CARLETON UNIVERSITY
SYSTEMS ENGINEERING

ANALYSIS AND DESIGN OF MOBILE/FIXED
RADIO TERMINALS FOR TRANSMISSION
OF VOICE AND DATA

by

J.S. Riordon, A.U.H. Sheikh, M.H. Hafez
and M. Tobis

SCE-83-7

Ours The Task Eternal

Department of Systems and Computer Engineering
Carleton University • Ottawa • Canada • K1S 5B6

ANALYSIS AND DESIGN OF MOBILE/FIXED RADIO TERMINALS FOR TRANSMISSION OF VOICE AND DATA

by

J.S. Riordon, A.U.H. Sheikh, M.H. Hafez
and M. Tobis

SCE-83-7

Table of Contents

Table of Contents

(cont'd)

Abstract

During recent years the land mobile communications has seen a rapid growth, which led to severe spectrum congestion. The available spectrum can be put to an efficient use when a voice channel is shared by more than one subscriber. This can be accomplished because the voice conversations contain considerable silent gaps.

The concept of Time Assigned Speech Interpretation [TASI] applicable to Telephone Switch Network could be applied for the mobile network to increase spectral utilization. A Demand TASI [D-TASI] suitable to mobile channels has been proposed. Three variations of D-TASI: ALOHA, POLLING and TABLE DRIVEN are proposed. ALOHA and POLLING concepts are similar to switched line networks with some time constraints peculiar to the mobile systems. The so-called TABLE DRIVEN system is radically different, and as opposed to the preceding two, does not suffer from technological constraints. The system simulations have indicated that the ALOHA and the POLLING methods are similar and are considerably superior to the TABLE DRIVEN method. It is however felt that superiority of ALOHA and POLLING systems may be due to choice of system parameters which may not be justifiable from the point of view of present day technology.

Finally, problems associated with signal design are addressed. Expressions for packet error rate in fading environment are obtained. The effect of synchronization time on the channel throughput is discussed. It is concluded that the coherent system performs better than the asynchronous system if the synchronization time is kept smaller than a critical value. The critical value of the synchronization time depends on the required throughput. Once the synchronization time exceeds the critical value, the synchronous system performs better.

CHAPTER 1

1.0 INTRODUCTION

For many years mobile radio systems have been used almost exclusively by police, ambulance, and other emergency and dispatch services. The vast majority of these systems were and still are analog FM voice systems, with push-to-talk transmission over narrow FM radio channels in the VHF and low VHF bands. Moreover, each of these systems is assigned one or more (20-30 KHz) radio channels for its exclusive use, and the configuration of the system consists of a base station, which is usually located near the center of the service area and a number of mobile units which roam freely within the (10-20 km radius) coverage area.

Recently, however, there has been an increased awareness of the advantages of radio communications by small businesses, and indeed by the general public. This awareness was triggered and enhanced by the tremendous advantages in solid/state technology, computers, remote control devices and information distribution systems. Indeed, the knowledge which a few years ago was the exclusive domain (and privilege) of the engineering and scientific communities is now fully appreciated by the general public. Moreover, the rapid decline in the cost of high technology products is making these products readily available to more and more people and small businesses. Cordless telephone, personal computers, microprocessor-controlled energy saving devices, robots and satellite receivers are quickly becoming household items.

The impact of this technological revolution is now being felt by the radio regulation authorities. A sharp increase in demand for radio licences has already saturated the band of frequencies allocated for mobile radio systems, and governments around the world are now scrambling in search of ways to accommodate the ever increasing demand on the radio spectrum. The use of single sideband modulation, spread spectrum, cellular systems, high capacity mobile telephony and an increase in spectrum resources through the allocation of new radio communications bands are some of the measures which are either being taken or are under consideration by many governments in the developed countries. The underlying aim is better spectrum utilization. In the very near future the perception of mobile/portable radio will change from being a simple tool for emergency services to an integral part of modern life.

Even if this evolving form of radio communications in urban areas is commercially exploited, still the concept of assigning one channel per active user may be an inefficient utilization of spectrum resources. For example, a radio channel assigned to a taxi dispatch system with 50 operating cabs may result in less than 10% utilization, as in this mode of communication brief conversations are followed by long periods of silence. Statistically studies show that when the two parties (mobile and base) are engaged in an active conversation, silent gaps between talk spurts occupy more than 50% of the time. In principle, this same radio channel could be shared with another user who makes use of intertalk spurt gaps left by the primary user, without noticeable degradation in service quality. Existing technology now provides the means to accomplish this end.

It is obvious that any workable time-shared scheme leads to better utilization of the scarce radio spectrum resource. Many such schemes have been studied in the past for the Switched Telephone Network (STN). Unfortunately the signal environment of the mobile network is quite different from that of the STN, and applications of the STN techniques to the mobile network have to be re-examined carefully under the stringent requirements of radio networks.

In this report we shall focus on one of the most sophisticated time-sharing schemes, namely, Time Assigned Speech Interpolation (TASI). This scheme has been studied extensively for the STN [1-9], and it has a potential capability of approximately doubling the number of users for the same number of voice channels. The adaptation of this scheme to the mobile radio network will be discussed in the following sections, with emphasis on the constraints imposed by the complex signal environments in mobile networks.

