interfere with other transmissions. Nelson called this generalization of CSMA, rude-CSMA. He concluded that for random planar networks, rude-CSMA does not improve performance.

1.2.4. Tree protocols

A number distributed, random access protocols, loosely termed "tree" protocols, have been described by Berger [Ber81], Capetanakis [Cap79a, Cap 79b], and Gallager [Gal78]. All of these protocols rely on a communication channel which is temporally slotted and it assumed that NIUs can synchronize their transmissions on slot boundaries. Also, it is assumed that NIUs do not have information about other transmissions on the channel as in CSMA and BTMA, but that NIUs receive feedback about the outcome of any transmission after a transmission has ended.

The operation of these protocols can be viewed as the cyclic searching of an imaginary n-ary tree which is superimposed on all NIUs. Each vertex of the tree corresponds to a subset of NIUs that may transmit in the next time slot if the vertex is scanned -- the root corresponds to the set of all nodes. The m children of any vertex ($m < n$) are disjoint sets, each containing approximately $1/m$ of the NIUs associated with the parent, which are scanned if and only if a collision occurs when their parent is scanned. The structure of the tree, the level at which the search begins, and the searching pattern (typically depth-first or breadth-first) may be varied to optimize the search for NIUs having packets to transmit.

In Capetanakis' seminal work, he assumed that whenever an NIU has a packet to transmit, it waits until the current tree has been expanded and a new tree expansion begins, whence it transmits the first packet in its queue whenever a vertex of which it is a member is scanned, and deletes that packet from its queue only if no collision occurs. He also assumed that all vertices in a tree are binary except for the root which may have more than two children; that a pseudo depth-first search pattern is used starting with the children of the root. Capetanakis showed that the capacity of an optimal tree protocol for the single-hop, infinite source, network model is 0.430. He also showed that for a finite source model the optimal tree protocol becomes a TDMA protocol as the load increases; hence, has a capacity of 1.0.

Berger and Gallager suggest modifications of Capetanakis's searching algorithm which increase the capacity of the tree protocol for the single-hop, infinite source, network model to 0.487.

1.2.5. Hybrid protocols

Akavia suggested that a combination of an ALOHA protocol and FDMA or TDMA might be an effective hierarchical, hybrid protocol for broadcast packet radio networks [Aka79]. In his networks, terminals are linked to repeaters and repeaters to one central station in the case of centralized systems or to each other in the case of distributed networks -- the basic configuration is similar to that described before except that all terminals transmit to local repeaters using the ALOHA protocol and each repeater transmits using a dedicated protocol. Thus Akavia uses ALOHA in an environment to which it is suited, one in which the traffic is bursty, and a dedicated protocol in an environment to which it is suited, one in which the traffic is steady: the traffic from many terminals is combined at the repeaters. He shows that for centralized systems and for one-dimensional distributed networks, the performance of two-level, hybrid protocols is superior to single level protocols over a wide range of operating conditions -- whenever the traffic is neither very bursty nor very steady. He also shows that while additional protocol levels may improve performance in some cases, the magnitude of the improvement is far smaller than that experienced when changing from a single level protocol to a two level protocol.

1.3. Performance limits

There have been a sequence of results that have placed bounds on the capacity of *random access* protocols in packet radio networks. The bounds pertain protocols which do not use supplementary feedback, e.g., channel state in CSMA.

In the single-hop environment, Gallager's tree protocol, the modified version of Capetanakis's protocol, is a random access protocol with a capacity of 0.487 which places a lower bound on the capacity that can be achieved by random access protocols. Pippenger used an information-theoretic argument to show that the capacity of any random access protocol for the single-hop, infinite source network model could not exceed 0.7448 [Pip79] and later Molle used a "genie" argument to show that the capacity for any random access protocol for this model could not exceed 0.6731 [Mol80].

In the multihop environment, Nelson found an upper bound on the one-hop nodal capacity over all random access protocols in connected random planar packet radio networks to be $0.9278/N$ where N is the average number of nodes within the transmission range of a randomly selected node [Nel82]. Lower bounds have have been established by Silvester, Nelson, and Akavia with various assumptions as described earlier.

1.4. Preview

In this dissertation, we address several important issues. First we examine the criteria that must be satisfied for multihop packet radio networks to operate without collisions and derive two general laws for the conflict-free operation of networks. Using these laws, we then place bounds on the operating characteristics of a subclass of all packet radio networks, the regular, planar networks. Next we show that the application of these laws leads to new random access protocols for multihop packet radio networks, a class of bi-level TREE/TDMA protocols and a carrier sense, collision avoidance multiple access (CS/CAMA) protocol. Analytical and simulation models of the protocols are described and are used to evaluate their operating characteristics. A Petri net model of the second of the protocols is used to prove that it operates correctly.

In Chapter 2 the criteria that must be satisfied for the conflict-free operation of multihop networks are examined. Both necessary and sufficient conditions are presented in the form of two "laws," the weak law and the strong law. These laws are then applied to protocols for two-dimensional regular networks, i.e., the triangular, square, and hexagonal networks. Upper bounds are placed on the capacities of protocols which obey the weak law for each of the regular networks and it is shown that there exist protocols achieving these bounds. It is also shown that there exist protocols for regular networks which obey the strong law and have capacities exceeding those that can be obtained by protocols obeying the weak law. It is shown that of protocols for the regular networks, those for square networks obeying the strong law have the highest potential capacities.

In Chapter 3, a new class of channel access protocols for packet radio networks is presented. The protocols are called TREE/TDMA protocols and are based on regular backbones of repeaters which are superimposed on irregular networks of terminals. In these protocols, two levels of control are used on a shared broadcast channel: a tree multiple access (TREE or TMA) link exists between each terminal and a nearby repeater and a spatial TDMA link exists between adjacent repeaters -- the TREE link is embedded in the TDMA framing. Hence a packet is transmitted from its source to a local repeater; then from repeater to repeater until it reaches the vicinity of its destination; whence, it is transmitted to its destination. An analytical model of selected operating characteristics of networks using TREE/TDMA is presented. Necessary and sufficient conditions for the stable operation of these are derived, capacities are calculated, and mean packet delay times are evaluated.

In Chapter 4, a new random access protocol is introduced and analytical models describing the operating characteristics of the protocol are presented. The protocol is a carrier sense, collision avoidance, multiple access (CS/CAMA) protocol and operates under the strong law of conflict-free operation described in Chapter 2 with a single level of control on a shared broadcast channel. Like the TREE/TDMA protocols, CS/CAMA can be used in irregular or random planar networks. In this chapter, we prove that CS/CAMA is essentially conflict-free and that it operates correctly, e.g., is free of deadlocks. We present both analytical and simulation models of the protocol and compare the operating characteristics of the protocol predicted by each model.

The TREE/TDMA and CS/CAMA protocols exhibit characteristics which make them attractive alternatives for use in a multihop packet radio environment. Both protocols are reliable since they are distributed, the latter more so than the former: the NIUs function independently of one another and require only local state information; hence, there does not exist a single node whose failure can render the entire network inoperative as in a centralized network. Both protocols are flexible and can accommodate mobile users. No elaborate mechanism is needed for adding or deleting nodes from a network as in certain "token-passing" and reservation protocols. Both protocols allow terminals to maintain a low EM profile; hence, limit detection: terminals need only transmit when conversing and any transmission can be confined to the neighborhood of the transmitter. Neither protocol requires expensive or complex hardware at an NIU.

Most importantly, the protocols are efficient. Most, if not all, of the earlier random access media access protocols are rendered inefficient in the multihop packet radio environment because either the "hidden area" is ignored by the protocols or in accommodating the hidden area, the potential for spatial reuse of the channel is sacrificed. Both the TREE/TDMA and CS/CAMA protocols eliminate the effect of the hidden area without sacrificing the potential for spatial reuse of the channel.

CHAPTER 2

TWO-DIMENSIONAL REGULAR NETWORKS

2.1. Introduction

In this chapter, we examine the criteria that must be satisfied for multihop packet radio networks to operate without collisions and discuss two general laws for the conflict-free operation of packet radio networks. The weak law stipulates a sufficient condition for the conflict-free operation of networks and is transformed into an operational protocol easily. The strong law stipulates necessary and sufficient conditions for the conflict-free operation of networks and is transformed into an operational protocol with more difficulty than the weak law.

Using these laws we place bounds on the maximum, instantaneous and average, one-hop and network, nodal capacities that can be achieved by protocols for a subclass of all packet radio networks, the two-dimensional regular networks. Upper bounds are placed on the capacities of protocols which obey the weak law for each of the regular networks and it shown that there exist protocols achieving these bounds. It is also shown that there exist protocols for regular networks which obey the strong law and have capacities exceeding those that can be obtained by protocols obeying the weak law. Among the three classes of regular networks, the triangular, square, and hexagonal networks, it is proved that protocols for hexagonal networks have the largest potential network capacities under the weak law, and that protocols for square networks have the largest potential network capacities under the strong law.

Both the weak law and the strong law, and regular two-dimensional networks, play a central role in this chapter and those that follow. In the next chapter, we show that regular networks can be used to form the backbones for more general networks in which nodes are randomly distributed in the plane. Finally in Chapter 4, we describe a protocol which obeys the strong law and also eliminates the problems which limit the performance of the CSMA protocols and BTMA protocols.

2.2. Laws of conflict-free concurrent transmissions

In this section, we present necessary and sufficient conditions for the steady-state, conflict-free, operation of multihop networks in which nodes share a common communication channel. We first describe a sufficient condition, hereafter called the "weak" law, for conflict-free concurrent transmissions. We then describe a necessary and sufficient condition, hereafter called the "strong" law, for conflict-free concurrent transmissions.

The weak law of concurrent conflict-free transmissions

If every transmitter in a network is at least three hops away from other transmitters, then the network is conflict-free.

Proof

- Assume that there exists one network whose transmissions satisfy the condition but is not conflict-free. Then, there exist transmissions which collide. Without loss of generality (WLG) assume that in the first (second) of these transmissions transmitter T_1 (T_2) is transmitting to receiver R_1 (R_2).

- If a collision occurs, then either the transmission from T_1 interferes with R_2, i.e., T_1 is one hop away from R_2, or the transmission of T_2 interferes with R_1, i.e., T_2 is one hop away from R_1. Since T_1 (T_2) is one hop away from R_1 (R_2) by assumption, then T_1 and T_2 are no more than two hops away from each other.
- This contradicts the condition that T_1 and T_2 are at least three hops away from each other. □

The strong law of concurrent conflict-free transmissions

A network is conflict-free if and only if every transmitter in the network is at least two hops away from other receivers excluding the receiver which is the intended recipient of the transmission.

Proof

- If a network is conflict-free, then a receiver can not hear other transmitters excluding its own transmitter. Therefore, all the other transmitters must be at least two hops away from the receiver.
- If every transmitter in a network is at least two hops away from other receivers excluding the receiver which is the intended recipient of the transmission, then one receiver can hear exactly one transmission. Therefore, the network is conflict-free. □

The weak law and the strong law describe conditions which, if satisfied, will ensure the conflict-free operation of a network. Note that these laws implicitly pose constraints on static states of the network in terms of topological constraints on a graph which specify which nodes are neighbors, i.e., who can "hear" whom, and not in terms of temporal constraints. When these laws are applied to a real network, one must adjust the constraints to account for the signal propagation time between neighbors.

2.3. Tessellations of the plane

In this section the structure of regular two-dimensional networks is examined. First we identify the equilateral polygons which can be used to generate regular tessellations of the plane. We then calculate the density of nodes in the plane for each regular network and find the mean number of hops that a packet will travel for each network.

Lemma 2.1.

There are three equilateral polygons which generate regular tessellations of the plane: the triangle, square, and hexagon [Cox69].

Proof

- The degree of an interior angle of an n-sided equilateral polygon is $180(n-2)/n$. Since each vertex belongs to m (a positive integer) equilateral polygons, m times the degree of an interior angle is 360, i.e, $180m(n-2)/n = 360$.
- Since n and m must be positive integers, the above equation has the following three sets of solutions:

  (i) $n = 3$, $m = 6$ which implies that the polygons are triangles;

  (ii) $n = 4$, $m = 4$ which implies that the polygons are squares; and

  (iii) $n = 6$, $m = 3$ which implies that the polygons are hexagons. □

When one of these three planar tessellations is used to represent the structure of a regular communication network, we associate a node or an NIU with each vertex in the tessellation and a communication channel with each edge in the tessellation.

If we assume the distance between two adjacent nodes to be r, then we can calculate the density of the nodes or vertices per unit area in the network as follows.

Lemma 2.2.

The node densities ν_* in triangular, square, and hexagonal networks are, respectively,

$$\nu_t = \frac{2}{\sqrt{3}\, r^2},\ \nu_s = \frac{1}{r^2},\ \text{and}\ \nu_h = \frac{4}{3\sqrt{3}\, r^2}.$$

Proof

- Triangular network - If that there are l triangles on a plane which is large enough so that the edge effect can be neglected, then there are $3l/6$ vertices on the plane since each triangle has three vertices and each vertex is shared by six triangles. The node density is the number of vertices divided by the area covered by the l triangles, i.e.,

$$\nu_t = \frac{\left(\dfrac{l}{2}\right)}{\left(\dfrac{\sqrt{3}\, l r^2}{4}\right)} = \frac{2}{\sqrt{3}\, r^2}.$$

- Square network - Apply the technique used for the triangular network. There are $4l/4$ vertices in a plane tessellated with squares and each square has an area of r^2.
- Hexagonal network - then there are $3l/6$ vertices There are $6l/3$ vertices in a plane tessellated with hexagons and each hexagon has an area of $3\sqrt{3}r^2/2$. □

Lemma 2.3.

In networks in which a packet originating at any node, its source, is equally likely to have any of the n nodes within m, $m \in \mathbb{Z}^+$, hops of the source as its ultimate destination, in which nodes transmit with power sufficient to just reach nodes which are k-distant (node A is k-distant from node B if the shortest path from A to B is of length k), $k \in \mathbb{Z}^+$, and in which nodes use a most forward routing algorithm, the mean numbers of hops a packet makes in traveling from its source to its ultimate destination in triangular, square, and hexagonal networks are, respectively,

$$h_\bullet = \frac{4km^2 + 3(k+1)m - k + 3}{6(mk+1)}$$

for $m = \frac{1}{2k}\left(\sqrt{\frac{4n + \alpha_\bullet}{\alpha_\bullet}} - 1\right)$, $\alpha_t = 3$, $\alpha_s = 2$, and $\alpha_h = 1.5$. If the network is large, i.e., $n >> 1$, then

$$h_\bullet \approx \frac{2m}{3} \text{ or } h_\bullet \approx \frac{2}{3k}\sqrt{\frac{n}{\alpha_\bullet}} \; .$$

Proof

- The numbers of nodes that are within m hops of nodes if transmissions cover nodes that are k-distant are for triangular, square, and hexagonal networks

  $$n_\bullet(m) = \alpha_\bullet mk(mk+1)$$

  for $\alpha_\bullet$ defined for the appropriate network as above. Note that this implies that

  $$m = \frac{1}{2k}\left(\sqrt{\frac{4n + \alpha_\bullet}{\alpha_\bullet}} - 1\right).$$

- Using $n_\bullet(m)$ we find that the numbers of nodes that are exactly i hops away are

$$n_{\bullet}(i) - n_{\bullet}(i-1) = \alpha_{\bullet} k(2ik - k + 1).$$

- Thus the expected numbers of hops are

$$h_{\bullet} = \frac{1}{n_{\bullet}(m)} \sum_{i=1}^{m} i(n_{\bullet}(i) - n_{\bullet}(i-1)) \text{ or}$$

$$h_{\bullet} = \frac{4km^2 + 3(k+1)m - k + 3}{6(mk+1)} \text{ for } m \text{ or n chosen as outlined above.}$$

- For $n >> 1$, $m \approx \frac{1}{k}\sqrt{\frac{n}{\alpha_{\bullet}}}$ and $h_{\bullet} \approx \frac{2m}{3}$ or $h_{\bullet} \approx \frac{2}{3k}\sqrt{\frac{n}{\alpha_{\bullet}}}$. □

2.4. Weak law -- capacity

In this section, we first derive upper bounds on the one-hop nodal capacities (instantaneous and average) that can be achieved by protocols obeying the weak law for triangular, square, and hexagonal networks. Next we place bounds on network nodal capacities of protocols for regular networks with uniform, isotropic traffic patterns. Finally we show that there exists a fair protocol whose one-hop nodal capacities (instantaneous and average) and network nodal capacities equal the respective bounds for each type of regular network.

Theorem 2.1.

Assuming that the transmission range of each node is fixed such that a transmission just reaches nodes that are k-distant, $k \in \mathbb{Z}^+$, then the one-hop nodal capacities (instantaneous and average) of protocols obeying the weak law for triangular, square, and hexagonal networks do not exceed $\varsigma_{w,\bullet} = 1/(n_{\bullet}(1) + 1)$, for $n_{\bullet}(1)$ defined in Lemma 2.3. Furthermore, the bounds on nodal capacities are greatest when k equals 1 and assume values of $1/7$, $1/5$, and $1/4$ for triangular, square, and hexagonal networks, respectively.

Proof

- Since the weak law requires that transmitters be at least three hops away from each other, no node can be within the transmission range of two different transmitters. Thus we can bound the instantaneous one-hop nodal capacity, hence all other types of nodal capacity, by estimating the largest fraction of nodes in a network that can transmit at any time.
- In a triangular network, there are at least $n_t(1) = 3k(k+1)$ nodes within the transmission range of any transmitter (see Figure 2.1). Thus the fraction of nodes that can transmit at the same time must not exceed $1/(n_t(1)+1)$. Since k is an positive integer, the maximum one-hop nodal capacity is $1/7$ when $k = 1$.

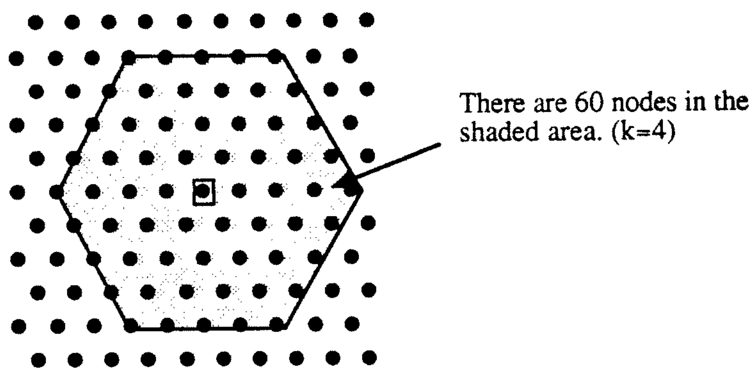


Figure 2.1

Weak law -- triangular networks

- In a square network, there are at least $n_s(1) = 2k(k+1)$ nodes within the transmission range of any transmitter (see Figure 2.2). Thus the fraction of nodes that can transmit at the same time must not exceed $1/(n_s(1)+1)$. Since k is an positive integer, the maximum one-hop nodal capacity is $1/5$ when $k = 1$.

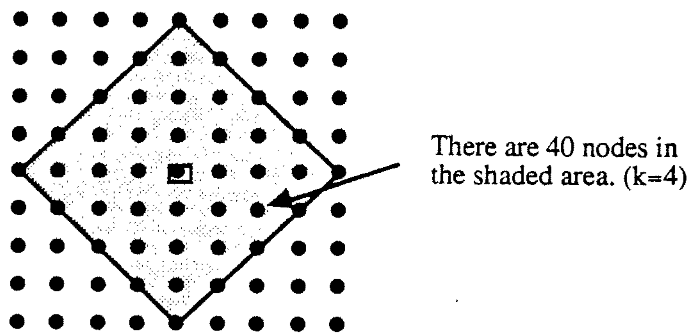


Figure 2.2

Weak law -- square networks

• In a hexagonal network, there are at least $n_h(1) = 3k(k+1)/2$ nodes within the transmission range of any transmitter (see Figure 2.3). Thus the fraction of nodes that can transmit at the same time must not exceed $1/(n_h(1)+1)$. Since k is an positive integer, the maximum one-hop nodal capacity is $1/4$ when $k = 1$. □

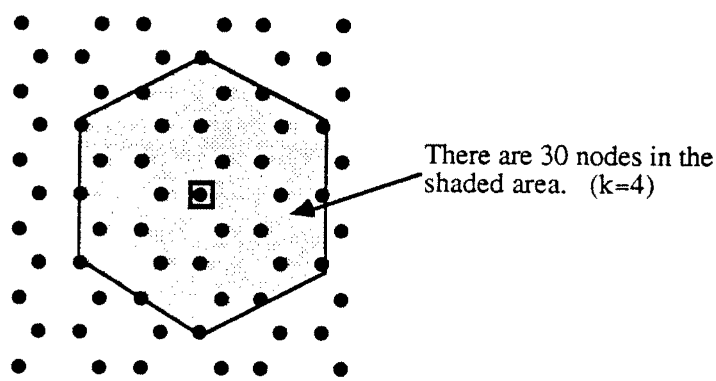


Figure 2.3

Weak law -- hexagonal networks

Theorem 2.2.

Assuming that the transmission range of each node is fixed such that a transmission just reaches nodes that are k-distant, $k \in \mathbb{Z}^+$, and that any of the n nodes within m hops of the source of data is equally likely to be its destination, then the network nodal capacities (instantaneous and average) of protocols obeying the weak law for triangular, square, and hexagonal networks do not exceed

$$\dot{S}_{w,\bullet} = \frac{\dot{\varsigma}_{w,\bullet}}{h_\bullet},$$

for $\dot{\varsigma}_{w,\bullet}$ and $h_\bullet$ as defined in Theorem 2.1 and Lemma 2.3, respectively.