CHAPTER II

2.0 TASI for Mobile Packet Radio

The concept of TASI stems from the desire to use the available bandwidth effectively. In conventional telephony, spectrum conservation does not play as important a role as in the case of mobile networks, where the number of allocated channels is limited and the demands on the spectrum are increasing. Cellular radio is one step towards spectrum conservation. With this technique a limited number of channels is used to serve a large area. TASI is a method of increasing the utilization of each channel. This section describes features of TASI which are particularly relevant to the mobile signal environment.

The application of TASI to Land Mobile (LM) networks encounters the following two major difficulties in addition to the complexities in the Switched Telephone Network (STN).

i) STN-TASI each of the users is connected directly to the switching system through a dedicated line. The system controller can therefore assign the speech channel immediately upon detecting the presence of speech from a user. In a land mobile network users are not directly connected to the site controller at the base station; no channel is assigned to a user during a silent gap. When speech activity is detected by the mobile unit, it must therefore obtain a channel assignment from the controller for immediate transmission of the talk spurt.

In addition to the time required to detect speech, the time necessary for channel acquisition must pose some limitation on system design. The protocol used for the acquisition by a mobile unit of an uplink channel to the cell cite is therefore of considerable significance.

ii) In a mobile land network the available spectrum must produce a common communications channel. Thus the geographical area covered is of some significance. Within a local area channels should be allocated to minimize adjacent channel interference. Over a large area a given channel may be used more than once.

Therefore, the major difference between TASI for STN and that in the LM Network is that the latter be based on demand assignment and is termed here as Demand TASI or D-TASI. It will be of interest to identify the parameters of importance which determine the expected performance of the Demand TASI System.

Clearly, the fundamental desire of the user is to hear natural uninhibited speech regardless of how such communication is being processed; that is, the system must be transparent to the user. Such a requirement imposes serious time constraints. Studies have shown that a delay of more than 250 milliseconds in speech results in a serious loss of intelligibility. Also there could be a delay introduced in the system when the direction of the flow of information is reversed. This delay has also an upper bound of 250 msec. Thus any time contiguous speech burst or talk spurt that fails to arrive at its destination within 250 msec is abandoned or blocked. Blocking probability must be kept to a very low level for natural speech to occur, since, if blocking is likely, speech patterns will probably adapt themselves to maintain the channel even when no information is being exchanged, thus defeating the whole purpose of TASI. These time constraints also appear in an STN with D-TASI. However, there are additional constraints in the application of TASI over Land Mobile communication. One important constraint on system performance is imposed by fading of the signal, which is more severe on the downlink channel (cell site to mobile) than on the uplink channel (mobile to cell site). This results in a channel model which is not reciprocal and hence may need different protocols for uplink and downlink channels. Also, due to fading, the system status information available at the mobile or at the base station may be imperfect. Collision of messages prior to allocation process or delay in information updating may also cause errors. Finite, and indeed significant, transmitter and receiver channel switching times is a technology related constraint.

2.1 System Network Configuration

The overall system is based on the high capacity mobile telephone network. Such networks have been installed in the Nordic countries and Japan and are being implemented in the United States, the United Kingdom, Canada, France and Spain. In several other developed and developing countries the feasibility of similar systems is being studied. The TASI system considered here could be thought of as a subset of these high capacity systems with a little variation. These high capacity systems are based on the cellular concept where a much larger area is subdivided into smaller contiguous areas known as cells and each cell is assigned a particular set of frequency which minimizes the adjacent and co-channel interference. Each cell contains on the average one cell site transmitter/receiver and has the task of serving mobiles in that particular cell. The system considered here, though, relies on the transceiver at the cell site; it has connected to it a number of base stations with the cell via land lines as shown in Figure 2.1.

These base stations communicate with their mobiles via cell site rather than going through a Mobile Telephone Switching office (MTSO) as is the case with the Bell's Advanced Mobile Phone Service (AMPS). This necessitates transferring some of the logic/control functions from MTSO to the cell site.

In D-TASI the following are the possible modes of communications:

(i) Mobile to Mobile (in the same cell)
(ii) Mobile to Base (same cell)
(iii) Mobile to Mobile (different cells)