The maximum values of $\dot{S}_{w,\bullet}$ occur when k equals 1 and are

$$\dot{S}_{w,\bullet} = \frac{3}{(2\alpha_\bullet + 1)(2m + 1)} \text{ for } \alpha_\bullet \text{ as defined before and}$$

$$m = \frac{1}{2}\left(\sqrt{\frac{4n + \alpha_\bullet}{\alpha_\bullet}} - 1\right) .$$

If the network is large, i.e., $n >> 1$, then

$$\dot{S}_{w,\bullet} \approx \frac{3\alpha_\bullet k}{2n\,(\alpha_\bullet k^2 + \alpha_\bullet k + 1)} \sqrt{\frac{n}{\alpha_\bullet}} .$$

For large networks the maximum values of $\dot{S}_{w,\bullet}$ $(k = 1)$ are

$$\frac{3\alpha_\bullet}{2n\,(2\alpha_\bullet + 1)} \sqrt{\frac{n}{\alpha_\bullet}} \quad \text{or} \quad \frac{0.3712}{\sqrt{n}}, \quad \frac{0.4243}{\sqrt{n}}, \quad \text{and} \quad \frac{0.4593}{\sqrt{n}}$$

for triangular, square, and hexagonal networks, respectively. If these values are normalized to the unit area, we find the corresponding bounds per unit area based on $\nu_\bullet$ to be

$$\frac{0.4286}{r^2\sqrt{n}},\ \frac{0.4243}{r^2\sqrt{n}},\ \text{and}\ \frac{0.3536}{r^2\sqrt{n}}.$$

Proof

- The network nodal capacity is the one-hop nodal capacity divided by the expected number of hops encountered by data. To prove this for node i, $i \in lb\,1,\ldots,nrb$, define indicator random variables X_j , Y_k, and Z_j, $j,k=1,\ldots,n$. Let X_j be 1 if and only if node j is transmitting data that originated at node i; let Y_k be 1 if and only if node k is receiving data that originated at node i and is an intended destination, i.e., on the most forward path from the source to the ultimate destination; let Z_k be 1 if and only if node k is receiving data that originated at node i and is also the ultimate destination of the data. Now

$$h = \mathrm{E}\left(\frac{\sum_1^n Y_k}{\sum_1^n Z_k}\right) = \frac{\mathrm{E}\sum_1^n Y_k}{\mathrm{E}\sum_1^n Z_k},$$

since the ultimate destination is randomly chosen from the n nodes and hence from the nodes receiving data from i. Since $\dot{S}_{w,\bullet} \triangleq \mathrm{E}\sum_1^n Z_k$ and $\sum_1^n X_j = \sum_1^n Y_k$,

$$\dot{S}_{\wp,\bullet} = \frac{\mathrm{E}\sum_1^n X_j}{h}.$$

Furthermore, $\dot{\varsigma}_{w,\bullet} = \mathrm{E}\sum_1^n X_j$ since the sum of the traffic originating at node i (primary transmissions by i and secondary transmissions by other nodes which are forwarding data originating at i) must equal the rate at which nodes transmits data in the network, i.e., the one-hop nodal capacity, since traffic is uniform and isotropic.

- By inserting the appropriate values from Theorem 2.1 and Lemma 2.3 the general formula for $\dot{S}_{w,\bullet}$ is found to be

$$\frac{6(mk+1)}{[\alpha_\bullet k(k+1)+1][4km^2+3(k+1)m-k+3]}$$

and the first portion of the theorem is proved.

- By substituting 1 for k the second portion of the theorem is proved.
- The third portion of the theorem follows from the fact that for $n >> 1$

$$h_\bullet \approx \frac{2}{3k}\sqrt{\frac{n}{\alpha_\bullet}} \quad \text{(see Lemma 2.3).}$$

- The fourth portion of the theorem follows by substituting 1 for k and 3, 2, and 1.5 for α_t, α_s, and α_h in the preceding.
- The last portion of the theorem follows if the network nodal capacities are multiplied by the appropriate densities of nodes found in Lemma 2.2. □

Theorem 2.3.

There exists a fair channel access protocol, obeying the weak law, whose one-hop and network nodal capacities (instantaneous and average) equal the bounds established in Theorem 2.1 and Theorem 2.2 for large regular networks.

Proof

- The protocol whose capacities equal the bounds is called weak spatial TDMA (WSTDMA) and can be envisioned as a TDMA protocol in which the location of a node in the network determines the the slot in a TDMA frame in which the node can transmit. Each TDMA frame in WSTDMA is divided into $n_\bullet(1)+1$ slots, i.e., seven slots for triangular networks, five slots for square networks, and four slots for hexagonal networks. Associated with each slot is a set of nodes or a clique whose members are allowed to transmit to nodes k-distant when the slot appears in the TDMA frame. The routing algorithm used in WSTDMA is a most forward progress algorithm.

Figures 2.4, 2.5, and 2.6 depict cliques when k is 1 for triangular, square, and hexagonal networks. In these figures, nodes bearing the number i belong to the same clique and may transmit in the ith slot of the TDMA frame.

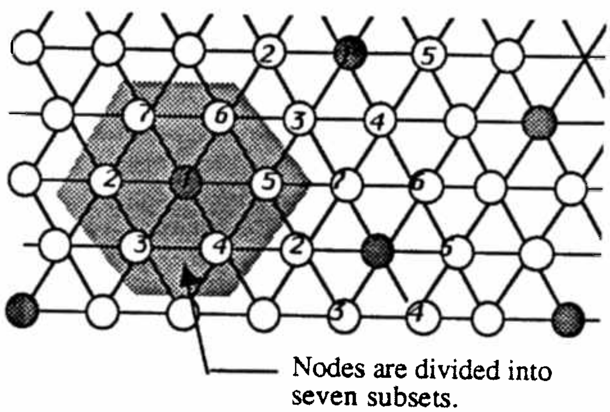


Figure 2.4

Weak spatial TDMA -- triangular networks

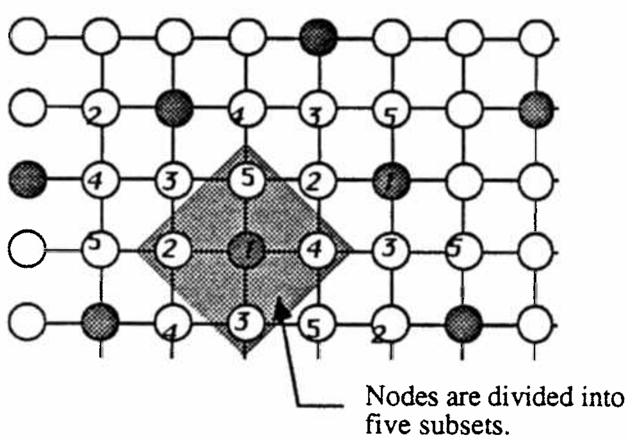


Figure 2.5

Weak spatial TDMA -- square networks

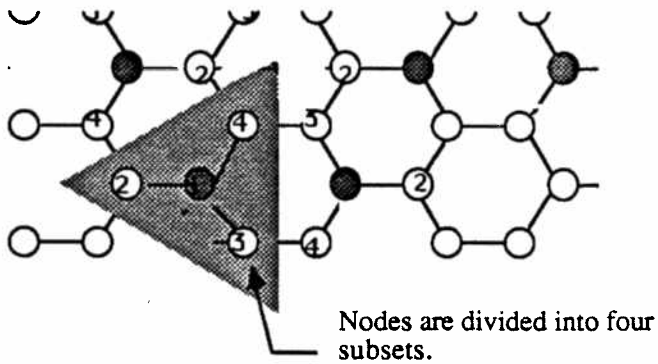


Figure 2.6

Weak spatial TDMA -- hexagonal networks

- WSTDMA is fair since each node in a network has equal access to each frame. Hence if a network is homogeneous and the traffic pattern is isotropic, then the average packet delays at all nodes in the network must be equal.
- WSTDMA obeys the weak law since no two nodes which are transmitting are within one or two hops of one another.
- The capacities of WSTDMA equal the respective bounds for triangular, square, and hexagonal networks because if one assumes that each node in a network always has a packet to transmit, then each node transmits once every seven time slots, five time slots, and four time slots for the respective networks. Since collisions cannot occur in WSTDMA, the instantaneous nodal capacities are always 1/7, 1/5, and 1/4, respectively, and since the nodal capacities are equal in each time slot for a chosen tessellating polygon, the average nodal capacities are equal to their corresponding instantaneous values. Network nodal capacities can

be obtained by dividing the appropriate one-hop nodal capacity by the mean number of times that a packet is transmitted -- this is equal to the expected number of hops calculated in Lemma 2.3 since a most forward routing algorithm is used in WSTDMA. □

The results of the preceding theorems are summarized in Table 2.1. It is clear from the table that hexagonal networks have the largest capacities of the three regular networks unless capacities are normalized to the unit area, whence square networks have the largest one-hop capacities and triangular networks have the largest network capacities.

Nodal Capacity	Triangular Networks	Square Networks	Hexagonal Networks
One-Hop	0.1429	0.2000	0.2500
Network	$\frac{0.3712}{\sqrt{n}}$	$\frac{0.4243}{\sqrt{n}}$	$\frac{0.4593}{\sqrt{n}}$
One-Hop Per Unit Area	$\frac{0.1650}{r^2}$	$\frac{0.2000}{r^2}$	$\frac{0.1925}{r^2}$
Network Per Unit Area	$\frac{0.4286}{r^2\sqrt{n}}$	$\frac{0.4243}{r^2\sqrt{n}}$	$\frac{0.3536}{r^2\sqrt{n}}$

Table 2.1

Weak law -- WSTDMA capacities for large regular networks

2.5. Strong law -- capacity

In this section, we place bounds on nodal and network capacities (instantaneous and average) that can be achieved by protocols obeying the strong law for triangular, square, and hexagonal networks. While we are unable find upper bounds to the capacities as we did in the preceding section, other than the obvious bound on the one-hop nodal capacities of $1/2$, we do place lower bounds on the capacities that can be obtained by protocols obeying the strong law for regular networks. The bounds demonstrate that there exist protocols obeying the strong law whose capacities exceed those that can be achieved by any protocol obeying the weak law.

Theorem 2.4.

Assuming that the transmission range of each node is fixed such that a transmission just reaches nodes that are k-distant, $k \in \mathbf{Z}^+$, then there exists a fair channel access protocol, obeying the strong law, whose one-hop nodal capacities (instantaneous and average) for triangular, square, and hexagonal networks are $\varsigma_{s,\bullet} = 1/(\beta_\bullet k + 1)$ for $\beta_t = 2$, $\beta_s = 1$, and $\beta_h = 1$. These capacities are greatest when k equals 1 and assume values of $1/3$, $1/2$, and $1/2$, respectively.

Proof

- The protocol is called strong spatial TDMA (SSTDMA) and is an extension of the WSTDMA protocol presented in Theorem 2.3. In SSTDMA each frame is divided into $\phi_\bullet(k) \triangleq \gamma_\bullet(\beta_\bullet k + 1)$ slots where $\gamma_t = 6$, $\gamma_s = 4$, and $\gamma_h = 3$ if $k = 1$ and $\gamma_h = 6$ otherwise. (Note that $\gamma_\bullet$ indicates the number of neighbors that are k-distant to which a specific node can transmit directly.) Associated with each slot is a set of nodes or a clique whose members are allowed to transmit to one k-

distant node when the slot appears in the TDMA frame. Figure 2.7 depicts successive states of nine nodes of a square network when k equals 1 during each of the eight slots in the frame. The states of the remaining nodes in the network can be inferred from those depicted. The state of each node is specified by a pair of letters in which the first member "T" ("R") indicates that a node is transmitting (receiving) and the second member "L" ("R") indicates that the operation is to or from the left (right). Of course, the actual operation is assumed to omnidirectional; however, due to placement of transmitters, but one node can actually receive the transmission another node without interference. The routing algorithm used in SSTDMA is a most forward progress algorithm.

	Slot 1			Slot 2			Slot 3			Slot 4	
T,L	T,R	R,L	T,U	T,U	T,U	R,R	T,L	T,R	R,D	R,D	R,D
T,L	T,R	R,L	T,D	T,D	T,D	R,R	T,L	T,R	T,U	T,U	T,U
T,L	T,R	R,L	R,U	R,U	R,U	R,R	T,L	T,R	T,D	T,D	T,D

	Slot 5			Slot 6			Slot 7			Slot 8	
R,L	R,R	T,L	R,U	R,U	R,U	T,R	R,L	R,R	T,D	T,D	T,D
R,L	R,R	T,L	R,D	R,D	R,D	T,R	R,L	R,R	R,U	R,U	R,U
R,L	R,R	T,L	T,U	T,U	T,U	T,R	R,L	R,R	R,D	R,D	R,D

Figure 2.7

Strong spatial TDMA -- states (square networks, $k = 1$)

- SSTDMA is fair since each node in a network has equal access to each frame. Hence if a network is homogeneous and the traffic pattern is isotropic, then the average packet delays at all nodes in the network must be equal.

- SSTDMA obeys the strong law since no node which is receiving is within one hop of one another node which is transmitting, excluding the node directing its transmission to the receiver.
- The one-hop capacities of SSTDMA are as presented because if one assumes that each node in a network always has a packet to transmit, then each node transmits in one third, one half, and one half of all time slots for triangular, square, and hexagonal networks, respectively. Since collisions cannot occur at a node which is the intended recipient of a transmission, the nodal capacities are always $1/3$, $1/2$, and $1/2$, and since the nodal capacities are equal in each time slot for a chosen tessellating polygon, the average nodal capacities are equal to their corresponding instantaneous values. □

There is an interesting feature of SSTDMA which is not present in WSTDMA and which influences its network nodal capacities. Namely, the node which is the ultimate destination of data must at a distance from the source which is an integral multiple of k or the data will never reach the destination. Indeed for some networks, only a subset of nodes that are k-distant, $k > 1$, from a chosen source can be the ultimate destination for data. Thus if the ultimate destination of any data is equally likely to be one of the n nodes within m hops and a single value of k must be chosen, then the only reasonable choice of k is 1.

The following theorem relates the network nodal capacities of SSTDMA based on the preceding observations.

Theorem 2.5.

Assuming that the transmission range of each node is fixed such that a transmission just reaches nodes that are 1-distant and that any of the n nodes within m hops of the source of data is equally likely to be its destination, then the network nodal capacities (instantaneous and average) of the SSTDMA protocol for triangular, square, and hexagonal networks are

$$S_{\iota,\bullet} = \frac{\xi_{\iota,\bullet}}{h_\bullet} = \frac{3}{(\beta_\bullet + 1)(2m + 1)}$$

where $\varsigma_{\iota,\bullet}$ and $h_\bullet$ are defined in Theorem 2.4 and Lemma 2.3, respectively.

For large networks, i.e., $n >> 1$, then

$$S_{\iota,\bullet} \approx \frac{3\alpha_\bullet}{2n\,(\beta_\bullet + 1)} \sqrt{\frac{n}{\alpha_\bullet}} \quad \text{or} \quad \frac{0.8660}{\sqrt{n}}, \quad \frac{1.0607}{\sqrt{n}}, \quad \text{and} \quad \frac{0.9186}{\sqrt{n}}$$

for triangular, square, and hexagonal networks, respectively. If these values are normalized to the unit area, we find the corresponding bounds per unit area based on $\nu_\bullet$ to be

$$\frac{1.0000}{r^2\sqrt{n}}, \quad \frac{1.0607}{r^2\sqrt{n}}, \quad \text{and} \quad \frac{0.7071}{r^2\sqrt{n}}.$$

Proof

- The network nodal capacity is the one-hop nodal capacity divided by the expected number of hops encountered by data. (see Theorem 2.2). By inserting the appropriate values from Theorem 2.4 and Lemma 2.3 the general formula for $S_{\iota,\bullet}$ is found to be

$$\frac{6(mk + 1)}{[\beta_\bullet k + 1][4km^2 + 3(k + 1)m - k + 3]}$$

and by substituting 1 for k the first portion of the theorem is proved.

- The second portion of the theorem follows from the fact that for $n >> 1$

$$h_\bullet \approx \frac{2}{3k}\sqrt{\frac{n}{\alpha_\bullet}} \quad \text{(see Lemma 2.3)}$$

and by substituting 1 for k and 3, 2, and 1.5 for α_t, α_s, and α_h in the preceding.

- The last portion of the theorem follows if the network nodal capacities are multiplied by the appropriate densities of nodes found in Lemma 2.2. □

The results of the preceding theorems are summarized in the following table. It is clear from the table that square networks have the largest capacities (of all types) of the three regular networks. Note too, that the lower bounds on the capacities of protocols for regular networks established for the SSTDMA protocol exceed the upper bounds (achieved by WSTDMA) of their counterparts under the weak law (see Table 2.1).

Nodal Capacity	Triangular Networks	Square Networks	Hexagonal Networks
One-Hop	0.3333	0.5000	0.5000
Network	$\frac{0.8660}{\sqrt{n}}$	$\frac{1.0607}{\sqrt{n}}$	$\frac{0.9186}{\sqrt{n}}$
One-Hop Per Unit Area	$\frac{0.3849}{r^2}$	$\frac{0.5000}{r^2}$	$\frac{0.3849}{r^2}$
Network Per Unit Area	$\frac{1.0000}{r^2\sqrt{n}}$	$\frac{1.0607}{r^2\sqrt{n}}$	$\frac{0.7071}{r^2\sqrt{n}}$

Table 2.2

Strong law -- SSTDMA capacities for large regular networks

As noted before, if SSTDMA is used in its pure form, i.e., nodes can only transmit at a fixed power level, and if destinations are randomly chosen from nodes m-distant or less, then the only realistic choice for k is 1. If, however, nodes can transmit at more than one power level, then there is an interesting variant of SSTDMA whose nodal capacities exceed those obtained by pure SSTDMA. These results are summarized in the following theorems.

Theorem 2.6.

Assuming that the transmission range of each node is variable and can be selected so that a transmission reaches nodes that are k-distant, $k = 1,...,m$, and that any of the n nodes m-distant from the source of data is equally likely to be the destination of the data, then there exists a variant of SSTDMA, SSTDMA', which is fair, obeys the strong law, and has one-hop nodal capacities (instantaneous and average) of

$$\varsigma_{l',\bullet} = \frac{6m}{2\beta_\bullet m^2 + (3\beta_\bullet + 6)m + \beta_\bullet}$$

for triangular, square, and hexagonal networks, respectively ($h_\bullet$ is defined in Lemma 2.3).

For large networks, i.e., $n >> 1$,

$$\varsigma_{l',\bullet} \approx \frac{3\alpha_\bullet}{n\,\beta_\bullet}\sqrt{\frac{n}{\alpha_\bullet}} \quad \text{or} \quad \frac{2.5981}{\sqrt{n}}, \quad \frac{4.2426}{\sqrt{n}}, \quad \text{and} \quad \frac{3.6742}{\sqrt{n}}$$

for triangular, square, and hexagonal networks, respectively. If these values are normalized to the unit area, we find the corresponding bounds per unit area based on $\nu_\bullet$ to be

$$\frac{3.0000}{r^2\sqrt{n}}, \ \frac{4.2426}{r^2\sqrt{n}}, \text{ and } \frac{2.8284}{r^2\sqrt{n}}.$$

Proof

- The protocol is a variant SSTDMA presented in Theorem 2.4. In SSTDMA' each node is allowed to transmit to every node within a range of a transmission reaching m-distant nodes via a path with the fewest number of hops. For example there twelve 1-distant transmissions and four 2-distant transmissions required to reach all the nodes within the range of 2-distant nodes from a specific node in a square network. Since this must be true for each node in the network and the strong law must not be violated, this requires twenty-four slots for the 1-distant transmissions and twelve slots for the 2-distant transmission if SSTDMA is used. In general $A_\bullet(k,m) \triangleq (2m - 2k + 1)\phi_\bullet(k)$ slots are required for all k-distant transmissions.
- SSTDMA' is fair since each node in a network has equal access to each frame. Hence if a network is homogeneous and the traffic pattern is isotropic, then the average packet delays at all nodes in the network must be equal.
- SSTDMA' obeys the strong law since no node which is receiving is within one hop of one another node which is transmitting, excluding the node directing its transmission to the receiver.
- If one assumes that each node in a network always has a packet to transmit to the destination associated with the next available slot, then the expected one-hop nodal capacities of SSTDMA' are

$$\varsigma_{\ell,\bullet} = \frac{1}{\sum_{k=1}^{m} A_\bullet(k,m)} \sum_{k=1}^{m} \frac{A_\bullet(k,m)}{k\beta_\bullet + 1} = \frac{6m}{2\beta_\bullet m^2 + (3\beta_\bullet + 6)m + \beta_\bullet}$$

- For $n >> 1$, $m \approx \sqrt{\frac{n}{\alpha_\bullet}}$ (see Lemma 2.3); hence,

$$\varsigma_{\ell,\bullet} \approx \frac{3\alpha_\bullet}{n\beta_\bullet}\sqrt{\frac{n}{\alpha_\bullet}}$$

and by substituting appropriate values of $\alpha_\bullet$ and $\beta_\bullet$ the one-hop nodal capacities can be found for triangular, square, and hexagonal networks.

- The last portion of the theorem follows if the one-hop nodal capacities are multiplied by the appropriate densities of nodes found in Lemma 2.2. □

Theorem 2.7.