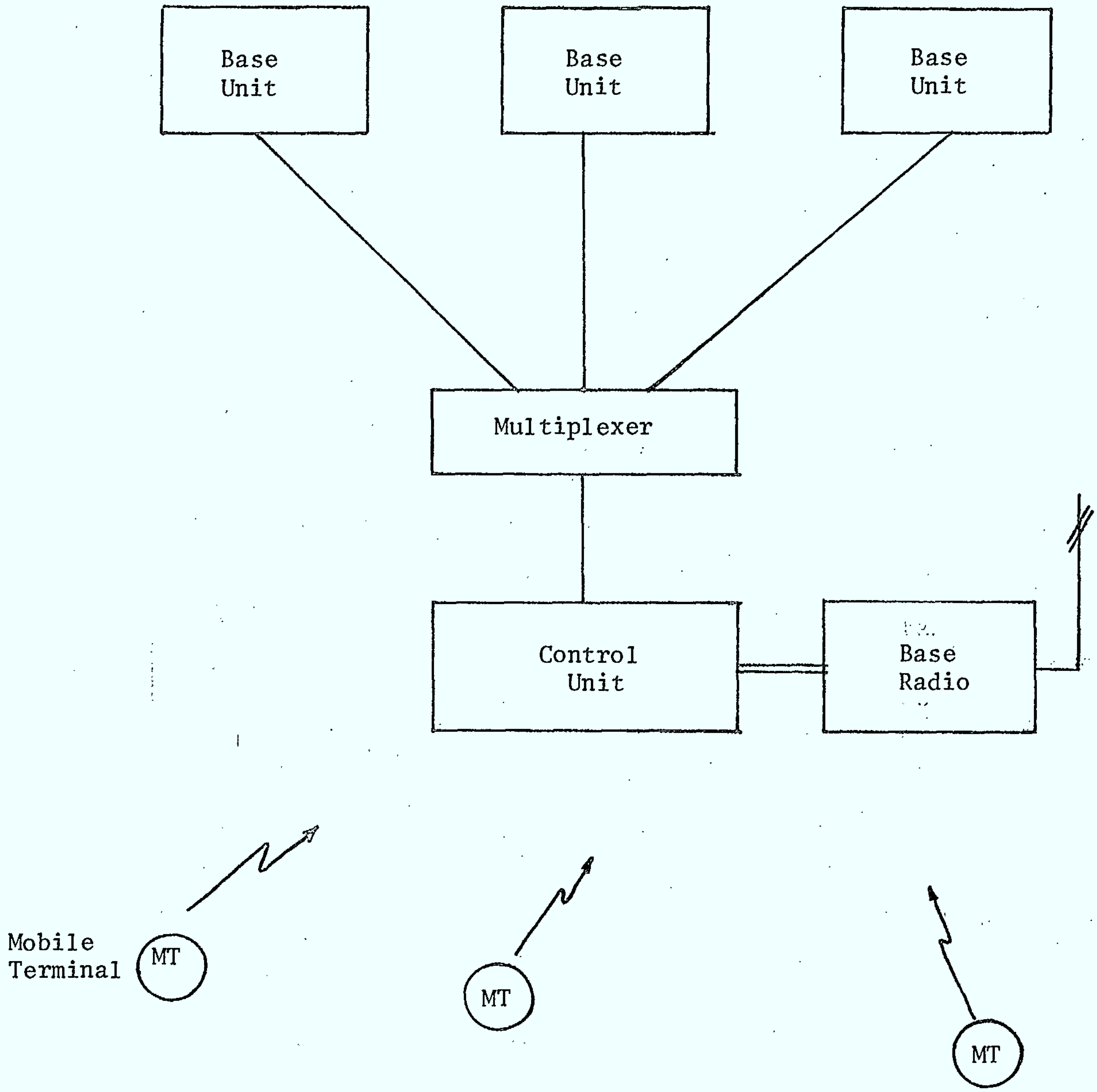


Fig. 2.1: BASIC SYSTEM CONFIGURATION

(iv) Mobile to Base (different cells)

(v) Base to Mobile (small cell)

(vi) Base to Mobile (same cell)

(vii) Mobile to Fixed telephone (Integration with cellular system necessary)

(viii) Fixed to Mobile (Integration with Cellular System necessary)

The base stations basically represent an office complex of a business organization which runs and controls its own mobile units. This arrangement not only provides interbusiness communications but also integrates these subsets into global telephone service. Out of the modes of communications listed above, (iii), (iv), (vii) and (viii) will not be considered further, as the analysis encompassing the total system is beyond scope of this report. We shall restrict our attention to a system consisting of a single cell containing a number of base stations and mobiles and assume that solutions to the problems arising out of intercell and cell-MTSO communications have been solved.

2.2 D-TASI Model

The D-TASI system is considered to have the following modes of operations: Mobile to Mobile, Mobile to Base and Base to Mobile.

In all cases the call must be processed in two stages:

(i) Call Establishment Procedure;

(ii) D-TASI on each talk spurt basis until signing off.

The call is established when the cell site receives a request from the calling party for entry to the system, giving its identification and of the called party. This information is retained in the cell site until the call is terminated. Obviously a call originating from a base station will be handled at no cost to the radio link since the base station is assumed to be wired to the cell site. On the other hand, a call originating at a mobile unit will be established using a dedicated setup channel. These three modes of operations are described below.

2.2.1 Call Originated by a Base Unit

1. The calling base sends a message to the cell site identifying itself and the mobile unit being called.

2. The cell site stores this information and checks if provision exists for an additional active user. If such provision exists, then the calling party is registered as a user and the called party is paged on a forward setup channel.

3. If acknowledgement to the paging message is not received within a certain period of time, the same message is broadcast again and an acknowledgement is awaited. In the absence of acknowledgement after a total of three or four trials, the cell site will abandon the call and inform the base to that effect; the base then makes an exit from the system.

4. If the called mobile acknowledges, the cell site informs the calling base of the readiness of the called mobile to receive the call. At this stage the calling and the called parties become bona fide users of the D-TASI system.

5. Once the call has been established, uplink and downlink channels must be allocated at the onset of a talk spurt. Allocation of downlink channels does not pose a problem as downlink channel status is available at the cell site. However, uplink channel allocation is more complex. The need for an uplink channel from a mobile unit is generated on a random basis by the occurrence of talk spurts, and the cell site has no a priori knowledge of such channel needs. This situation is addressed in more detail in subsequent chapters.

6. At the end of conversation, the base or the mobile sends a supervisory tone to announce the end of conversation and both parties are taken off the list of active users.

2.2.2 Mobile Originated Call

In this category two distinct modes occur:

(i) Mobile Calling Base

(ii) Mobile Calling Mobile

and these are described below.

Case i: Mobile Calling Base:

This case is the simpler of the two. The sequence of events differs slightly from the case described in Section 3.2.1 for base originating calls, and this difference lies in the call setup procedure. The mobile, when intending to communicate with a base, goes through the following procedure:

1. The calling party (mobile) sends a request on a dedicated or uplink setup channel identifying to the cell site both itself and the party it wants to contact, and opens a time window to receive acknowledgement.

2. The request is acknowledged on the downlink channel provided the system can sustain the entry of the calling party into the system. If no such acknowledgement is received, the request is repeated after a random delay. After R unsuccessful trials to access the system, the mobile abandons attempts to access and the call is considered to be blocked.

3. If the cell site decides that the requesting mobile is to be given access to the system, then it is considered to be an active user of the D-TASI until the call is terminated. At the instant of such a decision, the cell site acknowledges the request and also informs the called party. On the receipt of the acknowledgement, the mobile follows the procedure explained in Step 5 of the last section.

Case ii: Mobile Calling Mobile

The call setup procedure in this case starts with Step 1 identical to Case i. Here the mobile requests entry to the system giving its identification and the number called, and waits for the acknowledgement from the cell site. The cell site, on the receipt of the request, checks for provision of such entry to the system. If such provision exists, then it sends the acknowledgement on the downlink setup channel. Simultaneously it pages the called mobile and waits for acknowledgement. If acknowledgement is not received within a specified time, paging messages are repeated and after R failures to get the acknowledgement, the call is abandoned and the calling party is informed of this and the call is assumed to be blocked.

When the cell site receives the acknowledgement successfully, then both the calling and the called parties are registered as active users and normal TASI procedures take over. It is, however, to be noted that now two different mobiles are demanding channels from time to time depending on the arriving talk spurts. It is important that the elapsed time between the organization of the talk spurt and its arrival at the destination should not exceed 250 msec in order to preserve continuity of speech. The mobile to mobile case causes greatest strain on the D-TASI system.

It can be seen that control of communications in a land mobile network is distributed geographically amongst sites which are linked by noisy channels. This gives rise to the possibility of errors and blocking. To help overcome the difficulties of demand assignment, it is necessary to allocate additional communications facilities specifically for signalling purposes. These channels comprise an overhead cost which must be paid to compensate for the effects of fading, delay, and noise.

2.3 D-TASI System Parameters

It was pointed out earlier that the allowable time in which the various control/communications functions, starting from detection of the talk spurt to the channel allocation and transmission, must be completed is 250 msec. How this available time resource is shared among various control functions depends on the channel assignment protocol and on technological and environmental factors. One of the most important questions to be resolved is whether to use coherent or incoherent modulation. This aspect needs a detailed study of relative merits of each modulation technique on the basis of system performance. These issues are taken up in Chapter 5 of the report. However, it may be sufficient to state that though a coherent scheme always performs better, nevertheless overhead associated with carrier recovery may be prohibitively large, given the time constraints. For example, it is estimated that overhead for frame time acquisition is of the order of 2 milliseconds for the incoherent system and 30 milliseconds for the coherent system. In a time constrained system, the performance advantage of the coherent system may have to be sacrificed to reduce the time spent to acquire the channel. The adaptation of TASI for mobile radio may require a dedicated downlink channel over which system status is transmitted from the base station. This information will be required by active stations to avoid repeated timing acquisition delays. This information is transmitted continuously in the form of a bit stream, regardless of the presence or absence of new data. Thus, timing information will be available at all mobile stations at all times, and downlink transmission may hence be regarded as being coherent. Due to the unavailability of mobile signal phase information at the cell site, it is considered that the uplink channel will use incoherent modulation. However, the penalty paid for this limitation is not severe since, owing to the presence of scatterers close to the mobile, the scattered signal components will have wide spatial diversity at the base station with the result that uplink messages are relatively insensitive to fading. Transmission of a short preamble also results in a considerable benefit. Downlink channels are subject to fading to a larger extent than uplink channels, and this fact must be recognized in system design.

2.4 Environmental Parameters

Ideally, the design of a system should involve consideration of environmental factors influencing the signal characteristics. These environmental factors may include the orientation and physical nature of the streets, buildings and their structure, the location and velocity of the mobiles, etc. Such factors affect the signal environment, and inclusion of all of them will make the analysis and simulation too complex to handle. This calls for some simplifying assumptions. This study assumes the existence of Rayleigh fading, which approximates the signal characteristics associated with the vehicle moving at a constant speed. It is recognized, of course, that speed may vary substantially. Indeed, in the extreme case the vehicle may be stationary in a shadow zone and therefore unable to receive a transmission from the cell site.

2.5 Technological Parameters

The D-TASI system in a mobile environment must work under severe time constraints. Some finite time is required for switching the transceiver to newly acquired uplink and downlink channel frequencies. Switching time of the order of 100 msec in the present day mobile units is prohibitively long for D-TASI use. However, recent developments in the field of synthesized receivers by Nippon Electric show that channel switching times could be safely assumed to be less than 30 msec. The other time constraints which are technologically related are processing of system status information, and transmission of commands from the site to the mobiles. It is felt that the final system specifications will be dictated by technological factors.