Assuming that the transmission range of each node is variable and can be selected so that a transmission reaches nodes that are k-distant, $k = 1,...,m$, and that any of the n nodes within m hops of the source of data is equally likely to be the destination of the data, then the network nodal capacities (instantaneous and average) of SSTDMA' for triangular, square, and hexagonal networks are

$$S_{\ell,\bullet} = \frac{\xi_{\ell,\bullet}}{h'_\bullet}$$

where $\varsigma_{\ell,\bullet}$ is defined in Theorem 2.5 and

$$h'_\bullet = \frac{4\beta_\bullet m^2 + (3\beta_\bullet + 12)m - \beta_\bullet - 6}{2\beta_\bullet m^2 + (3\beta_\bullet + 6)m + \beta_\bullet}$$

For large networks, i.e., $n >> 1$, $h'_\bullet \approx 2$ and

$$S_{\ell,\bullet} \approx \frac{3\alpha_\bullet}{2n\beta_\bullet}\sqrt{\frac{n}{\alpha_\bullet}} \quad \text{or} \quad \frac{1.2991}{\sqrt{n}}, \quad \frac{2.1213}{\sqrt{n}}, \quad \text{and} \quad \frac{1.8371}{\sqrt{n}}$$

for triangular, square, and hexagonal networks, respectively. If these values are normalized to the unit area, we find the corresponding bounds per unit area based on $\nu_\bullet$ to be

$$\frac{1.5000}{r^2\sqrt{n}},\ \frac{2.1213}{r^2\sqrt{n}},\ \text{and}\ \frac{1.4142}{r^2\sqrt{n}}.$$

Proof

- The network nodal capacity is the one-hop nodal capacity divided by the expected number of hops encountered by data. (see Theorem 2.2). The expected number of hops for SSTDMA' is not $h_\bullet$ but

$$h'_\bullet = \frac{\sum_{k=1}^{m} (\phi_\bullet(k) + 2(A_\bullet(k,m) - \phi_\bullet(k)))}{\sum_{k=1}^{m} A_\bullet(k,m)}$$

$$= \frac{4\beta_\bullet m^2 + (3\beta_\bullet + 12)m - \beta_\bullet - 6}{2\beta_\bullet m^2 + (3\beta_\bullet + 6)m + \beta_\bullet}.$$

- The second portion of the theorem follows from the fact that for $n >> 1$, $h_\bullet \approx 2.0$.
- The last portion of the theorem follows if the network nodal capacities are multiplied by the appropriate densities of nodes found in Lemma 2.2. □

The results of the preceding theorems are summarized in the following table. Note that the lower bounds on the capacities of protocols for regular networks established for SSTDMA' exceed those achieved by SSTDMA (see Table 2.2); hence, they also exceed the upper bounds (achieved by WSTDMA) of their counterparts under the weak law (see Table 2.1).

Nodal Capacity	Triangular Networks	Square Networks	Hexagonal Networks
One-Hop	$\frac{2.5981}{\sqrt{n}}$	$\frac{4.2426}{\sqrt{n}}$	$\frac{3.6742}{\sqrt{n}}$
Network	$\frac{1.2991}{\sqrt{n}}$	$\frac{2.1213}{\sqrt{n}}$	$\frac{1.8371}{\sqrt{n}}$
One-Hop Per Unit Area	$\frac{3.0000}{r^2\sqrt{n}}$	$\frac{4.2426}{r^2\sqrt{n}}$	$\frac{2.8284}{r^2\sqrt{n}}$
Network Per Unit Area	$\frac{1.5000}{r^2\sqrt{n}}$	$\frac{2.1213}{r^2\sqrt{n}}$	$\frac{1.4142}{r^2\sqrt{n}}$

Table 2.3

Strong law -- SSTDMA' capacities for large regular networks

CHAPTER 3

TREE/TDMA CHANNEL ACCESS PROTOCOLS

3.1. Introduction

In this chapter a new class of channel access protocols, the TREE/TDMA protocols, is introduced and analytical models describing the operating characteristics of the class are presented. The TREE/TDMA protocols are natural extensions of the spatial TDMA protocols described in the previous chapter. While WSTDMA and SSTDMA are limited in application to regular planar networks, the TREE/TDMA protocols can be used in irregular or random planar networks, e.q., networks containing mobile terminals.

In the TREE/TDMA protocols a regular network of repeaters is superimposed on an irregular network of terminals and two levels of control are used on a shared broadcast channel. Any of the regular networks described in the preceding chapter can be used as the backbone of repeaters and the protocols can be structured so that they obey either the weak law or the strong law.

We first describe the TREE/TDMA protocols and illustrate their application under the weak law to random planar networks on which are superimposed triangular networks. We find necessary and sufficient conditions for the stable operation of the protocols, nodal capacities (of all types) of the protocols, and mean packet delays of the protocols. We also compare the network capacities of the protocols using the weak law for triangular, square, and hexagonal backbones.

3.2. Description

The TREE/TDMA protocols are based on a regular backbone of repeaters which is superimposed on a network of terminals. The backbone is the conduit for inter terminal communication. Any of the regular networks described in the preceding chapter can be used as the backbone of repeaters. It is assumed that the inter repeater distances are r, or $r_\bullet$ if normalized so that the respective densities $\nu_\bullet$ are equal, i.e,

$$r_t^2 = \frac{2}{\sqrt{3}\nu}, \quad r_s^2 = \frac{1}{\nu}, \text{ and } \quad r_h^2 = \frac{4}{3\sqrt{3}\nu}.$$

It is also assumed that repeaters service collections of terminals within radii $R_\bullet$ of the repeater. If $R_\bullet$ are chosen to be the smallest values such that all of the terminals are serviced by at least one repeater, then $R_\bullet$ are $r/\sqrt{3}$, $r/\sqrt{2}$, and r for the respective backbone networks with r optionally replaced by $r_\bullet$. A typical network with a triangular backbone of repeaters is depicted in Figure 3.1.

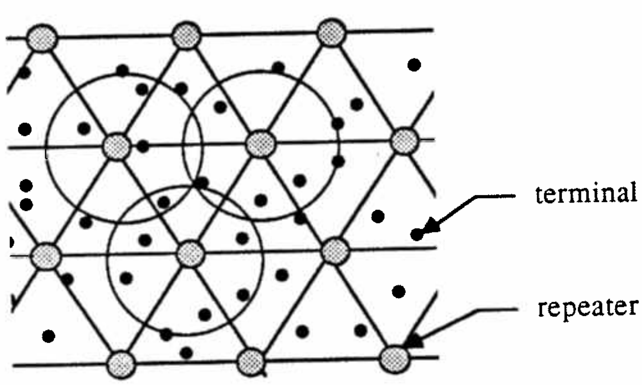


Figure 3.1

Network with a triangular backbone

In the TREE/TDMA protocols a packet is transmitted from its source to a nearby repeater; then from repeater to repeater until it reaches the vicinity of its destination; whence, it is transmitted to its destination. A most forward routing algorithm is used. Two different channel access protocols are used for terminal to repeater communication, for repeater to repeater communication, and for repeater to terminal communication. A tree multiple access (TMA or TREE) protocol, similar to a centralized version of Capetanakis' tree protocol, is used for terminal to repeater communication and a TDMA protocol is used for all other communications. The two access protocols share a common broadcast channel -- TMA is embedded in TDMA. The TDMA protocol may be structured so that either the weak law or the strong law is obeyed. The actual form assumed by a specific TREE/TDMA protocol depends on the type of backbone, the law of conflict-free operation, and the transmission range of repeaters.

The framework for a TREE/TDMA protocol is provided by a spatial TDMA protocol. If the weak law is the basis of operation and the transmission range of repeaters is chosen to maximize the capacity of the backbone, then each frame contains $\delta_{*} \triangleq 2\alpha_{*} + \beta_{*} + 3$ slots. Each of these slots is sufficiently long to allow a packet to be transmitted without interference from adjacent slots. One of these slots is allocated for terminal to repeater communication via TMA. "$2\alpha_{*} + 1$" slots are allocated for repeater to repeater communication using WSTDMA with k equal to 1 (to maximize capacity). "$\beta_{*} + 1$" slots are allocated for repeater to terminal communication -- a single slot will not suffice for repeater to terminal communication since repeater transmission ranges overlap and if adjacent repeaters transmit to local terminals at the same time, a subset of the terminals will receive transmissions from more than one repeater. Note that there exists a variant of this slot allocation for triangular backbones (see the section describing repeater to terminal transmissions).

Thus, TREE/TDMA protocols involve three types of communication embedded in a TDMA frame: terminal to repeater communication, repeater to repeater communication, and repeater to terminal communication. Each of these is described subsequently.

3.2.1. Terminal to repeater communication

TMA is used by terminals to communicate with repeaters in one of the TDMA slots in each frame. WLG we assume that slot T_0 is used. Each terminal is paired with the closest repeater and transmits with only enough power to reach that repeater. Each repeater and its surrounding terminals form a unit which operate there own local version of TMA during slot T_0 independent of all other units. TMA is describe for a single unit.

While all transmissions within T_0 are from terminals to repeaters, control or feedback information is transmitted from repeaters to terminals during the slots allocated for repeater to terminal communication. No more than two bits of feedback data must be transmitted on the reverse link and collisions cannot occur on the reverse link.

TMA may be considered to be either a centralized version of Capetanakis' adaptive tree protocol or an adaptive "probe-poll" protocol. The operation of TMA can be envisioned as the searching of an imaginary i-ary tree which is superimposed on all terminals. Each vertex of the tree corresponds to a subset of terminals that may transmit during the next T_0 if that vertex is scanned -- the root corresponds to the set of all nodes. The j children of any vertex ($j < i$) are disjoint sets, each containing approximately $1/j$ of the terminals associated with the parent, which are scanned if and only if a collision occurs when their parent is scanned. The structure of the tree, the level at which the search begins, and the searching pattern (typically depth-first or breadth-first) may be varied to optimize the search for terminals having packets to transmit.

In TMA it is assumed that all vertices in the tree are binary except for the root which may have more than two children and that adaptive, pseudo, depth-first search pattern is used starting at a level which minimizes the number of slots needed to search the tree. It is assumed that whenever a terminal has a packet to transmit, it waits until the current tree has been expanded and a new tree expansion begins, whence it transmits the first packet in its queue whenever a vertex of which it is a member is scanned, and deletes that packet from its queue only if no collision occurs. These assumptions are similar to those made by Capetanakis and allow us to use his results pertaining to delay [Cap79b]. They also show imply that TMA becomes TDMA with a capacity of 1.0 under heavy load.

3.2.2. Repeater to repeater communication

Repeaters use spatial TDMA to communicate with one another. If the weak law is the basis of operation and the transmission range of repeaters is chosen to maximize the capacity of the backbone, then "$2\alpha_* + 1$" slots in each frame of TREE/TDMA are allocated for repeater to repeater communication using WSTDMA ($k = 1$). There exists a variant of this slot allocation for triangular backbones (see the following section).

3.2.3. Repeater to terminal communication

If the weak law is the basis of operation for TREE/TDMA and the transmission range of repeaters is chosen to maximize the capacity of the backbone, then each repeater is assigned one of "$\beta_* + 1$" slots which are reserved for repeater to terminal communication in each TREE/TDMA frame. During its assigned slot, a repeater may transmit to its surrounding terminals without interference from other repeaters. A single slot will not suffice for all repeater to terminal transmissions since repeater transmission ranges overlap and if adjacent repeaters transmit to local terminals at the same time, a subset of the terminals will receive transmissions from more than one repeater. Slot assignment for repeaters for the regular networks is depicted in Figures 3.2 through 3.4. In these figures it is assumed that slots $T_1, \ldots, T_{\beta_* + 1}$ are reserved for repeater to terminal communication and that the repeaters indicated by the following symbols are assigned to the slots indicated: $\Delta \equiv 1$, $\bigcirc \equiv 2$, and $\square \equiv 3$.

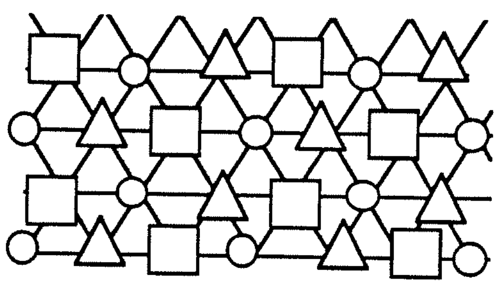


Figure 3.2

Slot allocation -- triangular backbone

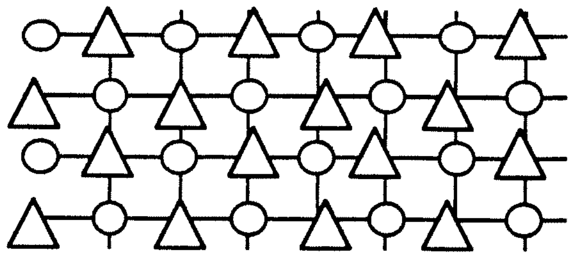


Figure 3.3

Slot allocation -- square backbone

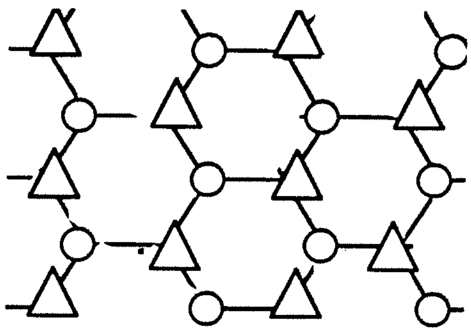


Figure 3.4

Slot allocation -- hexagonal backbone

There exists a variant of TREE/TDMA which may be used with a triangular backbone. This variant is of interest because of its performance characteristics. In this variant a single slot is allocated for terminal to repeater communication and nine slots are used for both repeater to repeater communication and for repeater to terminal communication. Figure 3.5 illustrates the assignment of repeaters to these nine slots -- repeaters with primary index i may transmit to adjacent repeaters in slot i and repeaters with secondary index j may transmit to local terminals in slot j and one additional slot.

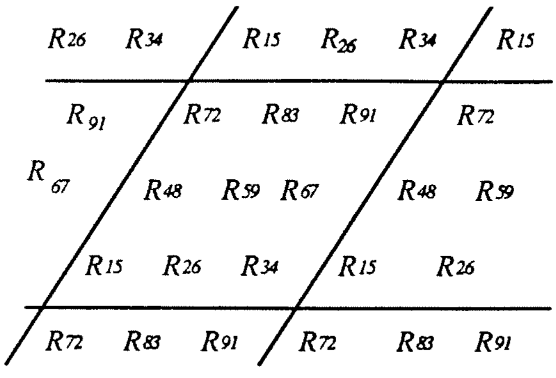


Figure 3.5

Slot allocation (variant) -- triangular backbone

Note that this assignment reduces the total number of slots required for TREE/TDMA using triangular backbones and the weak law from eleven to ten by allowing repeaters to "steal" slots for communicating with surrounding terminals from the slots reserved for repeater to repeater communication.

An example of concurrent repeater to repeater and repeater to terminal transmissions is depicted in Figure 3.6. In this example set C_1 contains repeaters with primary index 1, set C_2 contains repeaters with primary index 2, etc. It shows that if repeaters in C_1 allowed to transmit to adjacent repeaters, then repeaters in C_5 and C_9 can, at the same time, transmit to local terminals without causing interference.

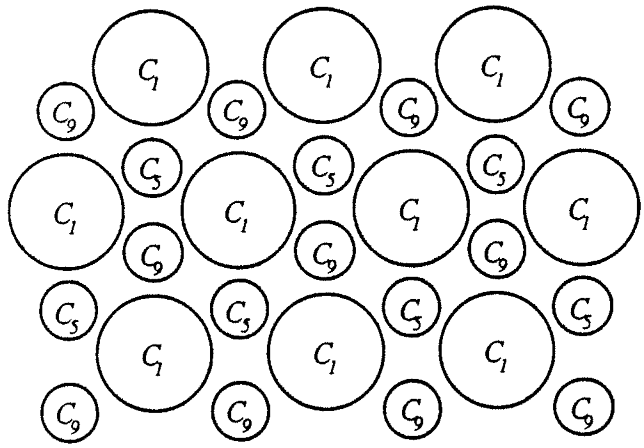


Figure 3.6

Cycle stealing with a triangular backbone

3.3. Analytical model

The analysis of the operating characteristics of networks using TREE/TDMA in subsequent sections is based on the model of a repeater depicted in Figure 3.7. The model is composed of a network of three queues. The internal queue (IQ) and its associated server represent the pool of local terminals awaiting to transmit packets to the repeater over their TREE link and the service provided by the TREE link. The sink queue (SQ) and its associated server represent the queue of packets that have arrived at the repeater and are awaiting transmission to a local terminal in one of the TDMA slots. The propagation queue (PQ) and its associated server represent the queue of packets that have arrived at the repeater and are awaiting to be forwarded to another repeater via the TDMA link.

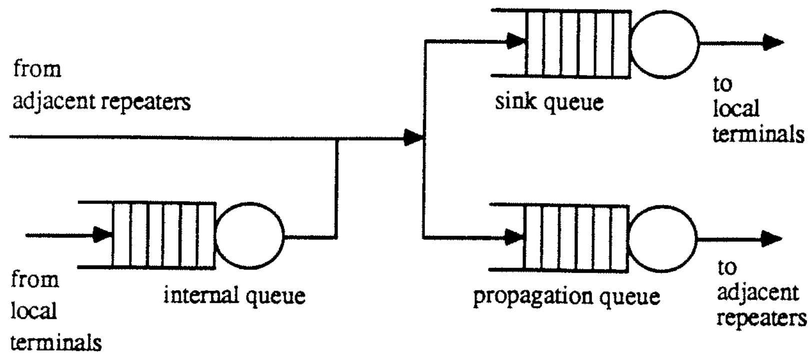


Figure 3.7

Queuing model of a repeater

The analysis is also premised on the following assumptions.

- Packets which arrive at a repeater from local terminals are independent of arrivals at other repeaters and the number of terminals wishing to transmit to a repeater has a Poisson distribution with parameter NP where N is the average number of terminals associated with any repeater and P is the probability that a terminal wishes to transmit to a repeater during a frame, i.e., the probability that i terminals wish to transmit to their repeater during a frame is $(NP)^i e^{-NP}/i!$. If $p_\bullet \triangleq P/\delta_\bullet$, then the mean arrival rate of packets to repeaters is $Np_\bullet$ packets per slot.
- Packets are of constant length and can be transmitted in a single slot.

- The transmission range of each repeater is fixed so that it just reaches adjacent repeaters (1-distant), the distance between adjacent repeaters is r or r_*, any of the n repeaters within m hops of the repeater associated with the source are equally likely to be the destination of data, and a most forward progress routing algorithm is used by each repeater. These assumptions and those that precede these imply that all repeaters have the same average internal arrival rate of packets and the traffic pattern is isotropic and uniform throughout the network.
- Each frame contains a random permutation of the slots $1, \cdots, 2\alpha_* + \beta_* + 3$. Note that this assumption only influences delay calculations: it ensures that paths of the same lengths have the same mean delay.

3.4. Performance -- stability

In this section, we analyze the conditions under which the TREE/TDMA protocols are stable. The network of queues forming our analytical model is stable if each of the queues is stable. A necessary and sufficient condition for a G/G/1 queue to be stable is that the traffic intensity ρ must be less than one. (Traffic intensity is defined to be the ratio of the mean rate at which packets join the queue to the mean rate at which they are removed.)

Since the TREE protocol is adaptive, it becomes a TDMA link under heavy load; hence, the maximum mean service rate of the internal queue is one packet per frame [Cap79a, Cap79b] or $1/\delta_*$ packets per slot. The mean service rates of the sink and propagation queues are also equal to one packet per frame or $1/\delta_*$ packets per slot since they are effectively TDMA links.

The mean arrival rate for the internal queue is NP by assumption. To find the mean arrival rates for the sink and propagation queues we define random variables X, Y, V, and W as follows: X denotes the number of packets which arrive at a repeater from other repeaters during a frame, Y denotes the number of packets which arrive at a repeater from local terminals and join the internal queue during a frame, V denotes the number of packets which join the sink queue during a frame, and W denotes the number of packets which join the propagation queue during a frame. Then $\mathrm{E}(X) = \mathrm{E}(W) = h_{\bullet}\mathrm{E}(Y) = h_{\bullet}\mathrm{E}(V)$ where $h_{\bullet}$ is as defined in Chapter 2, i.e., the mean number of repeater to repeater hops for a packet. This follows from the fact that there is conservation of flow through repeaters and that the traffic pattern is isotropic and uniform in the network.

Thus $\rho_{IQ} = \rho_{SQ} = NP$ and $\rho_{PQ} = h_{\bullet}NP$ and the system of queues is stable if and only if $\max(NP,\, h_{\bullet}NP) < 1$.

3.5. Performance -- capacity

For networks in which $h_{\bullet}$ exceeds one, the nodal capacities of the TREE/TDMA protocols are limited by the capacity of the WSTDMA backbone. The nodal capacities (of all types) for the latter, $S_{w,\bullet}$, were analyzed in Chapter 2 and the network nodal capacities were found to be

$$\frac{0.3712}{\sqrt{n}},\quad \frac{0.4243}{\sqrt{n}},\ \text{and}\ \frac{0.4593}{\sqrt{n}}$$

for large triangular, square, and hexagonal networks, respectively. When normalized to the unit area, these values become

$$\frac{0.4286}{r^2\sqrt{n}}, \quad \frac{0.4243}{r^2\sqrt{n}}, \text{ and } \frac{0.3536}{r^2\sqrt{n}}, \text{ respectively.}$$

These may be interpreted in units of packets per slot for the TREE/TDMA protocols if adjusted to account for the extra slots that appear in TREE/TDMA which do not appear in WSTDMA:

$$S_{t,\bullet} = \frac{(2\alpha_\bullet + 1)S_{w,\bullet}}{\delta_\bullet}.$$

Thus the network nodal capacities for large TREE/TDMA, triangular, square, and hexagonal backbones are in units of packets per slot

$$\frac{0.2362}{\sqrt{n}}, \quad \frac{0.2652}{\sqrt{n}}, \text{ and } \frac{0.2626}{\sqrt{n}},$$

and when normalized to the unit area

$$\frac{0.2727}{r^2\sqrt{n}}, \quad \frac{0.2652}{r^2\sqrt{n}}, \text{ and } \frac{0.2021}{r^2\sqrt{n}}.$$

The variant of TREE/TDMA for triangular backbones has a network nodal capacity and a network nodal capacity per unit area of

$$\frac{0.2598}{\sqrt{n}} \text{ and } \frac{0.3000}{r^2\sqrt{n}}.$$

These results are summarized in the following table. The results pertaining to the standard form of TDMA frame slotting and the alternate form of TDMA slotting for triangular networks are indicated by "I" and "II", respectively.