CHAPTER III

3.0 System Configurations

In light of various constraints for D-TASI discussed in the preceding chapter, three different systems are plausible. These are:

(i) System using ALOHA request channels

(ii) Polling system

(iii) Table driven system

The detailed discussion of these systems follows in this chapter. It includes system architecture, their ability to function within the given time constraints. The merits and demerits of each scheme are discussed qualitatively. Hardware requirements along with their degree of complexities are also discussed.

3.1 System Using ALOHA Request Channels

A subscriber is registered as active by going through the procedure explained in Chapter II. This system, as the name implies, relies on demand assigned channel for transmission of speech packets. The channel is requested on an uplink channel which is shared by many users. Due to sharing of channels, the request which originates as soon as a talk spurt is available at a user terminal, may result in a collision at the cell site with a request from some other user. For a large number of users sending short messages (or requests) on a shared channel, an ALOHA system can tolerate a traffic density of the order of 0.15 units. It is clear that it results in a very inefficient utilization of the channel. This inefficiency arises due to the requirement of retransmission of the collided message after a certain "fall back" time whose distribution has been optimized for the application. This type of system is usually termed as Pure ALOHA.

A modification of Pure ALOHA system, which is twice as efficient in channel utilization, is known as Slotted ALOHA. In the Slotted ALOHA system, the randomness in request transmission is reduced by allocating time slots for the messages. A message is sent in such a system not immediately upon the creation of transmit buffer, but must wait until the beginning of the next time slot. The time slot is of about the same duration as the message packet of duration, say, T seconds. It is clear that in this case, collision occurs only if another user assembles a packet during the same time slot. On the other hand, in a Pure ALOHA system, a collision will occur if another message is assembled either during the T second period of (Slotted ALOHA Transmission slot) T seconds or preceding T seconds. Thus the probability of collision in the Slotted ALOHA system is halved and hence capacity is doubled for the same probability of blocking.

The application of this system to D-TASI is as follows. Assuming that call has been established by following the procedure described in the preceding chapter and D-TASI have taken over the control of the call. A station detecting the presence of a talk spurt would immediately send a message on a dedicated uplink channel to inform the cell site of its identity and since the cell site has already registered the call between two parties, therefore it is already aware of the destination. The cell site immediately assigns an uplink and a downlink channel to the call originating and terminating stations respectively. This assignment is set up as a packet on a dedicated downlink channel which is transmitted as quickly as possible. It may be seen that the downlink message will not be much longer in duration than was the uplink request message and, given uplink capacity smaller than the deterministic allocation of the downlink channel, this delay will not be severe.

The assignment is transmitted repeatedly until the cell site receives the acknowledgements from both the parties, between whom the call has been set up. The destined station sends such acknowleldgement on the dedicated uplink channel and the transmitted on the assigned uplink frequency. Thus a complete connection has been made at the expense of, at most, two messages on ALOHA channel. Ideally, only one message has been sent on the dedicated downlink channel, although this number may increase in case of fading on either of the downlink paths.

It is to be noted that if the downlink message is not subject to fading at either the originator or receiver stations, then both stations will begin switching to the appropriate channels simultaneously. On receipt of acknowledgement from the cell site, the originator starts transmitting immediately. If the acknowledgement is received from the call completing station, the cell site buffers the information until B indicates its readiness to receive. The system is now prepared to receive other packets. In the systems where time constraint applies such simultaneous channel switching may carry considerable advantage. Intuitively, it seems theoretically possible ALOHA traffic density cannot be attained, due to the severe time constraints applied to the system. In order to get subjectively contiguous speech at the receiver, the fall back time must be short which increases probability of a collision on subsequent trials. The packet is abandoned if it has been delayed so much that the time constraint cannot be met .

In order for the ALOHA system to be feasible, even in the case of Slotted ALOHA, the uplink channel should be lightly loaded, as major difficulties arise when a request for channel allocation results in a failure. If a preamble is necessary for time acquisition purposes, then its length will make up most of the request message. The information contents of request message contains requesting station's identifier (perhaps 10 bits), a message identifier (2 bits) and some parity or redundant bits, all of which take considerably less time than the preamble (perhaps 30 msec).

An increase in the number of dedicated uplink channels could certainly alleviate this problem, the number of such channels needed might cause a significant loss of bandwidth advantage, which is the original motivation for considering TASI. This brings about the important point of the acquisition time which should be considerably shorter than 30 msec long preamble. The transmitters must also be capable of transmitting short packets on short notices and then leave the channel free of any interfering energy in the intervening time.

Such transmitters are not known to exist but certainly are feasible. Figure 3.1 shows a logical diagram of the timing of a channel acquisition sequence over an ALOHA D-TASI system.

3.1.1 Hardware Configuration

An ALOHA based D-TASI will consist of the following: a fixed channel receiver (for dedicated downlink channel), a fixed channel transmitter (for dedicated uplink channel) which has the property of being very quiet in quiescent state yet having capability of rapid start and termination of transmitting state, and a variable frequency transceiver of conventional design but having a rather rapid switching capability. Also, about 250 msec of buffering equivalent to 4K bits is required.

At the cell site, some modest computational capacity is necessary, as well as 250 msec of buffering per channel pair may be required with every set of fixed transmitters and receivers, one for each channel. The most stringent requirement to utilize maximum system capacity being very rapid acquisition of the ALOHA channel.

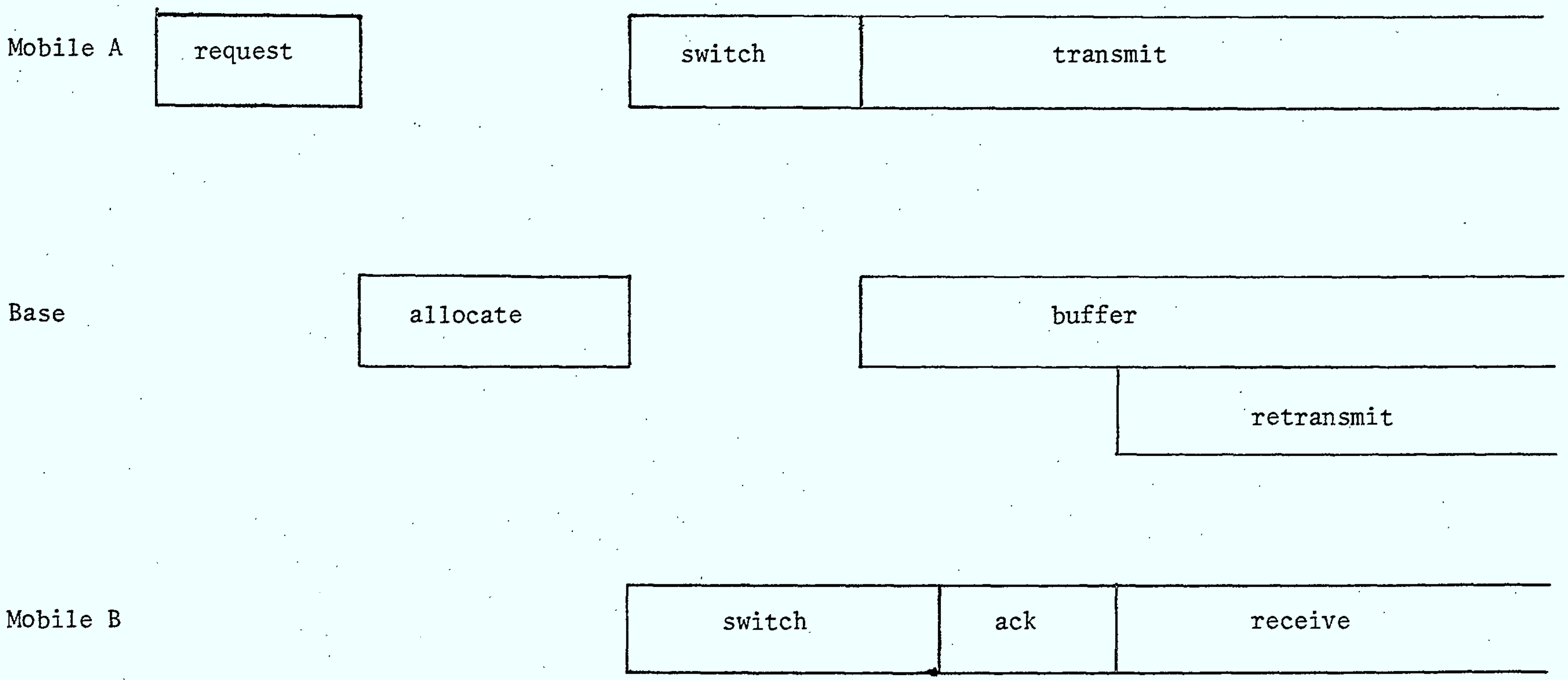


Fig. 3.1: LOGICAL TIMING DIAGRAM FOR ALOHA AND POLLING SYSTEMS IN THE ABSENCE OF FADING

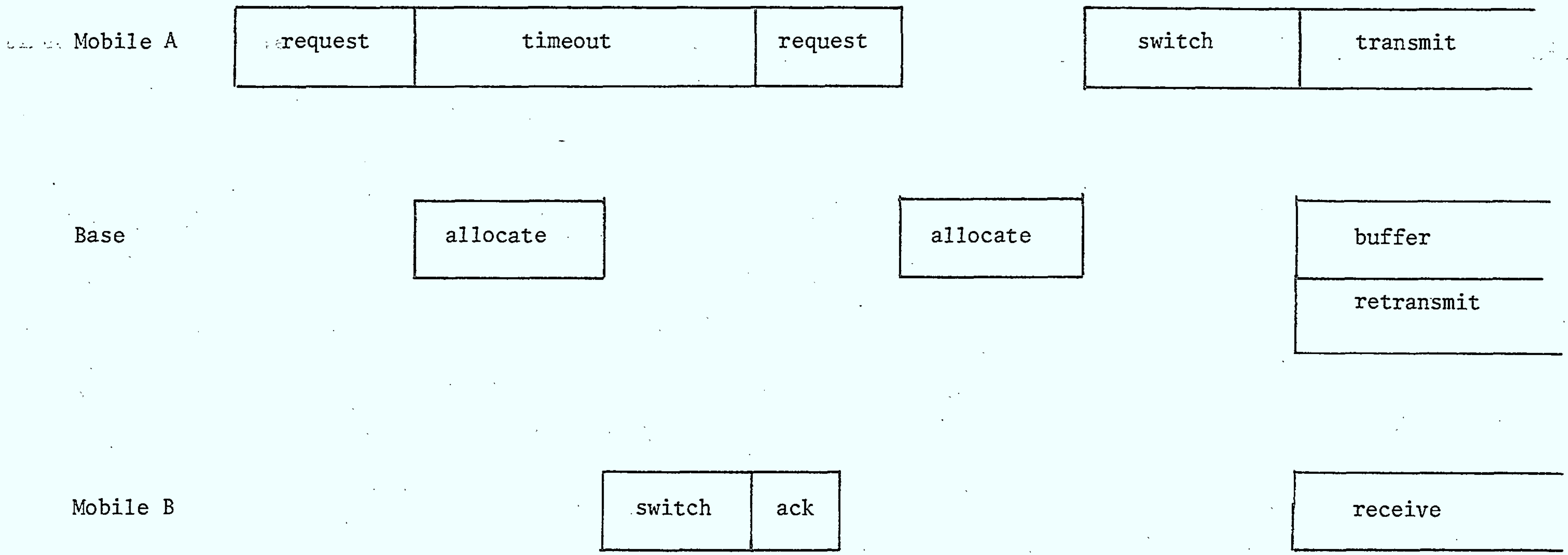


Fig. 3.2: LOGICAL TIMING DIAGRAM FOR ALOHA AND POLLING