Nodal Capacity	Triangular (I) Networks	Triangular (II) Networks	Square Networks	Hexagonal Networks
Network	$\frac{0.2362}{\sqrt{n}}$	$\frac{0.2598}{\sqrt{n}}$	$\frac{0.2652}{\sqrt{n}}$	$\frac{0.2626}{\sqrt{n}}$
Network Per Unit Area	$\frac{0.2727}{r^2\sqrt{n}}$	$\frac{0.3000}{r^2\sqrt{n}}$	$\frac{0.2652}{r^2\sqrt{n}}$	$\frac{0.2021}{r^2\sqrt{n}}$

Table 3.1

TREE/WSTDMA capacities for large regular networks

3.6. Performance -- mean packet delay

According to our analytical model, the delay encountered by a packet between the time at which it first arrives at a terminal and the time at which it is eventually received at its destination, having passed through one or more repeaters, is composed of three terms: the delay in the internal queue of the repeater associated with the source, the delay in one or more propagation queues of repeaters on the path followed by the packet, and the delay in the sink queue of the repeater associated with the destination. In subsequent sections the expected values of these terms are calculated and are combined to obtain the total delay. In these calculations it is assumed that the network is heavily loaded.

3.6.1. Packet delay -- internal queue

The delay in an internal queue is, by assumption, equivalent to the delay encounted by a packet transmitted over a TREE link. Capetanakis analyzed the delay in single-hop networks using an adaptive tree protocol as a function of NP via computer simulation [Cap79b]. He also noted that since the adaptive TREE protocol becomes a TDMA protocol when heavily loaded, the expected delay of the former never exceeds that of the latter. Thus an upper bound on the delay on the TREE link is the expected delay of a TDMA link. This delay in units of frames is

$$E(D_{IQ}) \leq \frac{N}{2} + 1 + \frac{NP^2}{2(1 - NP)} \quad \text{[Cap79b]}.$$

For networks in which $NP \leq 0.6$ (note that this must be the case if $h_{\bullet} \geq 5/3$ and the network is stable) and $N \leq 128$, a better bound on the expected delay in units of frames is $E(D_{IQ}) \leq 160NP + 4$.

3.8.2. Packet delay -- propagation queue

The delay encountered by a packet in the propagation queue is the sum of its service time, i.e., the time needed to transmit the packet once it reaches the head of the queue, and the time that it spends waiting in the queue before it reaches the head of the queue. The queuing time can be found by evaluating two generating functions $A(z)$ and $A'(z)$ associated with the discrete semi-Markov process of the propagation queue, and the random variables described subsequently.

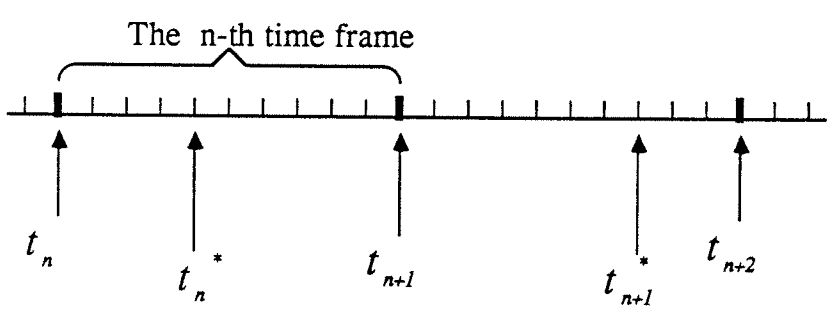


Figure 3.8

Time in the semi-Markov process

$t_n \quad \triangleq$ beginning of the n-th frame

$t_n^* \quad \triangleq$ beginning of enabled slot in PQ in n-th frame

$X_n \quad \triangleq$ number of packets in PQ at t_n

$X_n^* \quad \triangleq$ number of packets in PQ at t_n^*

$Y_n \quad \triangleq$ number of packets arriving to PQ in $(t_{n-1}, t_n]$

$U_n \quad \triangleq$ number of packets arriving to PQ in (t_{n-1}, t_{n-1}^*)

$A_n(k) \quad \triangleq \text{Prob}[X_n = k]$

$A_n^*(k) \quad \triangleq \text{Prob}[X_n^* = k]$

$A_n \quad \triangleq [A_n(1) \;\; A_n(2) \;\; A_n(3) \quad \cdots]$

$A_n^* \quad \triangleq [A_n^*(1) \;\; A_n^*(2) \;\; A_n^*(3) \quad \cdots]$

$A \quad \triangleq \lim_{n\to\infty} A_n$

$A^* \quad \triangleq \lim_{n\to\infty} A_n^*$

$A(z) \quad \triangleq$ generating function of $\lim_{n\to\infty} X_n$

$A^*(z) \quad \triangleq$ generating function of $\lim_{n\to\infty} X_n^*$

Now $A(z)$ can be found by embedding Markov chain X_n in the discrete semi-Markov process. The Markov property of X_n is expressed as follows:

$$X_{n+1} = X_n + Y_n - 1 \qquad \text{if } X_n \neq 0, \text{ and}$$

$$X_{n+1} = Y_n - I_{\{U_n > 0\}} \qquad \text{otherwise,}$$

where $I_{\{U_n > 0\}} = 1$ if $U_n > 0$ and $I_{\{U_n > 0\}} = 0$ otherwise. The embedded Markov chain has the one step transition matrix

$$P = \begin{bmatrix} g_0 & g_1 & g_2 & \cdot \\ f_0 & f_1 & f_2 & \cdot \\ 0 & f_0 & f_1 & \cdot \\ 0 & 0 & f_0 & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{bmatrix}$$

where $f_k = \text{Prob}[Y_n = k]$ and

$$g_k = \text{Prob}[Y_n = k]\text{Prob}[U_n = 0] + \text{Prob}[Y_n = k+1]\text{Prob}[Y_n \neq 0].$$

(Both f_k and g_k are described in appendix B.) Noting that $A = AP$ and taking the Z-transform we find that (see Appendix C)

$$A(z) = \frac{1 - m_f}{1 + m_g - m_f} \frac{zG(z) - F(z)}{z - F(z)}$$

where $A(z) \triangleq \sum_{k=0}^{\infty} A_k z^k$, $G(z) \triangleq \sum_{k=0}^{\infty} g_k z^k$, $F(z) \triangleq \sum_{k=0}^{\infty} f_k z^k$, $m_f \triangleq \sum_{k=0}^{\infty} k f_k$, and $m_g \triangleq \sum_{k=0}^{\infty} k g_k$.

$A^*(z)$ can be derived from $A(z)$. As defined above, $A^*(z)$ is the generating function of $\lim_{n \to \infty} X_n^*$ and $A(z)$ is the generating function of $\lim_{n \to \infty} X_n$. The random variable X_n^* is the sum of the random variables X_{n-1} and U_n. Since the generating function of two independent random variables is the product of the two generating functions, $A^*(z) = A(z)\, U(z)$ where $U(z)$ is the generating function of $\lim_{n \to \infty} U_n$ (see appendix B).

Since the service rate of the propagation queue is one packet per frame, the expected delay in the propagation queue equals the expected number of packets which are in the propagation queue when a randomly chosen packet arrives [Kle75b]. This turns out to be difficult to evaluate; hence, we find an equivalent value, the expected number of packets in the propagation queue after a randomly chosen packet is sent [Kle75b]. If $\pi(k)$ is the probability that k packets are in the propagation queue after a randomly chosen packet is sent, then

$$\pi(k) = \mathrm{Prob}[\lim_{n\to\infty} X_n^* = k-1 \mid \lim_{n\to\infty} X_n^* > 0] = \frac{A^*(k+1)}{1 - A^*(0)}.$$

and the generating function of $\pi(k)$ is

$$\Pi_k(z) = \frac{(A^*(z) - A^*(0))z^{-1}}{1 - A^*(0)} = \frac{(A(z)U(z) - A(0)U(0))z^{-1}}{1 - A(0)U(0)}$$

for $A(z)$ and $U(z)$ as derived before.

The mean number of packets in the propagation queue after a packet is sent is $\lim_{z\to 1}\frac{d\,\Pi_k(z)}{dz}$. For the reasons mentioned above, this value is the expected delay in the propagation queue. Thus we find that the total delay in the propagation queue is

$$\mathrm{E}(D_{PQ}) = 1 + \lim_{z\to 1}\frac{d\,\Pi_k(z)}{dz}.$$

which can be evaluated using numerical techniques.

3.6.3. Packet delay -- sink queue

The total delay in the sink queue can be found by the same method we used in the last section. The only major difference is that the mean arrival rate is equal to NP packets per frame rather than $h_* NP$ packets per frame.

3.6.4. Packet delay -- total

When a TREE/TDMA network is heavily loaded, a packet is never forwarded more than once in any frame; hence, the total expected packet delay is the sum of the expected delay in the internal queue; plus h_* times the expected delay in the propagation queue; plus the expected delay in the sink queue. When $h_* >> 1$ then the total expected packet delay is effectively the delay encounted by a packet in being forwarded from repeater to repeater.

The total expected packet delay for TREE/TDMA in which a triangular backbone is used and the framing employs cycle stealing is plotted against N and the packet arrival rate P in Figures 3.9 and 3.10 respectively. The values were obtained by numerical techniques.

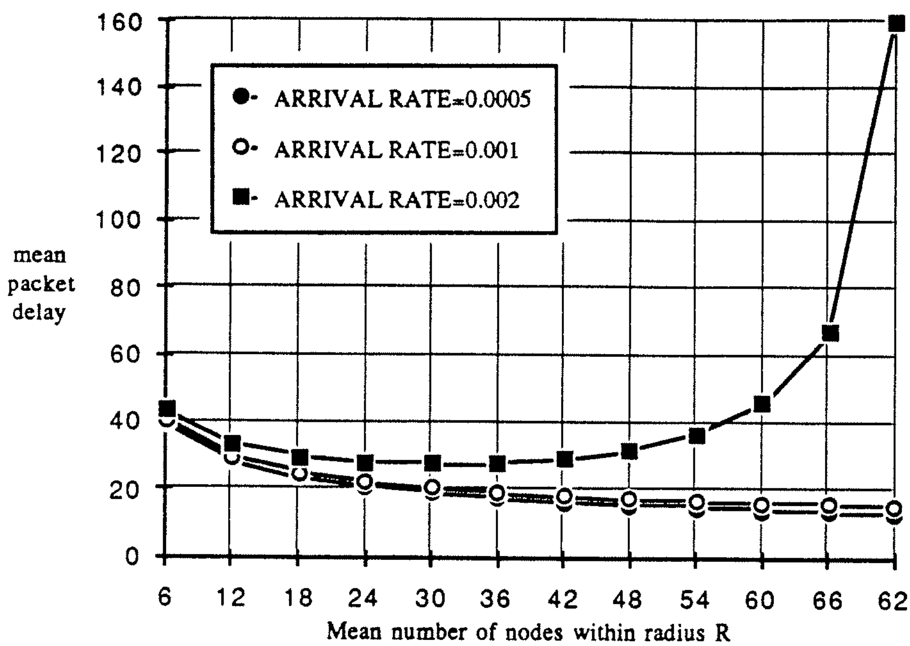


Figure 3.9

TREE/TDMA -- mean delay versus N

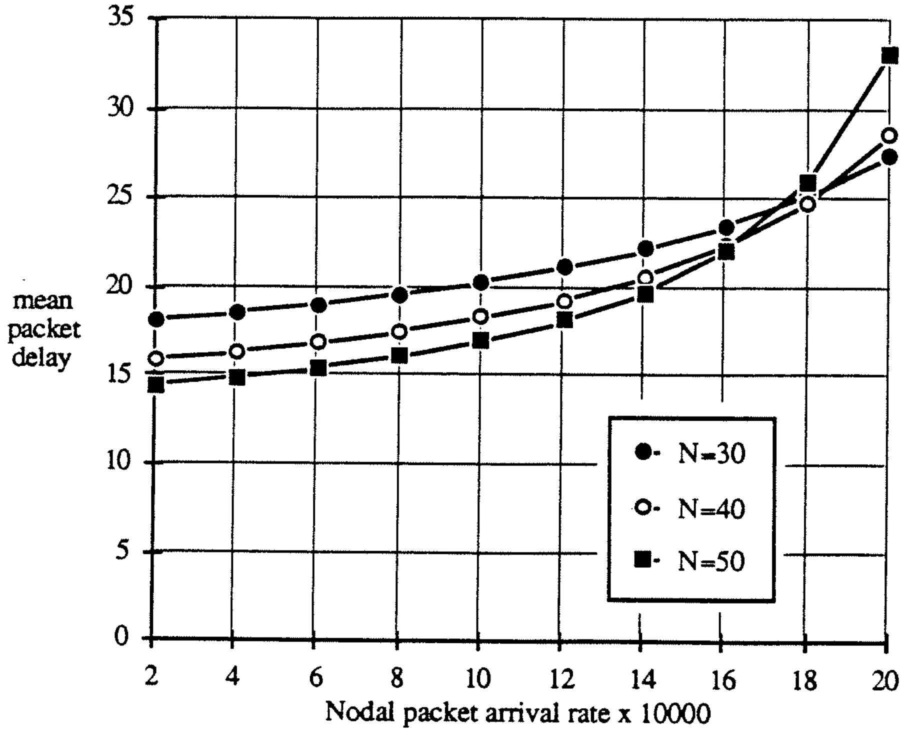


Figure 3.10

TREE/TDMA -- mean delay versus packet arrival rate

CHAPTER 4

CS/CAMA CHANNEL ACCESS PROTOCOLS

4.1. Introduction

In this chapter a new channel access protocol is introduced and analytical models describing the operating characteristics of the protocol are presented. The protocol is a carrier sense, collision avoidance multiple access (CS/CAMA) protocol and operates under the strong law of conflict-free operation described in Chapter 2. Like the TREE/TDMA protocols, CS/CAMA can be used in irregular or random planar networks. Unlike many existing random access protocols for packet radio networks, CS/CAMA creates a conflict-free environment without sacrificing potential spatial reuse of the communication channel.

We first describe the CS/CAMA protocol and illustrate its operation. We then show that CS/CAMA is essentially conflict-free, i.e, that the transmissions of two nodes will never interfere with one another in such a way that the transmissions will not reach their intended destinations. Using Petri nets we next show that the protocol operates correctly, e.g., is free of deadlocks. We then present an analytical model of the protocol and investigate its operating characteristics. Finally we present a simulation model of the protocol and compare the operating characteristics of the protocol found using the simulation model to those obtained with the analytical model.

4.2. Description

In CS/CAMA a packet is transmitted from its source (station) to a nearby station; then from station to station until it reaches its destination. A most forward routing algorithm is used by each station and each station operates independently of all other stations.

The communication channel is shared by all nodes and is not slotted as in the TREE/TDMA protocols. The channel is split into two separate channels. On one channel, the main channel, packets and request signals (REQs) are transmitted. On the other channel, the subchannel, acknowledge and release signals (ACKs and RELs) are transmitted. While both packets and REQs are addressed to a specific station, ACKs and RELs are not addressed to a specific station and bear no additional information other than their existence. (The bandwidth required on the subchannel is very small when compared to that required on the main channel: only the existence of ACKs and RELs need be established.) A station can transmit on either channel and listen to the other channel simultaneously, but cannot both transmit and receive a packet on the same channel.

Before using the channel, a station must check two status indicators which it maintains: the channel status (CS) indicator and the station status (SS) indicator. CS provides information about the main channel in the vicinity of a station. CS is "busy" if the main channel is busy, otherwise, it is "free". (If CS is "free", then a station can receive a transmission.) SS provides information about the possible effect of the transmission of a station on other stations in its transmission range. SS depends on a counter, the SS counter (SSC), maintained at each station. The value of SSC indicates the number of stations that would experience interference if the station were to transmit. SS is defined to be "permit" if SSC is 0 and "prohibit" otherwise.

The two status indicators together provide sufficient information about the state of the network in the area surrounding a station to schedule transmissions without conflict according to the strong law described in Chapter 2. Table 4.1 illustrates what stations can do as a function of CS and SS.

CHANNEL STATUS	STATION STATUS	CANDIDATE FOR RECEIVER	CANDIDATE FOR TRANSMITTER
Busy	Prohibit	No	No
Busy	Permit	No	Yes
Free	Prohibit	Yes	No
Free	Permit	Yes	Yes

Table 4.1

CS/CAMA station status indicators

In CS/CAMA each station obeys the following rules.

(Rule 1) Execute the following when a station wishes to transmit a packet.

(1.1) Repeat (1.1) until its SS is "permit" and then proceed to (1.2).

(1.2) Send REQ on the main channel (REQ identifies the receiver to which the packet is to be sent). If REQ is acknowledged within $2\tau + t_m$ to $4\tau + t_m$ seconds and if SS is still "permit" at that time, then transmit the packet immediately and set SSC to SSC + 1; otherwise, revert to (1.1). Note that τ is the maximum propagation delay between a station and another station within its transmitting range and t_m is the duration of a REQ plus the response time of the station.

(Rule 2) Execute the following when REQ arrives at a station and is addressed to that station.

If CS is "free" from 2τ seconds before receiving REQ to $2\tau + t_m$ seconds after receiving REQ, then send ACK on the subchannel and await a packet. Send REL on the subchannel if a packet is received.

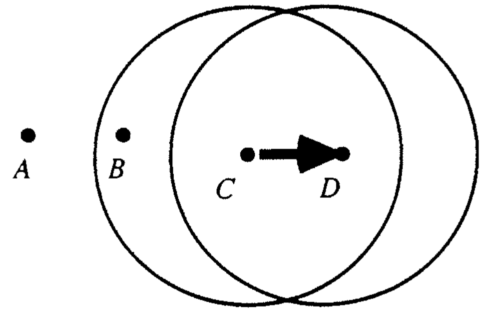
(Rule 3) Execute the following when ACK appears at a station on the subchannel.

Set SSC to SSC + 1.

(Rule 4) Execute the following when REL appears at a station on the subchannel or if SSC is not 0 and has not been decremented for T_m seconds after it was last incremented (T_m is the time needed to transmit a packet of maximum length including propagation delays), then

Set SSC to SSC - 1.

The following two examples illustrate the operation of the protocol. The first example is depicted in Figure 4.1. Station C is assumed to be transmitting to D and A desires to transmit to B. Since SS of A is "permit", REQ is transmitted by A on the main channel. Since CS of B is "busy", B will not ACK and A will wait. This decision is correct since B may not receive a transmission of A correctly even if A were to transmit.

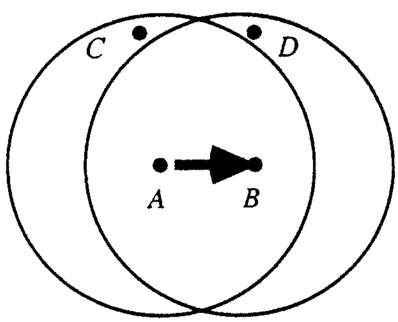


	Channel status	Station status
Station A	Free	Permit
Station B	Busy	Permit

Figure 4.1

Example 1 -- operation of CS/CAMA

The second example is depicted in Figure 4.2. Station A is assumed to be transmitting to B and C desires to transmit to D. According to the rules, C will transmit and no collision will occur. This decision is also correct.

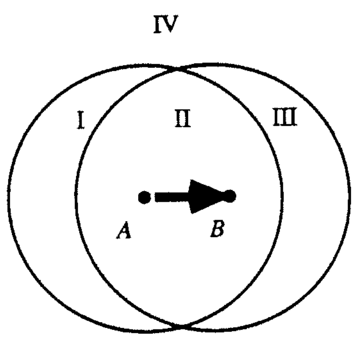


	Channel Status	Station Status
Station C	Busy	Permit
Station D	Free	Prohibit

Figure 4.2

Example 2 -- operation of CS/CAMA

The actions taken by a station using CS/CAMA for various states are summarized in the following figure and table. In this figure and table we have partitioned the region surrounding stations A and B into four disjoint regions, region I through IV. It is assumed that A is transmitting to B and C desires to transmit to D. C can be in any one of the four regions as can D. The table shows the decisions made by C based on its location and the location of D. It indicates that whenever C transmits to D, the transmission is successful and does not interfere with the transmission from A to B. Whenever C defers, it does so for one of two reasons: to avoid interference (the transmission would interfere with the ongoing transmission); to improve efficiency (an ongoing transmission may interfere with the transmission).



A is transmitting to B.

Figure 4.3

Environment of a transmission

Transmitter Receiver	I (Busy, Permit)	II (Busy, Prohibit)	III (Free, Prohibit)	IV (Free, Permit)
I (Busy, Permit)	Wait inefficiency	Wait interference	Wait interference	Wait inefficiency
II (Busy, Prohibit)	Wait inefficiency	Wait interference	Wait interference	Wait inefficiency
III (Free, Prohibit)	Transmit	Wait interference	Wait interference	Transmit
IV (Free, Permit)	Transmit	Wait interference	Wait interference	Transmit

Table 4.2

CS/CAMA states and decisions

4.3. Analysis of conflicts

In this section we show that the protocol is essentially free of conflict. In the subsequent proof it is assumed that both station T_1 and T_2 in Figure 4.4 are in the "permit" state and that stations T_1 and T_2 desire to transmit to stations R_1 and R_2, respectively. In this case, only one of the two stations can be given permission to transmit if a collision is to be avoided and, furthermore, exactly one of the two stations must be given permission unless we are to sacrifice spatial reuse of the channel.

We first show that collisions are impossible on the main channel if collisions do not occur on the subchannel. Then, we show that even if collisions occur on the subchannel, it is unlikely that collisions will occur on the main channel.

Notation:

Types of signals

req_i -- REQ from transmitter T_i

ack_i -- ACK from receiver R_i

pac_i -- packet from transmitter T_i

Time

$t_{i,j}$ -- propagation time between T_i and R_j

t_{req_i} -- time at which req_i signal is transmitted

t_{j,req_i} -- time when receiver R_j receives req_i

t_{j,ack_i} -- time when transmitter T_j receives ack_i

t_{j,pac_i} -- time when receiver T_j receives pac_i

t_m -- duration of the REQ plus station response time

τ -- maximum propagation delay within transmission range

(We reset time to 0 when req_1 is transmitted.)