SYSTEMS. CASE OF A SINGLE FADE ON DOWNLINK TO ORIGINATING MOBILE.

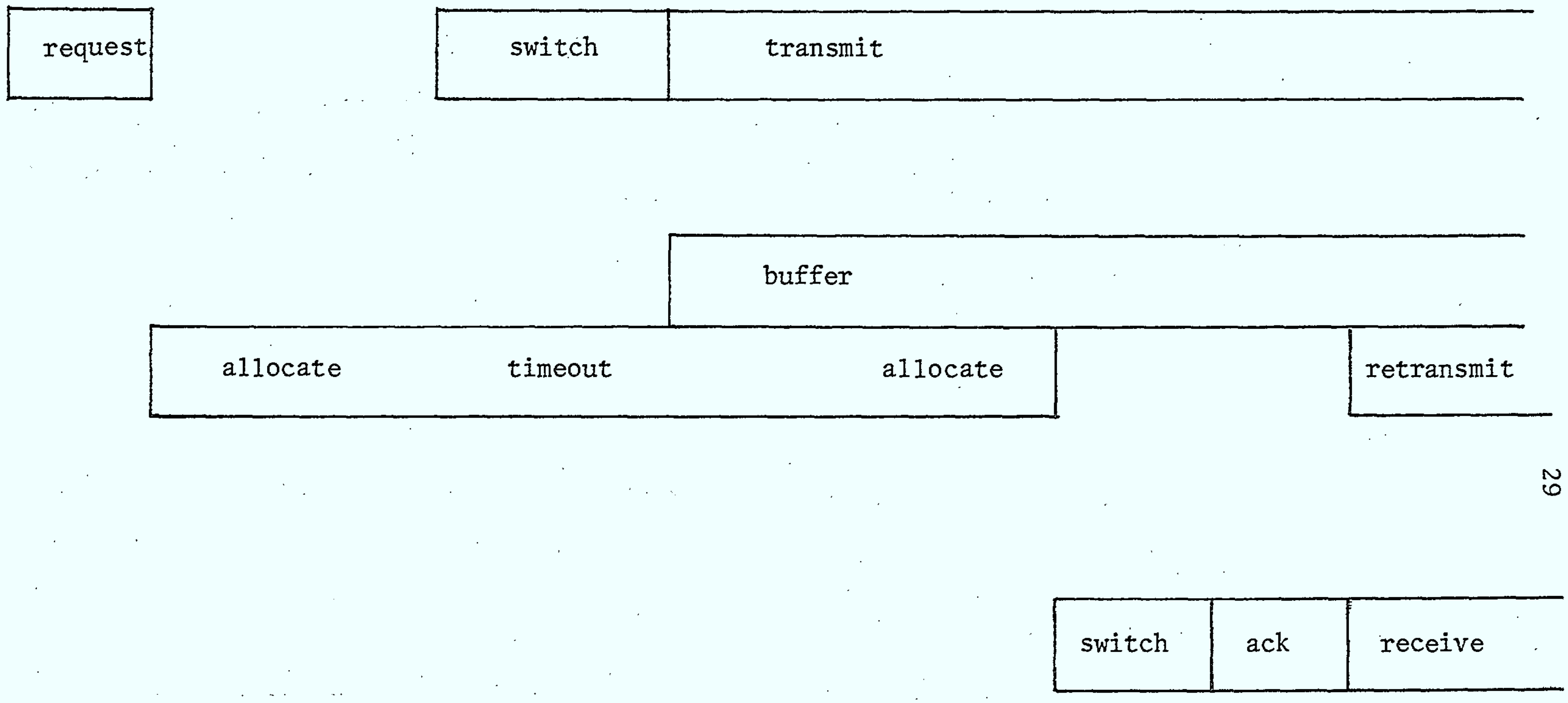


Fig. 3.3: LOGICAL TIMING DIAGRAM FOR ALOHA AND POLLING SYSTEMS.
CASE OF A SINGLE FADE ON DOWNLINK TO DESTINATION MOBILE.