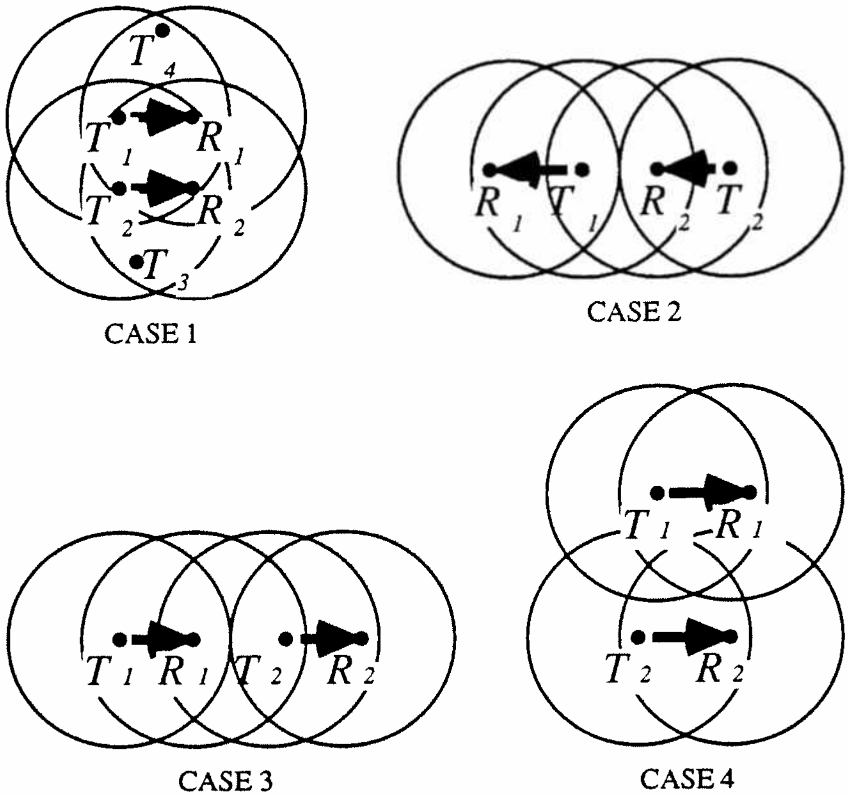


Figure 4.4

Cases illustrated for discussion

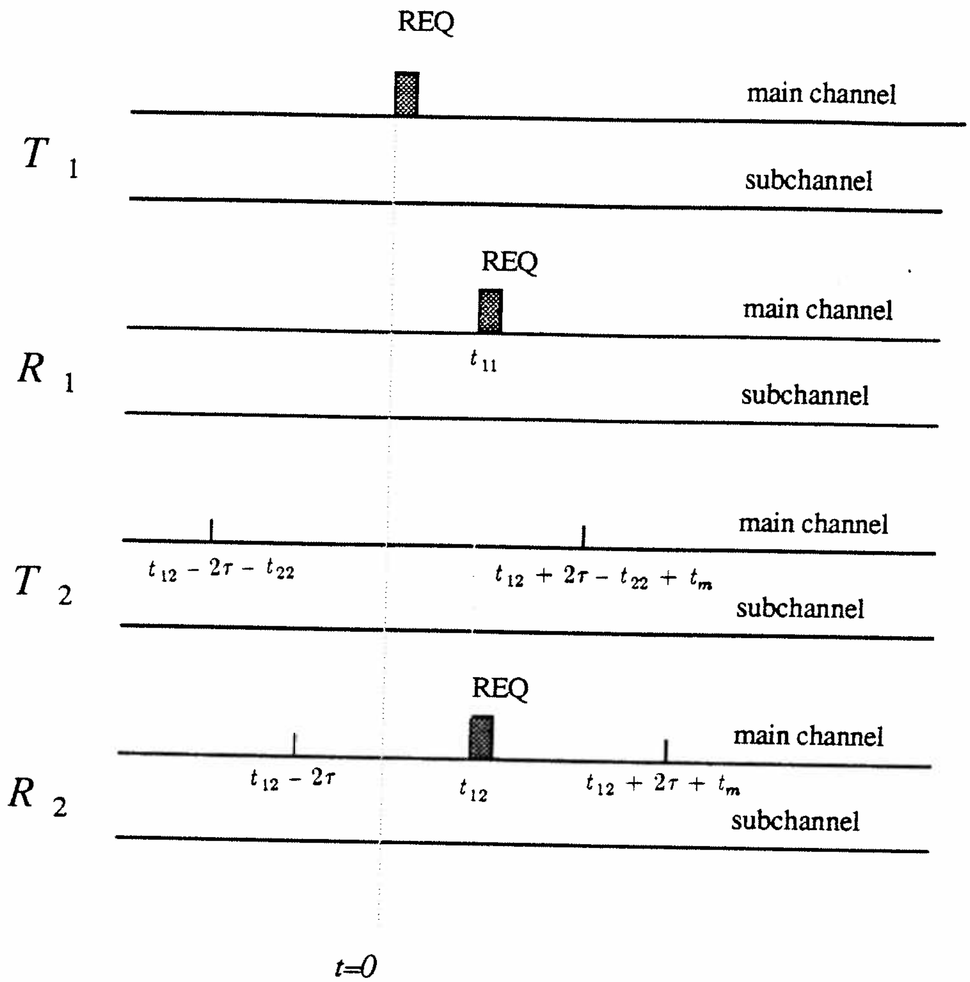


Figure 4.5

Time diagram for analysis

CASE 1: t_{11}, t_{12}, t_{21}, t_{22} are less than τ.

If $t_{req_2} < t_{12} - 2\tau - t_{22}$, then $t_{2,req_2} = t_{req_2} + t_{22}$

If ack_2 is not transmitted (station T_3 is transmitting), then no collision occurs.

If ack_2 is transmitted, then $t_{ack_2} = t_{2,req_2} + 2\tau + t_m < t_{12} + t_m$

$t_{1,ack_2} = t_{ack_2} + t_{12} < 2t_{12} + t_m < 2\tau + t_m$

Since $t_{1,ack_2} < 2\tau + t_m$, SS of T_1 will be set to "prohibit" before $2\tau + t_m$ seconds, according to rule 1.2, transmitter T_1 will refrain from transmission. No collisions occur.

If t_{req_2} is in $[\ t_{12} - 2\tau - t_{22},\ t_{12} + 2\tau - t_{22} + t_m\ ]$, then $t_{2,req_2} = t_{req_2} + t_{22}$ and t_{2,req_2} is in the interval $[t_{12} - 2\tau,\ t_{12} + 2\tau + t_m]$

Since $t_{2,req_1} = t_{12}$, we have $-2\tau < t_{2,req_2} - t_{2,req_1} < 2\tau + t_m$

Two REQs are received by station R_2. Hence, station R_2 will not issue an ACK according to rule 2. T_2 will refrain from transmission and no collisions occur.

If $t_{req_2} > t_{12} + 2\tau - t_{22} + t_m$

If ack_1 is not transmitted (T_4 is transmitting), then no collisions occur.

If ack_1 is transmitted, then $t_{ack_1} = t_{1,req_1} + 2\tau + t_m = t_{11} + 2\tau + t_m$

$t_{1,ack_1} = t_{ack_1} + t_{11} = 2t_{11} + 2\tau + t_m$

$t_{pac_1} < t_{1,ack_1} + t_m = 2t_{11} + 2\tau + 2t_m$

(We assume the response time to send a packet is shorter than t_m.)

$t_{2,pac_1} = t_{pac_1} + t_{12} < 2t_{11} + 2\tau + t_{12} + 2t_m$

$t_{2,req_2} = t_{req_2} + t_{22} > t_{12} + 2\tau + t_m$

t_{2,pac_1} - $t_{2,req_2} < 2t_{11} + t_m \ = 2\tau + t_m$

t_{2,pac_1} - $t_{2,req_2} < 2\tau + t_m$

The packet transmitted from station T_1 arrives at station R_2 within $2\tau + t_m$ seconds after req_2 signal was received. According to rule 2, ACK is not issued by R_2. Thus station T_1 will transmit and T_2 will wait. No collisions occur.

CASE 2: t_{21} is undefined.

This case occurs when station R_1 is outside the transmission range of T_1. Since we did not use t_{21} in the proof of case one; hence, the same proof applies.

CASE 3: t_{12} is undefined.

The same argument as in CASE 2 applies.

CASE 4: Both t_{12} and t_{21} are undefined.

Since both R_1 and R_2 are outside the transmission range of T_2 and T_1. There will be no collision even if both transmitters are active.

We have proved that packets will not collide if signals on the subchannel are received correctly. We now examine cases in which signals overlap on the subchannel.

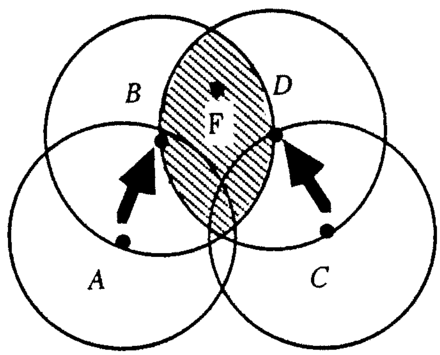


Figure 4.6

Signals overlapping on the subchannel

CASE 5: Suppose ACKs from stations B and D in Figure 4.6 overlap. Stations in the shaded area may detect one ACK rather than two.

(a) If the stations which have detected but one ACK have nothing to transmit the error will be detected and corrected after these stations receive two RELs.

(b) If the stations which have detected but one ACK attempt to transmit after receiving one REL, a collision may occur on the main channel.

CASE 6: Suppose RELs from stations B and D overlap. Station F may detect one REL rather than two. Thus SSC is decremented by one. Rule 4 contains an auto-reduction rule which corrects the problem.

CASE 7: Suppose REQs from stations B and D overlap. Since REQs are not primitive signals on the subchannel but are sent on the main channel as packets, it is assumed that they interfere and will be ignored.

4.4. Petri net model of correctness

In the following section, we use a Petri net to model the CS/CAMA protocol and show that the protocol is correct and free of deadlock. In Figure 4.7 the flow chart of the CS/CAMA protocol is depicted. The associated Petri net along with its time table and predicate table are shown in Figure 4.8. In Figure 4.9 we link two of the Petri nets by a channel which is also represented as a Petri net.

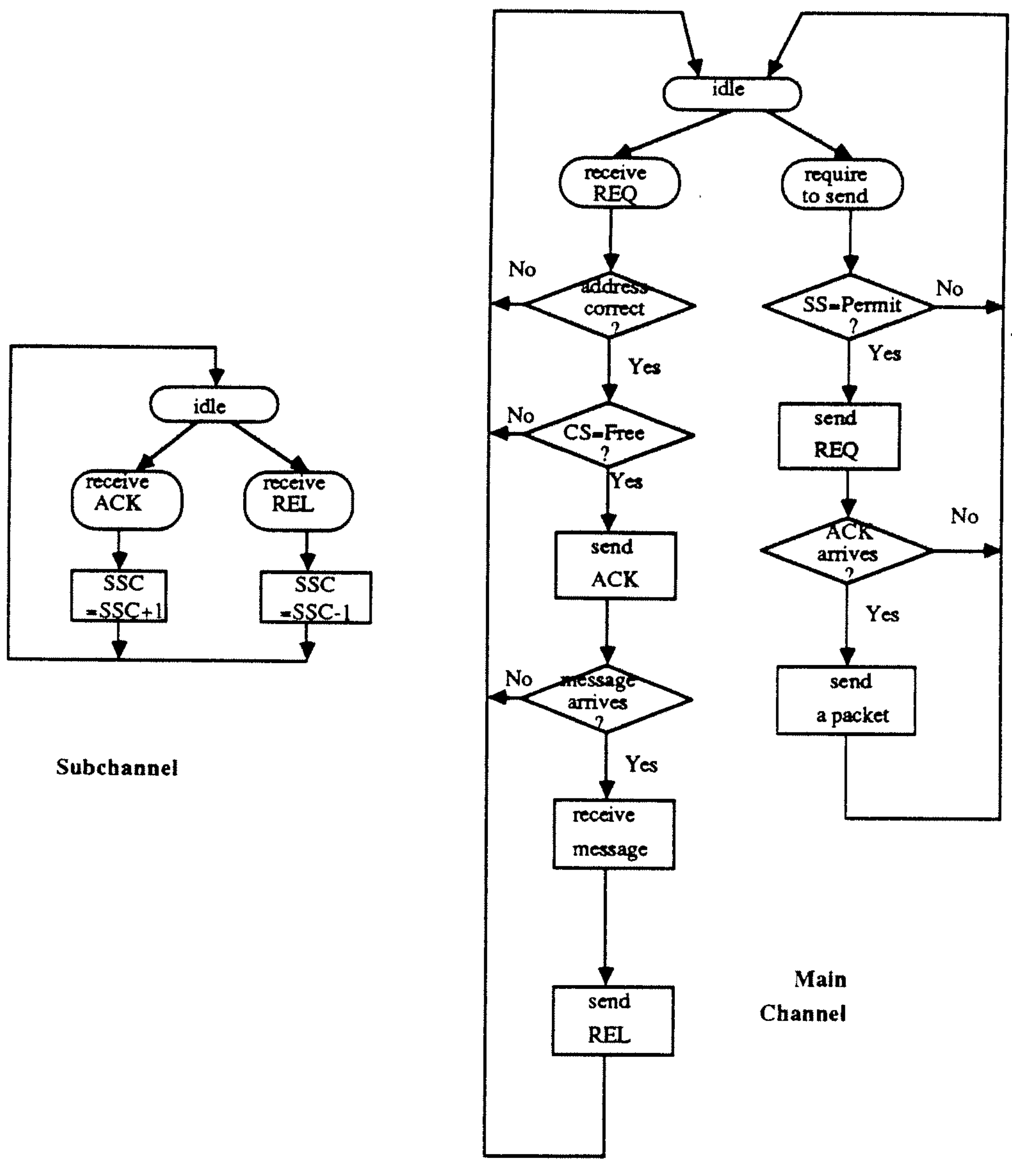


Figure 4.7

Flow chart of the CS/CAMA protocol

Representation of states

PLACE	STATE
p_1	ACK arrives
p_2	REL arrives
p_3	Subchannel is idle
p_4	Content of SSC
p_5	Main channel is idle
p_6	REQ arrives
p_7	(used to model a free space channel)
p_8	REQ is sent
p_9	(used to model a free space channel)
p_{10}	ACK arrives
p_{11}	Waiting state
p_{12}	Checking address
p_{13}	ACK is sent
p_{14}	A packet is sent
p_{15}	Waiting state
p_{16}	Sending state
p_{17}	A packet arrives
p_{18}	(used to model a free space channel)
p_{19}	REL is sent

Transition time

Transition	T_{min}	T_{max}
t_1	0	0
t_2	0	0
t_3	0	0
t_4	0	0
t_5	0	τ
t_6	0	τ
t_7	$4\tau+t_m$	$4\tau+t_m$
t_8	$4\tau+t_m+\epsilon$	$4\tau+t_m+\epsilon$
t_9	$2\tau+t_m$	$2\tau+t_m$
t_{10}	$2\tau+t_m$	$2\tau+t_m$
t_{11}	0	max packet transmission time
t_{12}	2τ	2τ
t_{13}	0	2τ
t_{14}	0	τ

Predicate table

Transition	Predicate
t_3	Station in sending state and SS is permit
t_7	REQ is granted
t_9	Station can not receive the packet
t_{10}	The predicate of t_9 is not true
t_{11}	Transmission process is completed

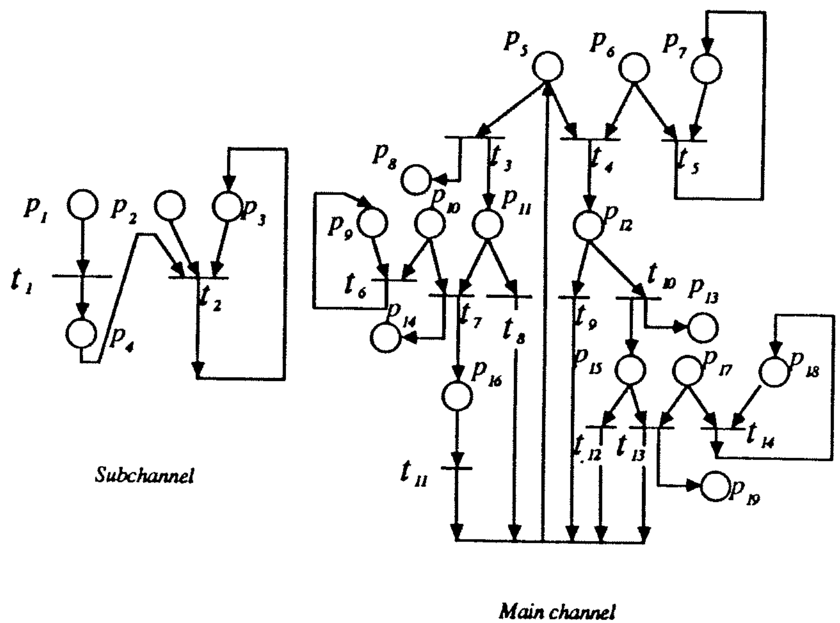


Figure 4.8

Petri net model of a CS/CAMA station

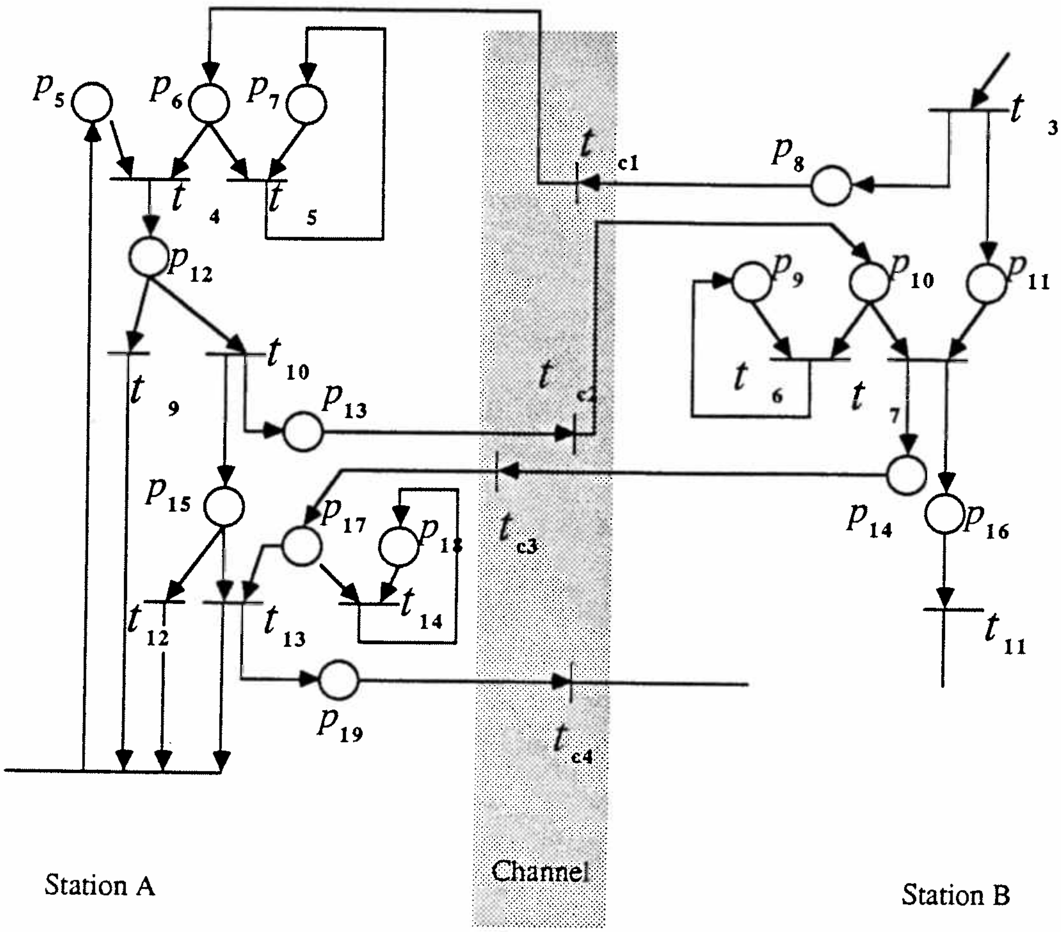


Figure 4.9

Two stations linked CS/CAMA

We now show that the system depicted in Figure 4.8 is "bounded" and "live"; hence, that the CS/CAMA protocol can be implemented and is free of deadlock. A Petri net is bounded if and only if every place of the net is bounded. Thirteen places of the Petri net in Figure 4.8 are bounded as follows.

- $M(p_3) = M(p_7) = M(p_9) = M(p_{18}) = 1$
- $M(p_5)+M(p_{11})+M(p_{12})+M(p_{15})+M(p_{16}) = 1$
- $M(p_8) \leq 1$ since the maximum transition time of t_{c1} (in Figure 4.8) is less than the sum of minimum transition times of t_3 and t_8 or t_5, t_{11}, and t_{11}
- $M(p_{13}) \leq 1$ since the maximum transition time of t_{c2} (in Figure 4.9) is less than the sum of minimum transition times of t_4, t_{12}, and t_{13}, or t_4, t_{12}, and t_{14}
- $M(p_{14}) \leq 1$ since $M(p_{16})$ is bounded and $M(p_{14}) \leq M(p_{16})$ because the predicate of t_{11} is $M(p_{14}) = 0$
- $M(p_{19}) \leq 1$ since the maximum transition time of t_{c4} is less than the sum of the minimum transition time of t_4, t_{10}, and t_{12}

The remaining places are places whose tokens represent either an ACK, a REL, a REQ, a message, or the content of the SS counter. These places are also bounded since only a finite number of signals can be heard by a station and the content of the SS counter is finite. We conclude that the CS/CAMA protocl can be implemented.

A Petri net is live if and only if the transitions of the Petri net are live for all markings that can be reached from M_0. We show that the Petri net of the CS/CAMA protocol is live in two steps. First we show that there exists a marking called M' that can be reached from any marking that can be reached from M_0. Second we show that all transitions are live for the marking M' .

The marking M' is defined as follows: $M'(p_5) = 1$, $M'(p_{11}) = 0$, $M'(p_{12}) = 0$, $M'(p_{15}) = 0$, and $M'(p_{16}) = 0$. The fact that it exists can be proved by the place invariant $M'(p_5) + M'(p_{11}) + M'(p_{12}) + M'(p_{15}) + M'(p_{16}) = 1$. This invariant implies that only one of these five places has a token at any time. In addition, none of these places can hold a token forever. For example, a token can not stay at place p_{11} for more than $4\tau + t_m$ seconds since the transition t_8 will always fire if the token remains that long. Thus, marking M' exists.

All transitions of the Petri net are live for M' since for each transition we can find a feasible firing sequence which allows this transition to fire. For example, the transition t_3 will fire if the predicate of t_3 is true, i.e., the station is attempting to send message and this condition will occur; the transition t_7 will fire if $M(p_{10}) = 1$ and $M(p_{11}) = 1$. From Figure 4.9, we know that $M(p_{10}) = 1$ and $M(p_{11}) = 1$ can be reached from M' because t_{min} of t_8 is set to be greater than the sum of t_{max}'s of t_{c1}, t_4, t_{10}, and t_{c2}. All the remaining transitions can be proved to be live by the same method.

Since M' can be reached from every reachable marking of M_0 and all transitions are live for M', we conclude that the Petri net is live and the CS/CAMA protocol is free of deadlock.

4.5. Performance -- analytical models

To facilitate the evaluation of the CS/CAMA protocol, the following assumptions are made.

- The nodal packet arrival process is Poisson with rate λ which consists of the internal packet arrival rate and the external packet arrival rate.
- The time required to transmit a packet is exponentially distributed with mean 1.
- The propagation time delay is short and can be neglected.

- If a packet can not be transmitted at a some point of time, a random amount of time is waited before attempting to transmit the packet again. The waiting time is assumed to be exponentially distributed with mean 1.
- The network traffic is two dimensional isotropic.
- A MFP routing algorithm is used.

4.5.1. Queuing Models

Two M/M/2//1 queues are used to represent the operation of the CS/CAMA protocol. Each one of these M/M/2//1 queues can be translated to an equivalent but simpler M/G/1 queue. By analyzing these two M/G/1 queues, we can obtain the performance of the CS/CAMA protocol.

Model I

Model I, shown in Figure 4.10, is an M/M/2//1 queue, i.e, has a Poisson arrival process, a Poisson service process, two servers of which only one server can work at each time, a FIFO discipline, and infinite buffers. This queuing model will be used to find an upper bound on the mean time delay, a lower bound on the expected throughput, and a lower bound on the capacity of the protocol. This M/M/2//1 queuing model consists of two parts.

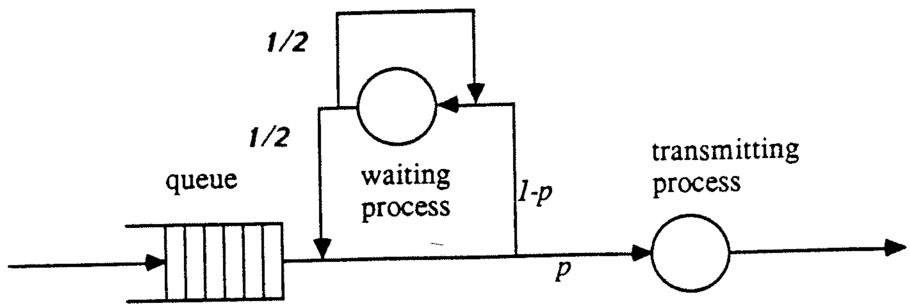


Figure 4.10

Model I of the CS/CAMA protocol

The first part of the M/M/2//1 queue is a feedback queue which represents the waiting process of the CS/CAMA protocol. As operated in the CS/CAMA protocol, an idle station which receives a packet can transmit only if SS of the station is "permit" and CS of the receiver is "free". If p is the probability of such an event, then with probability $1-p$ a station can not transmit on the first try and will wait. Since the waiting time is exponentially distributed with mean 1, we know that with probability 1/2 that no transmission that keeps a station waiting, will finish before a randomly chosen delay which is also exponentially distributed with mean 1. Hence, the waiting process will remain in the feedback queue until at least one such transmissions finishes. When one such transmission finishes before the waiting time, the packet can be transmitted on the next try with probability greater than p. In model I, this probability is taken to be p since this model is used to find the worst performance of the protocol. Therefore, the feedback link of the queuing model has the entrance point indicated in Figure 4.10.

The second part of the M/M/2//1 queue shown in Figure 4.10 represents the transmission process of a station. Since the propagation time delay is much shorter than the packet transmission time, we neglect the propagation time delay in this model.

Combining these two parts, the first M/M/2//1 queuing model is formed. Since its direct analysis is complicated, we transform this M/M/2//1 queuing model to an equivalent but simpler M/G/1 queuing model, whose performance is known [Kle75c]. This M/G/1 queue has a mean arrival rate λ and a service time generating function $B(s) = \frac{p + 2ps}{p + 2s} \frac{1}{s + 1}$ (see Appendix D).

Model II

Model II, shown in Figure 4.11, is also an M/M/2//1 queue. The waiting process of this model is the same as that of Model I. Model II will be used to find a lower bound on the mean time delay, an upper bound on the expected throughput, and an upper bound on the capacity of the CS/CAMA protocol.

By using the same method in Model I, the M/M/2//1 queue can be transformed to an equivalent M/G/1 queue with $B(s) = \frac{(1 - p)}{(s + 1)^2} + \frac{p(s + 1)}{s}$ (see Appendix D).

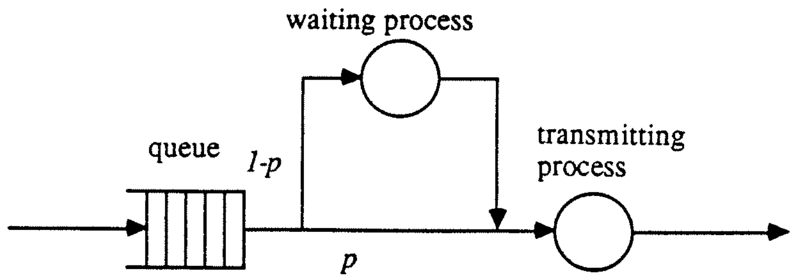


Figure 4.11

Model II of the CS/CAMA protocol

4.5.2. The probability of transmission

As defined before, p is the probability that a station is permitted to send a waiting packet at time t. According to the CS/CAMA protocol, a station with packet is allowed to transmit the packet to a selected receiver if SS of the station is "permit" and CS of the receiver is "free". In the following theorem, we find bounds on the probability of the aforementioned condition and use the result to find an upper bound and a lower bound on the value of p.

Theorem 4.1

Let $A(t)$ be the event that CS of a station is "free" and $B(t)$ be the event that SS of a station is "permit". Then,

$$(1 - \lambda N) \leq \mathrm{Prob}[A(t)] \leq e^{-\lambda N}, \text{ and } (1 - \lambda N) \leq \mathrm{Prob}[B(t)] \leq e^{-\lambda N}.$$

Proof

- The random process $X(t)$ is the number of transmitters within the transmission range of a station. Since the packet arrival process and the packet service process are both Poisson, $X(t)$ is a regenerative process with regenerative points t_i which are the starting points of the i-th busy cycle. From the renewal reward theorem [Ros85], we find

$$\text{Prob}[A(t)] = \text{Prob}[X(t)=0] = \frac{\text{E[length idle period]}}{\text{E[length busy cycle]}}.$$

- Let T_n be the total length of n busy cycles. Then, E[length busy cycle] $= \lim_{n\to\infty} \frac{T_n}{n}$ and E[length idle period] $= \lim_{n\to\infty} \frac{1}{n} \int_0^{T_n} I_{\{X(t)=0\}} dt$ where I is an indicator random variable which is 1 if $X(t)=0$ and is 0 otherwise. Thus $\text{Prob}[A(t)] = \lim_{n\to\infty} \int_0^{T_n} \frac{1}{T_n} I_{\{X(t)=0\}} dt$.

- The counting process $M(t)$ is the number of packets that have been sent out from those stations within the transmission range of the station and

$$\begin{aligned}
\int_0^{T_n} I_{\{X(t)=0\}} dt &= T_n - \int_0^{T_n} I_{\{X(t)>0\}} dt \\
&\geq T_n - \int_0^{T_n} X(t)\, dt \\
&= T_n - \text{E[sum of packets service time up to time } T_n] \\
&= T_n - \text{E[E[total service time of } M(t) \text{ packets} \mid M(t)]] \\
&= T_n - N\lambda T_n .
\end{aligned}$$

Hence, $\text{Prob}[A(t)] \geq 1 - N\lambda$.

- The random process $N(t)$ is the number of stations which are within the transmission range of the station. For each one of the $N(t)$ stations, a random process $X_i(t)$ is defined to be 1 if the i-th station is transmitting at time t and is defined to be 0 otherwise. Then, $X(t) = \sum_{j=1}^{N(t)} X_j(t)$ and the $X_j(t)$'s are not mutually independent.
- Under the operation of the CS/CAMA protocol, more than two concurrent transmissions within an interference area are unlikely. Therefore from Figure 4.12 we can see that if the $X_j(t)$'s are mutually independent, then the event $A(t)$ is more likely to occur than in the dependent case.

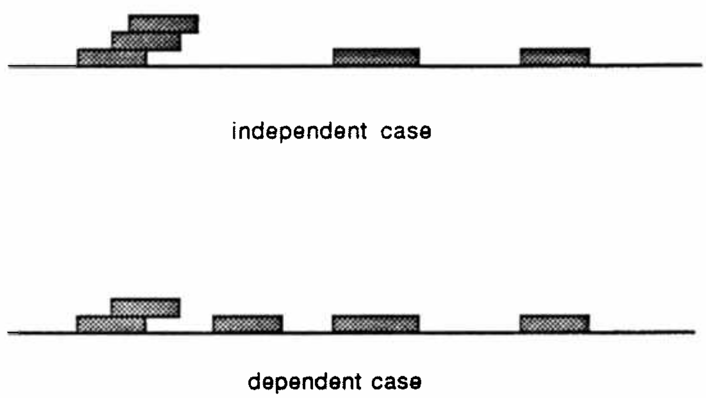


Figure 4.12

Independent versus dependent cases

- Upper bounds on Prob$[A(t)]$ are

$$
\begin{aligned}
\text{Prob}[A(t)] &= \mathrm{E}[I_{\{X(t)=0\}}] \\
&= \mathrm{E}[\mathrm{E}[I_{\{\sum_0^m X_i(t)\}} \mid N(t) = m]] \quad \text{if } X_j(t)\text{'s dependent;} \\
&\leq \mathrm{E}[\mathrm{E}[I_{\{\sum_0^m X_i(t)\}} \mid N(t) = m]] \quad \text{if } X_j(t)\text{'s mutually independent.}
\end{aligned}
$$

In the latter case we find

$$\text{Prob}[A(t)] \leq \text{E}[(1-\lambda)^{N(t)}] = \sum_{m=0}^{\infty} \frac{N^m}{m!} e^{-N} (1-\lambda)^m = e^{-\lambda N}$$

- Using the same method we can show that

$$(1-\lambda N) \leq \text{Prob}[B(t)] \leq e^{-\lambda N}. \quad \square$$

The probability p is related to the events $A(t)$, $B(t)$, and $\{N(t)\}$ as

$$p = \text{Prob}[\, A(t) \cap \{N(t)>0\} \cap \{\text{selected receiver not busy}\} \cap B(t)\,].$$

It is difficult to calculate the value of p since events $A(t)$ and $B(t)$ are dependent; hence, we assume that $A(t)$ and $B(t)$ are mutually independent. (The simulation results in the next section indicate that this is an appropriate simplification.) Upper and lower bounds can now be found on p: $p_L \leq p \leq p_U$ where $p_L = (1-\lambda)(1-e^{-N})(1-\lambda N)^2$ and $p_U = (1-\lambda)(1-e^{-N})e^{-2\lambda N}$. These bounds are plotted with respect to λ in Figure 4.13.

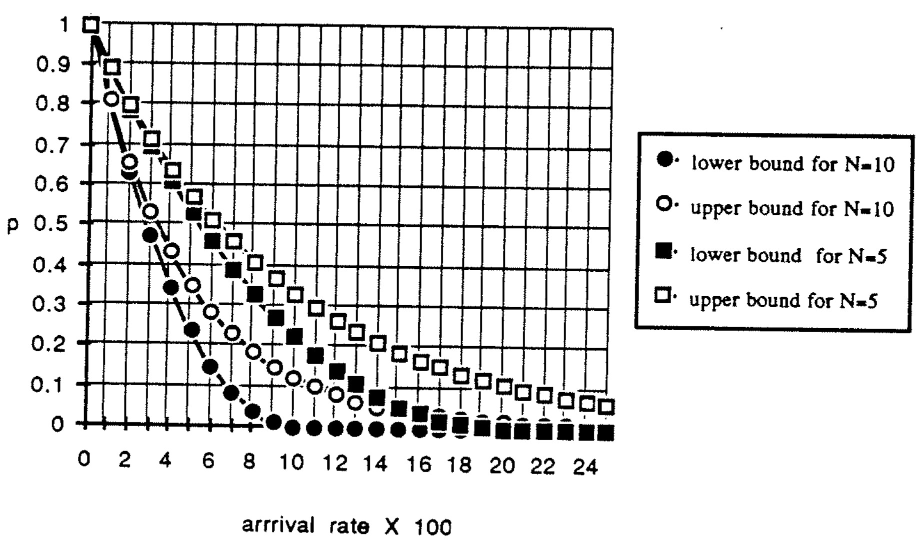


Figure 4.13

Graph of p_U and p_L

Since the larger p, the better the performance of the CS/CAMA protocol, we use p_L with Model I to find the worst performance and use p_U with Model II to find the best performance of the CS/CAMA protocol.

4.5.3. Mean time delay

The mean time delay to transmit a packet from source to ultimate destination is the mean number of hops multiplied by the mean time delay of one hop transmission. The mean number of hops for the MFP routing algorithm is given as follows [Kle78].

$$h = \frac{2}{3}\left(\frac{n}{N}\right)^{1/2}\left[1 + e^{-N} - \int_{-1}^{1} e^{\frac{-N}{\pi}q(t)}\,dt\right]^{-1}$$

where $q(t) = \cos^{-1}(t) - t\sqrt{1 - t^2}$, n is the total number of stations in space, and N is the number of stations within the transmission range of a station.

The mean time delay of one hop transmission equals the mean system time (waiting time + service time) of the M/G/1 queuing. The system time of an M/G/1 queue is represented by the Pollaczek-Khinchin formula [Kle75b].

$$\text{Mean system time } t_d = m_b\left(1 + \rho\frac{1 + C_b^2}{2(1 - \rho)}\right)$$

where m_b = mean service time of the M/G/1 queue, ρ = traffic intensity of the M/G/1 queue, and $C_b = \dfrac{\text{standard deviation of the service time}}{m_b}$.

Model I

An upper bound on the mean delay is found by replacing p by p_L and is

$$hm_b \left[1 + \rho \frac{1 + C_b^2}{2(1 - \rho)} \right]$$

where $m_b = 1 + 2(1 - p_L)\frac{1}{p_L}$, $\rho = \lambda m_b$, and $C_b = \frac{1}{p_L m_b} \sqrt{4 - 3{p_L}^2}$.

Model II

A lower bound on the mean delay is found by replacing p by p_U and is

$$hm_b \left[1 + \rho \frac{1 + C_b^2}{2(1} - \rho) \right]$$

where $m_b = (1 - p_U)\frac{2}{\mu} + p_U(\frac{1}{\mu})$, $\rho = \lambda m_b$, and $C_b = \frac{1}{m_b \mu} \sqrt{2 - {p_U}^2}$.

The above calculations lead us to find the upper bound and the lower bound on the mean time delay which are plotted in Figure 4.14 with respect to different loading conditions and in Figure 4.15 with respect to different transmission ranges. In Figure 4.14, we can see that the mean time delay of a packet is monotonically increasing with respect to the traffic intensity of the protocol. In Figure 4.15, we can see that the mean time delay of a packet is a V-curve with respect to parameter N.

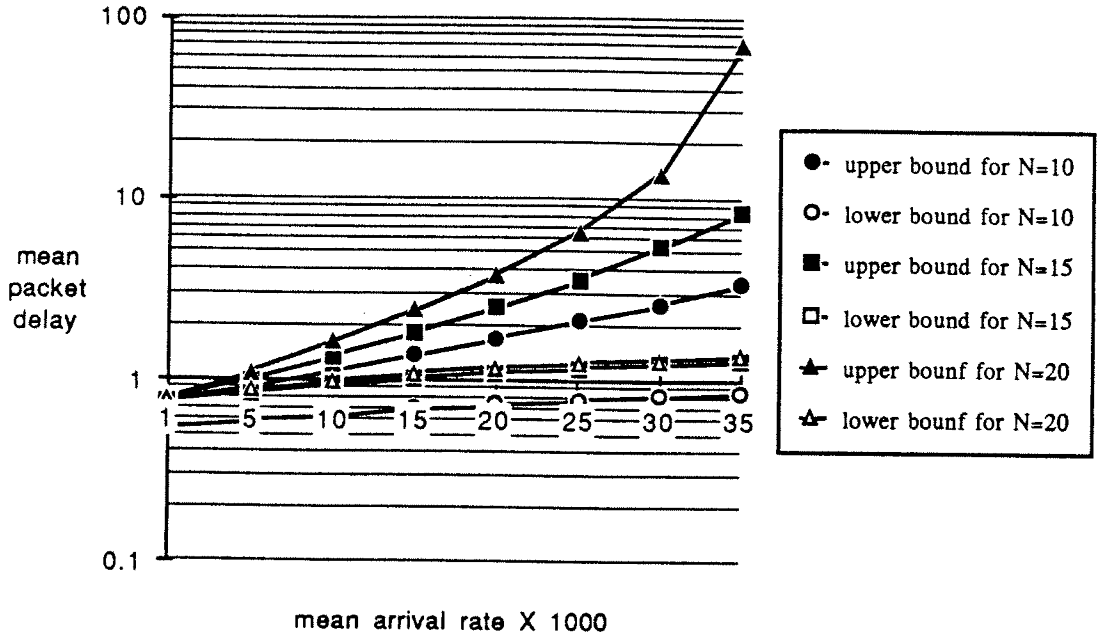


Figure 4.14

Mean delay versus λ

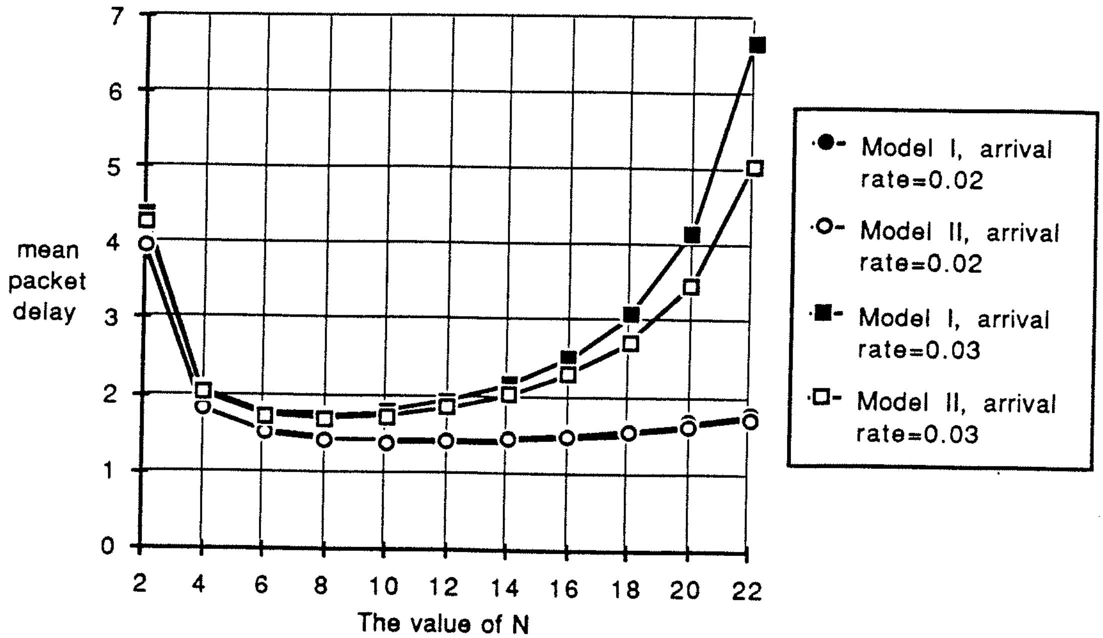


Figure 4.15

Mean delay versus N

4.5.4. Throughput of CS/CAMA

The expected throughput is the expected rate at which packets arrive at their ultimate destinations. If the traffic intensity is less than one, the expected global throughput of the CS/CAMA protocol is equal to the global packet arrival rate, which equals $n\lambda$, divided by the mean number of hops. If the traffic intensity is greater than one, the the network is congested and the expected throughput is the maximum rate that packets can reach their ultimate destination. According to these two relations, the bounds of the expected throughput of the CS/CAMA protocol is calculated as follows.

Model I

A lower bound on the global throughput is found by replacing p by p_L and is

$$\min(n\frac{\lambda}{h},\ n\frac{\lambda_{\max}}{h}) \quad \text{subject to the condition } \lambda_{\max} + 2(1 - p_L)\frac{\lambda_{\max}}{p_L} < 1$$

where $p_L = (1 - \lambda_{\max})(1 - e^{-N})(1 - \lambda_{\max}N)^2$.

Model II

An upper bound on the global throughput is found by replacing p by p_U and is

$$\min(n\frac{\lambda}{h},\ n\frac{\lambda_{\max}}{h}) \quad \text{subject to the condition } 2\lambda_{\max}(1 - p_U) + p_U(\lambda_{\max}) < 1$$

where $p_U = (1 - \lambda_{max})(1 - e^{-N})e^{-2\lambda_{max}}$.

The relationships between S and N and between S and λ are shown in Figures 4.16 and 4.17, respectively.

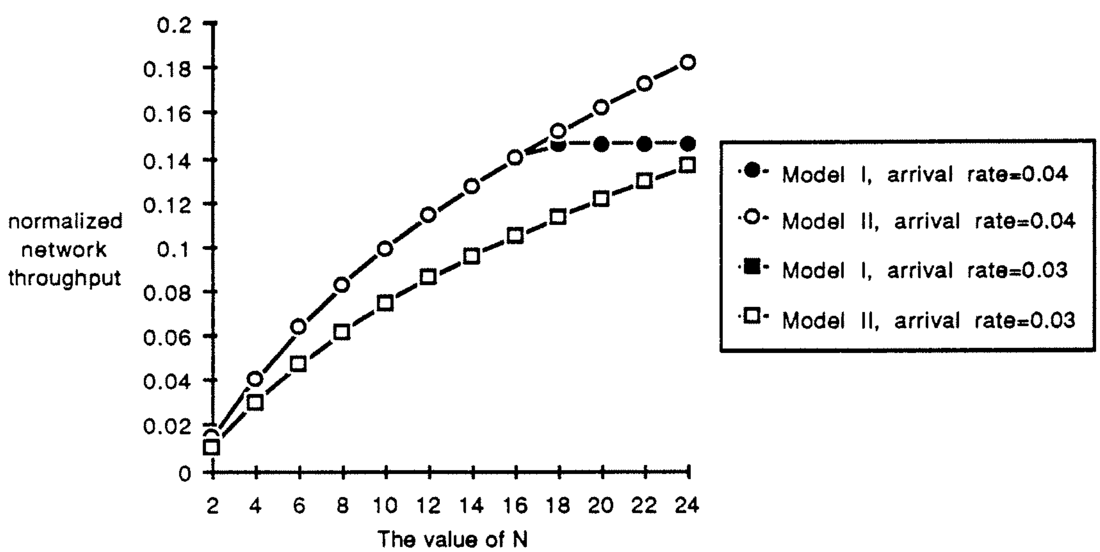


Figure 4.16

Mean throughput versus N

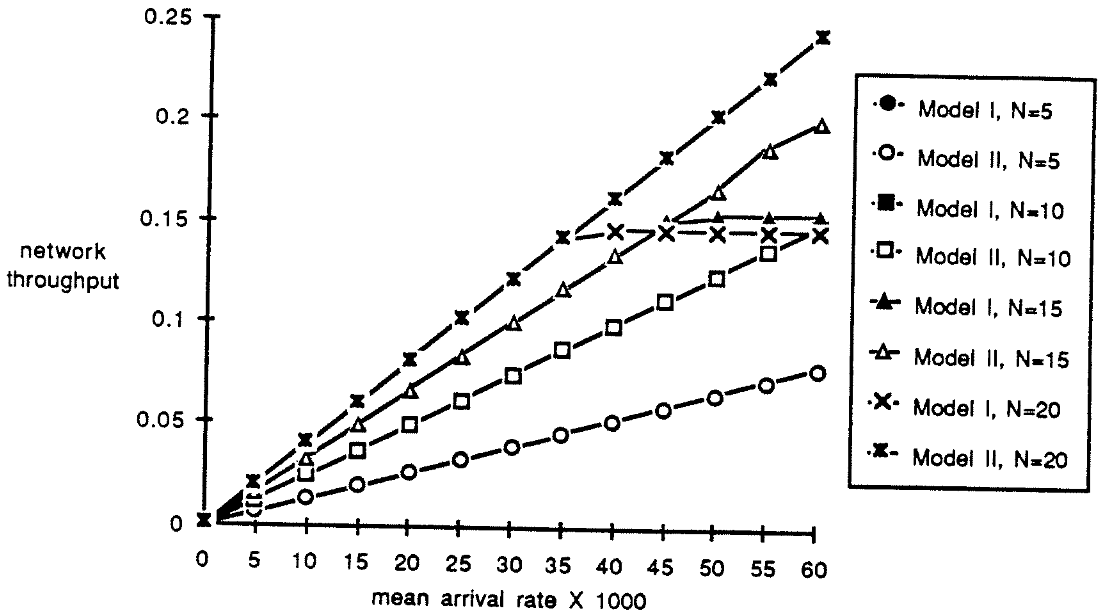


Figure 4.17

Mean throughput versus λ

4.5.5. Capacity of the CS/CAMA protocol

In the previous section, the expected throughput of the CS/CAMA protocol is found to be a function of the mean packet arrival rate λ and N. The value of N at which the expected throughput of the protocol is maximum is called the "magic number" of the protocol and is the value used when we compute the capacity of CS/CAMA in this section.

Model I

A lower bound on the global network capacity is found by replacing p by p_L and is

$$\max(n\frac{\lambda}{h}) \quad \text{subject to the condition } \lambda + 2(1 - p_L)\frac{\lambda}{p_L} < 1$$

where $p_L = (1 - \lambda)(1 - e^{-N})(1 - \lambda N)^2$

Model II

An upper bound on the global network capacity is found by replacing p by p_U and is

$$\max(n\frac{\lambda}{h}) \quad \text{subject to the condition } 2\lambda(1 - p_U) + p_U\lambda < 1$$

where $p_U = (1 - \lambda)(1 - e^{-N})e^{-2\lambda}$.

These bounds are calculated by computer. Figure 4.18 indicates that the lower bound on the global network capacity is $0.1597\sqrt{n}$ with corresponding magic number 10. Figure 4.19 indicates that there does not exist an upper bound on the global network capacity.

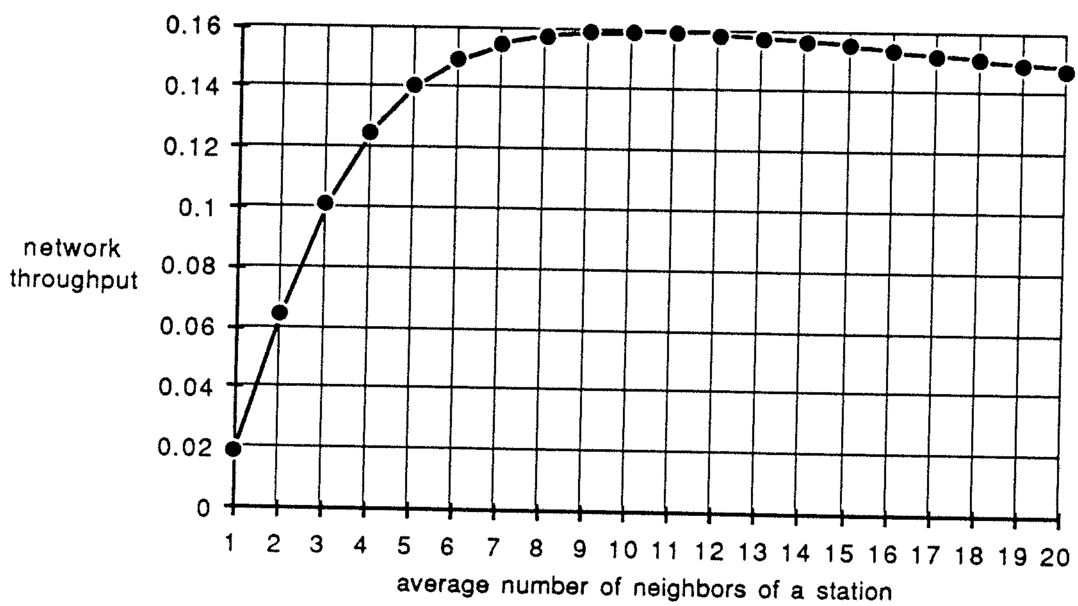


Figure 4.18

The magic number of Model I

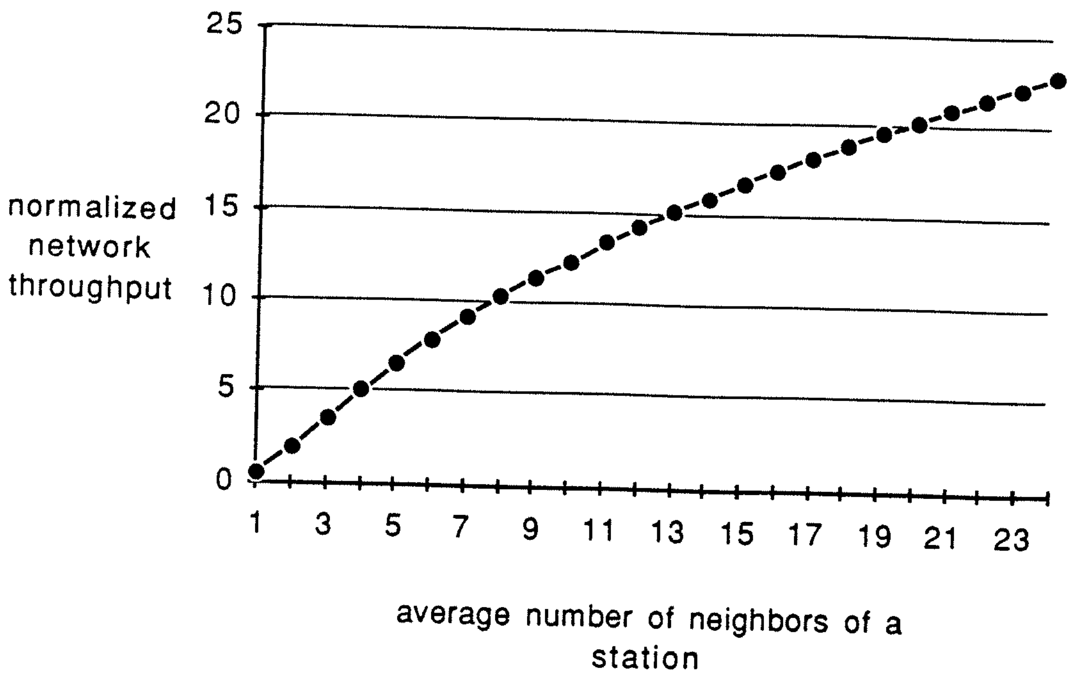


Figure 4.19

Capacity of Model II

4.6. Performance -- simulation model

By simulating four protocols, the CS/CAMA protocol, the multi-hop ALOHA protocol, the multi-hop CSMA protocol, and the multi-hop BTMA protocol, on a computer, we show that the analytical models used in the previous section are valid, and that the capacity of the CS/CAMA protocol is higher than that of a multi-hop ALOHA protocol, a multi-hop CSMA protocol, and a multi-hop BTMA protocol. In the simulation model, 40 stations are randomly placed on a square region with side length 3.5 and each station transmits packets with transmission radius 1. Therefore, the mean number of stations within the transmission range of a station is 10.

Furthermore, we let the opposite sides of the square be coincident and let a packet which leaves the square region from one side enter the region from the opposite side. Thus, this is effectively an infinite network.

The simulation environments are the same as described in Section 4.4 with three exceptions.

- The mean waiting time to retest the channel if busy is one in the simulation model of CS/CAMA. In the simulation models of the other protocols, the mean waiting times are the optimal values, i.e, the values which maximize the capacities of the protocols.
- The CS/CAMA protocol is analyzed with zero and with nonzero propagation delay. All the other protocols are analyzed with nonzero propagation time delay.
- Each station has 100 buffers to store packets. If a buffer is not available when a packet arrives, the packet disappears.

We first simulate the operation of different protocols to estimate t_d, the mean time delay for one hop transmission. After obtaining the estimated value $\hat{t}_d$, we then estimate the capacity of the network, which is the maximal value of $\frac{n\lambda}{h}$ with the constraint $\hat{t}_d < \infty$.

Let n' be the number of obtained regeneration cycles, $N(i)$ be the number of packets arrived in the i-th cycle and $T(i)$ be the sum of $N(i)$ packets' time delay spent in the i-th cycle. Whenever the simulation output data of a new regeneration cycle are obtained, we use the regenerative simulation method [Cra77] to estimate the mean time delay of a packet $\hat{t}_d$ and the half length of 100(1−α)% confidence interval l_α as follows:

$$\hat{t}_d = \frac{\overline{T}(n')}{\overline{N}(n')}$$

where $\overline{T}(n') = \frac{1}{n'} \sum_{i=1}^{n'} T(i)$ and $\overline{N}(n') = \frac{l1}{n'} \sum_{i=1}^{n'} N(i)$;

$$l_\alpha = z_{1-\frac{\alpha}{2}} \frac{\sqrt{\delta_v^2(n') / n'}}{\overline{N}(n')}$$

where $\delta_v^2 = \delta_{11}(n') + \hat{t}_d{}^2 \delta_{22}(n') - 2\hat{t}_d \delta_{12}(n')$, $\delta_{11}(n') = \frac{1}{n'} - 1 \sum_{j=1}^{n'} [T(j) - \overline{T}(n')]^2$,

$\delta_{12}(n') = \frac{1}{n'} - 1 \sum_{j=1}^{n'} [T(j) - \overline{T}(n')][N(j) - \overline{N}(n')]$, $\delta_{22}(n') = \frac{1}{n'} - 1 \sum_{j=1}^{n'} [N(j) - \overline{N}(n')]^2$, and

$z_{1-\frac{\alpha}{2}} = 100(1 - \frac{\alpha}{2})\%$ percentile of the standard normal distribution.

New data are generated and the above procedure is performed repeatedly until the relative half length, $\frac{\text{H.L.}}{t_d}$ is less than 0.1. The estimated one hop mean time delay we obtain here is thus within 90% confidence interval.

With the procedure stated above, the CS/CAMA protocol is simulated with zero propagation time, the same as that we used in the analysis models. The simulation results are compared with the analytical results in Figure 4.20, which show that the analytical models are valid for $N = 10$. The same procedure is carried out to verify the analysis models for $N = 2$, $N = 5$, and $N = 15$.

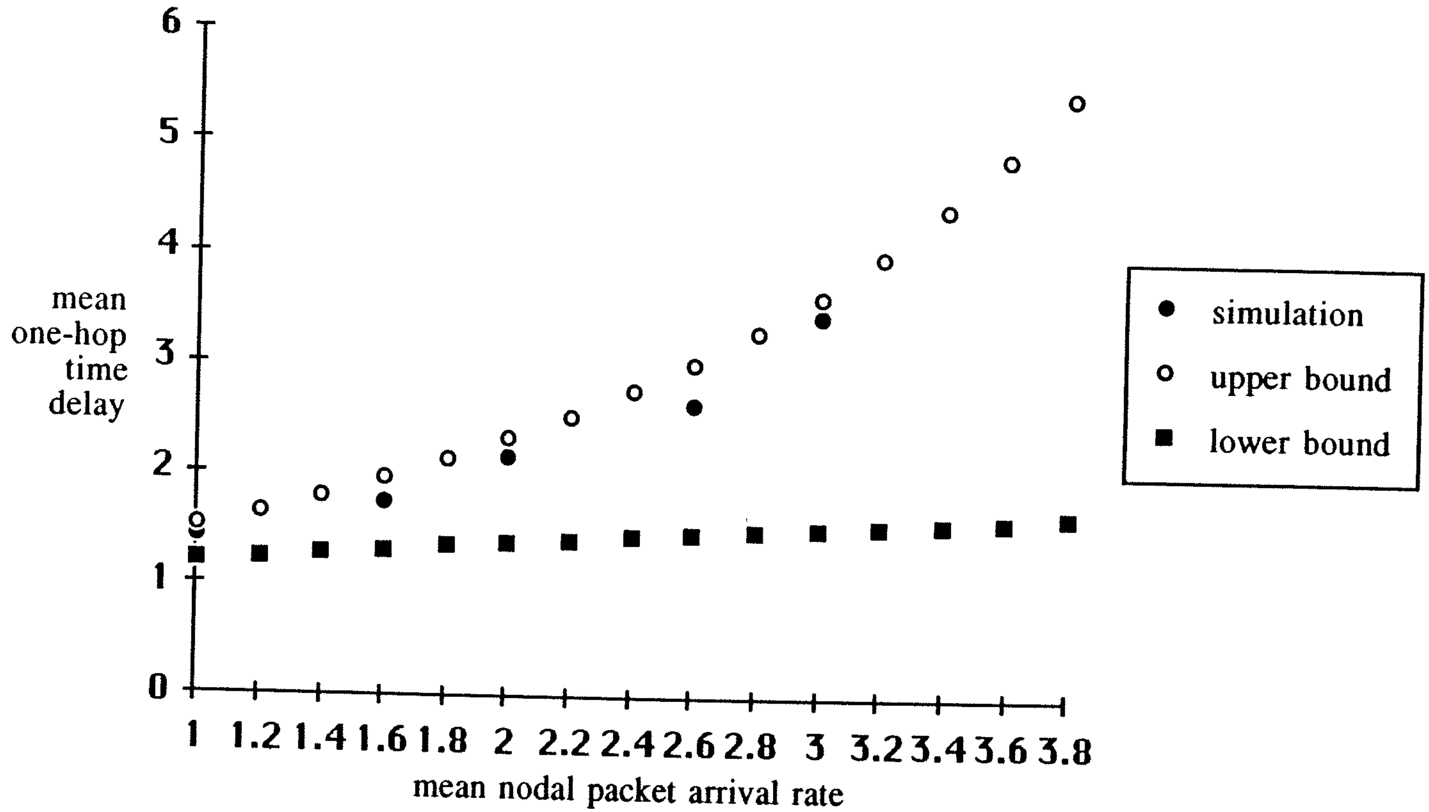


Figure 4.20

Analytical results and simulation results

Second, we simulate the operation of different protocols with propagation time delay 0.1. In Figure 4.21, we compare the performance of multi-hop ALOHA protocol, the multi-hop CSMA protocol, the multi-hop BTMA protocol, and the CS/CAMA protocol. It is shown that the performance of the CS/CAMA protocol is worse than others when the traffic is light. This is mainly because the window control technique we introduced wastes channel capacity while its advantage is nill when the traffic load is light. When the traffic load is heavy, the performance of the CS/CAMA protocol is much better than others because collisions do not occur. (Note that even though the window control technique takes four units of propagation time, the capacity of the CS/CAMA protocol is still much higher than that of others.)

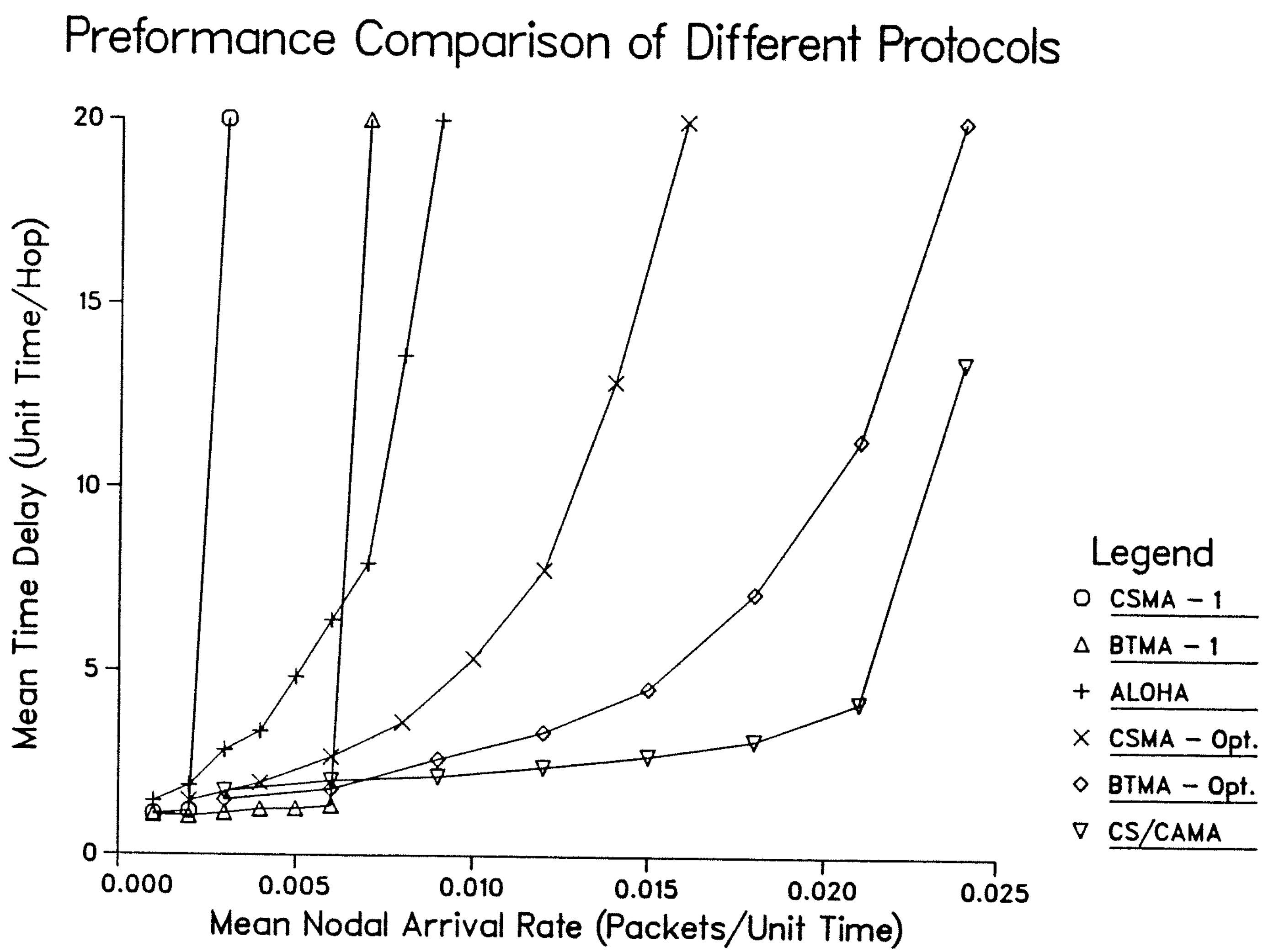


Figure 4.21
Comparison of protocols

CHAPTER 5

CONCLUSION

In this dissertation we have explored topics related to multihop, packet radio, communication networks. We have introduced two "laws," the weak law and the strong law, which if obeyed will ensure that such networks will operate without interference due to coincident transmissions, and have investigated the effects of these laws when then are applied to the regular planar networks. In the course of this investigation we have described two new forms of TDMA for the regular planar networks and have derived the capacities of these protocols when applied to each of the regular networks. One of these, weak spatial TDMA (WSTDMA) obeys the weak law and the other, strong spatial TDMA (SSTDMA) obeys the strong law. Our analysis shows that SSTDMA and a variant of SSTDMA have the highest capacities, especially when applied to square networks.

Since the spatial TDMA protocols, in their pure forms, can only be used in regular two-dimensional networks, we describe a method of modifying these protocols so that they can be used in irregular planar networks, e.g., networks containing mobile nodes. The resulting protocols, i.e., the TREE/TDMA protocols discussed in Chapter 2, possess the best features of both the tree protocols and the spatial TDMA protocols. In these protocols, two levels of control are used on a shared broadcast channel: tree networks are used smooth the flow of traffic to a backbone of repeaters using a spatial TDMA protocol. This allows the TDMA backbone to operate more efficiently. To complete our description of the TREE/TDMA protocols we derive necessary and sufficient conditions for the stable operation of networks using the protocols, calculate capacities for these networks, and evaluate mean packet delay times.

We also introduce a second protocol, the CS/CAMA protocol, which does not make use of a regular backbone but does obey the strong law. Like the TREE/TDMA protocols, CS/CAMA can be used in irregular or random planar networks. Since this protocol involves a complicated "handshaking" between nodes to exchange information about local states, we have demonstrated that the protocol is logically correct, e.g., that it will not deadlock and is essentially conflict-free. We have also presented both analytical and simulation models of the protocol and compare the operating characteristics of the protocol predicted by each model.

The TREE/TDMA and CS/CAMA protocols exhibit characteristics which make them attractive alternatives for use in a multihop packet radio environment. Both protocols are reliable since they are distributed, the latter more so than the former: the nodes function independently of one another and require only local state information; hence, there does not exist a single node whose failure can render the entire network inoperative as in a centralized network. Both protocols are flexible and can accommodate mobile users. No elaborate mechanism is needed for adding or deleting nodes from a network as in certain "token-passing" and reservation protocols. Both protocols allow terminals to maintain a low EM profile; hence, limit detection: terminals need only transmit when conversing and any transmission can be confined to the neighborhood of the transmitter. Neither protocol requires expensive or complex hardware at a node.

Most importantly, the protocols are efficient. Many media access protocols are rendered inefficient in the multihop packet radio environment because either the "hidden area" is ignored by the protocols or in accommodating the hidden area, the potential for spatial reuse of the channel is sacrificed. Both the TREE/TDMA and CS/CAMA protocols eliminate the effect of the hidden area without sacrificing the potential for spatial reuse of the channel.

There are several areas that might be explored were one to continue the work described in this thesis. First, while we were able to show that WSTDMA is optimal, optimal in the sense that there do not exist other protocols for the regular planar networks obeying the weak law which have higher capacities, we were unable to make a similar claim for SSTDMA or its variant SSTDMA'. Note that the one-hop capacities of SSTDMA are optimal but these do not imply that the network capacities of SSTDMA are optimal (in fact they are not as demonstrated by SSTDMA'). The key difference between WSTDMA and SSTDMA is that it is more "efficient" to use short transmission distances for the former and long distances for the latter when attempting to increase *network* capacities. Unfortunately, an idiosyncrasy of SSTDMA prevents the use of a single long transmission distance and in fact prevents the direct (one-hop) communication of certain pairs of stations.

Second, there are direct extensions of this work to space of other dimensions. Extensions of the work relevant to the planar networks discussed in chapters 2 and 3 of this thesis to three dimensional networks would be interesting. The CS/CAMA protocol presented in Chapter 4 can be used in n-space without alteration.

Third, improvements could be made to the analytical models of the CS/CAMA protocol presented in Chapter 4. Recall that we presented models which enabled us to calculate upper and lower bounds on capacity, packet delay, etc., but which did not capture the true operating characteristics of the protocol. A better analytical model of the CS/CAMA protocol would be welcomed.

Finally, there are extensions which could be made to the work presented in this thesis to networks in which traffic patterns are not uniform and isotropic.

APPENDICES

APPENDIX A

INFINITE SOURCE MODEL

An infinite source model is frequently used in the analysis on contention protocols and, in particular, to find a lower bound on the capacity of a protocol. In this approximation, packets are usually assumed to be of fixed size and one unit of time is needed to transmit a packet on the channel. It is assumed that the random processes $V_i(t)$, $i=1,2,...,N$, denoting the number of packets generated at NIUs in slot t are mutually independent and identically distributed with average rates $\lambda_i = \mathrm{E}[V_i(t)]$. (The aggregate network packet arrival rate is thus $\lambda = \sum_{i=1}^{N} \lambda_i$.) Now we allow $N \rightarrow \infty$ and $\lambda_i \rightarrow 0$ such that λ remains fixed. As $N \rightarrow \infty$, $\mathrm{Prob}(X_i(t) > 1) \rightarrow 0$ for $i=1,...,N$ where $X_i(t)$ is a random process denoting the number of packets queued at NIU i at time t. (Note that if, as is frequently done, it is assumed that an NIU wishes to transmit a packet with probability λ_i in each slot and $\lambda_i = \lambda_j$ for all i and j, then the total number of NIUs wishing to transmit in any slot has a binomial distribution for a finite source model and a Poisson distribution for an infinite source model.)

If $Z(t)$ is a random process denoting the delay encountered in serving all packets in the network at time t, then the protocol is stable in the sense that all packets are transmitted with finite average delay if $\lim_{t \rightarrow \infty} \mathrm{E}[Z(t)]$ is finite. In this case the network throughput λ is "achievable" (albeit the delay may be very large). If λ can be achieved by a protocol, then λ is a lower bound on the capacity of the protocol.

APPENDIX B

DERIVATION OF THE PROBABILITY MASS FUNCTIONS

In this appendix we find the probability mass functions f_k, g_k, and the generating function $U(z)$, which are defined in Chapter 3. Since function g_k will be represented as a function of f_k and $U(z)$, we will derive g_k last. The variables involved in the deviation are defined as follows.

$p_1 \quad \triangleq$ probability that a packet leaves the IQ in a frame

$p_2 \quad \triangleq$ probability that a packet leaves the PQ in a frame

$p_3 \quad \triangleq$ probability that a packet move to the SQ

$A_n \quad \triangleq$ number of internal arrivals in the n-th frame

$B_n \quad \triangleq$ number of external arrivals in the n-th frame

$C_n \quad \triangleq A_n + B_n$

Since the probability of finding a stable queue busy equals the traffic intensity of the queue [Kle75a], the values of p_1 and p_2 are NP and $h_\bullet NP$, where N and P are defined in Chapter 3. The value of $p_3 = \dfrac{1}{1 + h_\bullet}$, since the packet arrival rate of the sink queue is NP, the same as the internal queue, and that of the propagatio queue is

$h_* NP$. The probability mass functions of A_n and B_n can be represented as

$$f_{A_n}(a_n) = \begin{cases} p_1 & ; \text{if } A_n = 1 \\ 1-p_1 & ; \text{if } A_n = 0 \end{cases}$$

$$f_{B_n}(b_n) = \begin{cases} \binom{6}{b_n} (\frac{p_2}{6})^{b_n} (1-\frac{p_2}{6})^{6-b_n} & ; \text{if } B_n = 0,1,2,..6 \\ = 0 & ; \text{otherwise} \end{cases}$$

Since $C_n = A_n + B_n$, the generating function of C_n, denoted as $G_{C_n}(z)$, can be represented as $G_{C_n}(z) = [1+p_1(z-1)][1+\frac{p_2}{6}(z-1)]^6$.

From the definition given in Chapter 3, function f_k is the probability mass function of Y_n, which equals $\sum_{i=0}^{C_n} I_i$ where the indicator random variable I_i equals 1 if and only if packet i goes to the PQ. Accordingly, the generating function of Y_n, denoted as $F(z)$, can be obtained as follows.

$$F(z) = \mathrm{E}[z^{Y_n}] = \mathrm{E}[z^{\sum_{i=0}^{C_n} I_i}] = \mathrm{E}[\mathrm{E}(z^{\sum_{i=0}^{C_n} I_i} \mid C_n))]$$

$$= \mathrm{E}[(p_3 + (1 - p_3)z)^{C_n}] = [1 + p_1(p_3 + (1 - p_3)z)][1 + \frac{p_2}{6}(p_3 + (1 - p_3)z]^6$$

After obtaining the generating function $F(z)$, the probability mass function f_k can be obtained from $F(z)$ by taking the inverse Z-transform.

In order to derive the generating function $U(z)$, the random variables A_n and B_n are respectively represented as the sum of two random variables, $A_n = A^p{}_n + A^e{}_n$ and $B_n = B^p{}_n + B^e{}_n$. The random variables $A^p{}_n$ and $B^p{}_n$ represent the number of internal and external arrivals which enter the propagation queue in the n-th time frame, respectively.

The random variable U_n, defined in Chapter 3, can be represented as $U_n = A^p{}_n + \sum_{i=0}^{B^p{}_n} I_i$ where I_i is 1 if the i-th external arrival reaches the propagation queue before the queue is enabled in the n-th frame. From this, the generating function $U_n(z)$ is found to be

$$U_n(z) = G_{A^p{}_n}(z) G_{B^p{}_n}(\frac{1}{2}(1+z)$$

where $G_{A^p{}_n}(z) = [1 - \frac{h_\bullet}{1+h_\bullet} p_1(1-z)]$, and $G_{B^p{}_n}(z) = [1 - \frac{h_\bullet}{1+h_\bullet} \frac{p_2}{6}(1-z)]$. Since $U(z) = \lim_{n \to \infty} U_n(z)$ and $U_n(z)$ is independent of n, $U(z) = U_n(z)$.

In Chapter 3, the probability mass function g_k is defined as $f_k \mathrm{Prob}[U_n = 0] + f_{k+1} \mathrm{Prob}[U_n \neq 0]$. Hence, the generating function of g_k, denoted as $G(z)$, can be represented as $G(z) = U(0)F(z) + \frac{1 - U(0)}{1 - F(0)} \frac{F(z) - F(0)}{z}$. By taking the inverse Z-transform, we can obtain the probability mass function g_k from $G(z)$.

APPENDIX C

STEADY STATE PERFORMANCE

In this appendix, the generating function of the row matrix A will be obtained from the equilibrium equation $A = AP$ where

$$A = [\, A(0)\ A(1)\ A(2)\ \ldots\ldots]$$

$$P = \begin{bmatrix} g_0 & g_1 & g_2 & \cdot & \cdot \\ f_0 & f_1 & f_2 & \cdot & \cdot \\ 0 & f_0 & f_1 & \cdot & \cdot \\ 0 & 0 & f_0 & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix}$$

Since $A = AP$, we have

$$A(z) \triangleq A \cdot \begin{bmatrix} 1 \\ z \\ z^2 \\ z^3 \\ z^4 \\ z^5 \\ \cdot \\ \cdot \end{bmatrix} = A \cdot P \cdot \begin{bmatrix} 1 \\ z \\ z^2 \\ z^3 \\ z^4 \\ z^5 \\ \cdot \\ \cdot \end{bmatrix}$$

which gives us

$$A(z) = A \begin{bmatrix} G(z) \\ F(z) \\ zF(z) \\ z^2F(z) \\ z^3F(z) \\ \cdot \\ \cdot \end{bmatrix}$$

From the above equation, we then have

$$A(z) = A(0)G(z) + A(z)F(z)z^{-1} - A(0)F(z)z^{-1} = A(0)\frac{zG(z) - F(z)}{z - F(z)}$$

The value of $A(0)$ can be found from $\lim_{z \to 1} A(z) = 1$ to be $A(0) = \dfrac{1 - m_f}{1 + m_g - m_f}$ where $m_f = \lim_{z \to 1} \dfrac{dF(z)}{dz}$ and $m_g = \lim_{z \to 1} \dfrac{dG(z)}{dz}$. Thus, the generating function of the row matrix A is $\dfrac{1 - m_f}{1 + m_g - m_f} \dfrac{zG(z) - F(z)}{z - F(z)}$.

APPENDIX D

EQUIVALENT M/G/1 QUEUE

The generating function of the service time distribution of both equivalent M/G/1 queues of models I and II will be obtained in this appendix. The Lemmas we used here follow.

Lemma D.1

Let a server serves a customer for time Y or time Z with probability p and $1-p$, respectively. Then the generating function of the server's service time $G_X(s)$ is $pG_Z(s) + (1-p)G_Y(s)$ where $G_Z(s)$ and $G_Y(s)$ are the generating function of random variables Z and Y, respectively. [Kle75b]

Lemma D.2

Consider a server which serves customers in two stages with times Y and Z, respectively. Then, the service time distribution of the server has the generating function $G_X(s) = G_Y s)G_Z(s)$. [Kle75b]

Lemma D.3

Let a customer's service time be $X_k (X_k = (k+1)Z + kY)$ with probability $p_k (p_k = p(1-p)^{k)}$ where k is a positive integer. Then, the generating function $G_X(s) = \dfrac{pG_Z(s)}{1-(1-p)G_Z(s)G_Y(s)}$ where $G_Y(s)$ and $G_Z(s)$ are the generating functions of Y and Z, respectively. [Kle75b]

Using these three Lemmas we can obtain the service time distribution of the equivalent M/G/1 queues of the two queuing models in Chapter 3.

Model I

In order to find the generating function of the service time of Model I, we translate Model I in Figure 4.15 to its corresponding block diagram in Figure D.1.

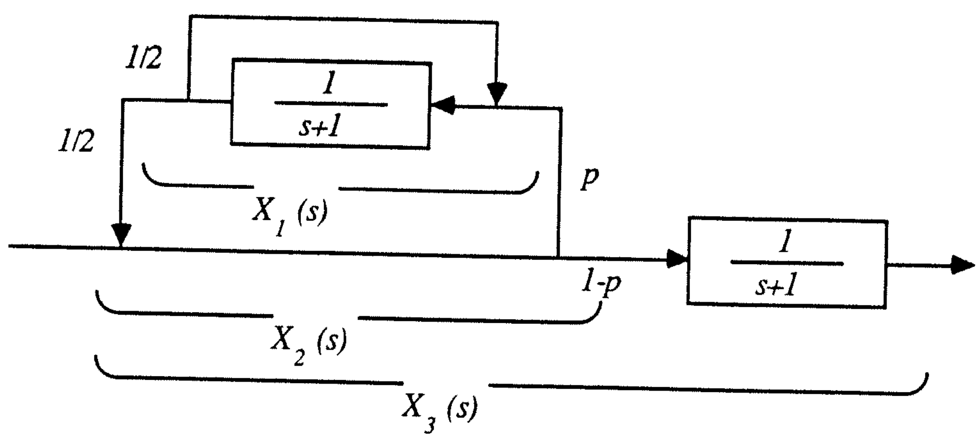


Figure D.1

Block diagram of Model I

Then, by applying the three relations we derived earlier, the generating functions of X_1, X_2, and X_3 can be obtained.

$$G_{X_1}(s) = \frac{1}{2s+1} \quad \text{(Lemma D.3)}$$

$$G_{X_2}(s) = \frac{p}{1 - \frac{1-p}{(2s+1)}} \quad \text{(Lemma D.3)}$$

$$G_{X_3}(s) = \frac{p(2s+1)}{(2s+p)(s+1)} \quad \text{(Lemma D.1)}$$

Thus the generating function of the service time distribution is

$$\frac{p(2s+1)}{(2s+p)(s+1)} \quad \text{(Lemma D.1)}$$

Model II

The block diagram of Model II in Figure 4.16 is shown in Figure D.2.

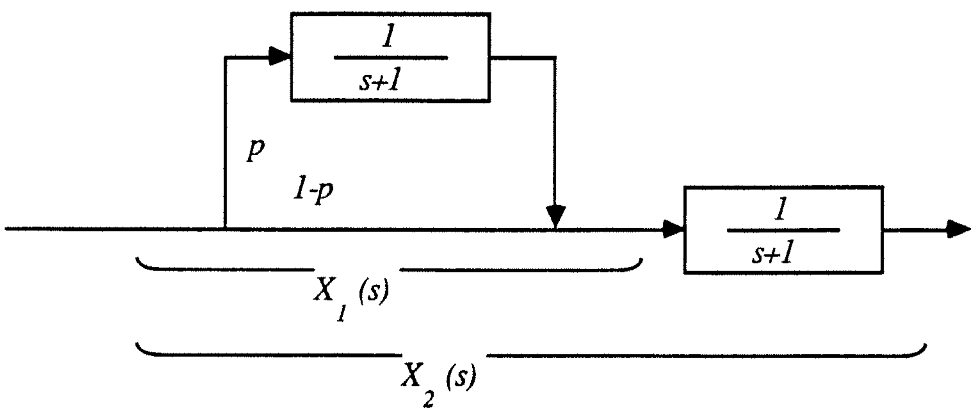


Figure D.2

Block diagram of Model II

We find

$$G_{X_1}(s) = p + \frac{1-p}{s+1} \quad \text{(Lemma D.3)}$$

$$G_{x_2}(s) = \frac{p}{s+1} + \frac{1-p}{(s+1)^2} \quad \text{(Lemma D.1)}$$

Thus the generating function of the service time distribution is

$$\frac{p}{s+1} + \frac{1-p}{(s+1)^2} \quad \text{(Lemma D.1)}$$

REFERENCES

[Abr70] N. Abramson, "The ALOHA System - Another Alternative for Computer Communication," *AFIPS Conference Proceedings FJCC 37,* pp. 281-285, 1970.

[Aka79] G. Akavia and L. Kleinrock, *Performance Tradeoffs and Hierarchical Designs of Distributed Packet-Switching Communication Networks,* UCLA-ENG-7952, Sept. 1979.

[Ber81] T. Berger, "The Poisson Multiple-Access Conflict Resolution Problem," *Multi-User Communication Systems,* Springer-Verlag, New York, 1981.

[Cap79a] J.I. Capetanakis, "Generalized TDMA: The Multi-Accessing Tree Protocol," *IEEE Transactions on Communications,* Vol. COM-27, pp.1476-1484, Oct. 1979.

[Cap79b] J.I. Capetanakis, "Tree Algorithms for Packet Broadcast Channels," *IEEE Transaction on Information Theory,* Vol. IT-25, pp.505-515, Sept. 1979.

[Cox69] H.S.M. Coxeter, *Introduction to Geometry,* John Wiley & Sons, New York, 1969.

[Cra77] M.A. Crane, and A.J. Lemoine, "An Introduction to the Regenerative Method for Simulation Analysis," *Lecture Notes in Control and Information Sciences,* Vol. 4, Springer-Verlag, New York, 1977.

[Gal78] R.G. Gallager, "Conflict Resolution in Random Access Broadcast Networks," *Proceedings of the AFOSR Workshop in Communications Theory and Applications,* Provincetown, 1978.

[Ken63] M. Kendall and P. Moran, *Geometrical Probability,* Griffin, London, 1963.

[Kle75a] L. Kleinrock and F.A. Tobagi, "Packet Switching in Radio Channels : Part I - Carrier Sense Multiple Access Modes and Their Throughput-Delay Characteristics," *IEEE Transactions on Communications,* Vol. COM-23, pp. 1400-1416, Dec. 1975.

[Kle75b] L. Kleinrock, *Queueing System: Vol. 1 -- Theory,* John Wiley & Sons, 1975.

[Lam75] S.S. Lam and L. Kleinrock, "Dynamic Control Schemes for a Packet Switched Multi-Access Broadcast Channels," *AFIPS Conference Proceedings NCC 44,* pp. 143-154, May 1975.

[Lam78] S.S. Lam, "A New Measure for Characterizing Data Traffic," *IEEE Transactions on Communications,* Vol. COM-26, January 1978.

[Mol80] M.L. Molle, *On the Capacity of Infinite Population Multiple Access Protocols,* Computer science Department, University of California, Los Angeles, CA, March 1980.

[Nel82] R. Nelson, *Channel Access Protocols for Multi-hop Broadcast Packet Radio Networks,* UCLA-ENG-8259 (Ph.D dissertation), July 1982.

[Pip79] Pippenger, *Bounds on the Performance of Protocols for a Multiple-Access Broadcast Channel,* Report RC-7742, Mathematical Sciences Department, IBJ Thomas J. Watson Research Center, Yorktown Heights, NY, June 1979.

[Rob72] L.G. Roberts, *ALOHA-Packet System With and Without Slots and Capture,* ARPA Network Information Center, Stanford Research Institute, Menlo Park, CA, April 1972.

[Ros85] S.M. Ross, *Stochastic Processes,* John Wiley & Sons, 1985.

[Sil80] J. Silvester, *On Spatial Capacity of Packet Radio Networks,* UCLA-ENG-8021 (Ph.D. dissertation), May 1980.

[Tak84] H. Takagi and L. Kleinrock, "Optimal Transmission Ranges for Randomly Distributed Packet Radio Terminals," *IEEE Transitions on Communications,* Vol. COM-32, pp246-257, March 1984.

[Tob74] F.A. Tobagi, *Random Access Techniques for Data transmission over Packet Switched Radio Networks,* UCLA-ENG-7499 (Ph.D. dissertation), Dec. 1974.

[Tob75] F.A. Tobagi and L. Kleinrock, "Packet Switching in Radio Channels: Part II - The Hidden Terminal Problem in Carrier Sense Multiple Access and the Busy Tone Solution," *IEEE Transactions on Communications,* Vol. COM-23, pp1417-1433, Dec. 1975.

[Tob80] F.A. Tobagi, "Analysis of a Two-Hop Centralized Packet Radio Network," *IEEE Transactions on Communications,* Vol. COM-28, pp196-216, 1980.