3.2 Polling Approach

3.2.1 System Description: The polling approach is a slight variation on the ALOHA approach. The general idea remains that creation of a talk spurt at a mobile station causes a request to be initiated from a station (mobile or base) to the cell site for channel allocation. Thus, idea of dedicated uplink channel for these requests, as well as for acknowledgement, is maintained, as is the idea of a dedicated downlink channel for the assignment information. The difference is in the nature of the requests. Rather than a stochastic access approach, each active user would be assigned a very short time slot in which to signify the presence of a message by a pulse. Thus if active users existed, for example, and the slot for each were of a duration of 250 user, then each station would have a 250 usec window in which to signify the presence of a message, with such a window appearing every 10 msec.

The cell site receiver examines the uplink channel time slots and determines the caller's identity from the particular slot filled. As the identity of the called party is already known, therefore the cell site allocates an uplink and a downlink channel to the call initiator and call recipient respectively. Both the stations, on receipt of channel allocation information starts switching to the new channel, thus advantage of simultaneous switching of the channels, as in case of ALOHA configuration, is retained.

In such a case, delay due to polling would not be too severe, being uniformly distributed on (0, 10) msec, to be followed by a similar delay upon confirmation. In general, this delay is proportional to the maximum number of slot positions which would probably be identical to the number of active users. Furthermore, while the minimum window duration is set by the uplink channel bandwidth, this duration is quite sensitive to the geographical size of the region being served. The minimum propagation delay of a pulse relative to a window whose limits were determined by a downlink timing frame is near zero, while the maximum is determined by the time for propagation over twice the cell size. This delay is approximately 6.67 msec per kilometre of radius. It may be considered that since this channel is for uplink messages, and thus relatively insensitive to fading, a rather high capacity channel might be set up for the purpose (with resulting expense from the point of view of bandwidth). In this case, the performance of a polling system is limited by its geographical size and its maximum active population.

The question of definition of a call establishment protocol is not trivial in this system, since an idle station does not have an assigned slot in order to communicate with the base station. The solution may be to allocate one or several time slots for request to enter the system.

Another point requiring consideration is whether the polling cycle should be of fixed or variable number of slots. In case of variable number of slots, the number of slots may be added or deleted as the new users enter or established users hang up respectively. It should be kept in mind that variable length polling cycle adds to complexities and may not be desirable.

Additional complications arise if more than one type of message is to be contemplated on the uplink channel. Consider, for example, a situation involving two users with an established call which will be labelled for convenience users A and user B. Suppose that user A requests attention, so that the cell site sends on the downlink channel allocation for A and B. Suppose further that by mischance not only the message to B is lost but also that B in the next polling sequence also sends a channel request. This request may then be mistaken for an acknowledgement. Either some escape protocol from this rare but by no means impossible event must be designated, or some means must be found for distinguishing between requests and acknowledgements.

A similar dilemma occurs in considering the possibility of mixed voice/data traffic in a network of this kind. Since voice is to have a higher priority than data, some indication must be made at the time of request as to what sort of message is being sent. At this juncture, four types of messages, i.e., voice request, voice acknowledge, data request, and data acknowledge have been identified.

To suggest that message window allows for two bits of information leads back to the timing acquisition problems discussed in the previous section. Perhaps the detection of the presence or absence of a pulse could be extended to a decision among positive pulse, negative pulse and no pulse which would allow the base to distinguish between a request and an acknowledgement. Another alternative is to double the polling time with polling cycle alternating between requests and acknowledgements or perhaps between data and voice messages. It is clear that the number of uplink messages must be minimized to keep the polling cycle of reasonable duration.

3.2.2 Hardware Configuration: The hardware requirements of this system are very similar to the ALOHA systems requirements. That is, a fixed channel transmitter and a fixed channel receiver required for uplink channel to be used for requesting a channel and receiving acknowedgements on the downlink channel respectively. In addition, variable frequency transceiver is required for each user to transmit and receive speech packets. The cell site will require a similar array of communication equipment along with considerable computational facilities. As compared to ALOHA, the principal difference is in the uplink request transmitters which will still be required to be very quiet when not transmitting, and yet very quick to attain full energy. The design of such transmitters may be simplified by a fact that only detectable energy need be sent, rather than decodable message. These design constraints may turn out to be considerably less severe. Furthermore, the timing recovery problem at the base station is considerably mitigated in this approach as the presence or absence of transmitted energy in a window should be easy to detect regardless of clock phase.

3.3 Table Driven Approach

3.3.1 System Description: This approach differs considerably from the systems described in Sections 3.1 and 3.2. In the previous cases, all the channel allocations took place at the cell site with the mobile stations acting as (somewhat unreliable slaves). In the table driven approach, decisions relating to channel allocations are made at the mobile station itself. The tradeoff is that, in exchange for using technology which is known to exist, considerable complexity is added to the mobile station itself. The basic idea is as follows. A dedicated channel is maintained for the purpose of sending a complete list of system status. That is, a table is transmitted indicating whether each channel is free or busy. Thus for 160 channels, table information will consist of 160 bits, which may take up 10 msec to transmit. A mobile station detecting a talk spurt (or for that matter, a station wishing to establish a call) would find a free channel, based on its version of the table. It would then switch to that channel and immediately begin transmission. The first message of the call (part of call establishment procedure) would have to include some identifier of the calling and the called in parties. Here, it would seem that the cell site has been caught unaware of impending traffic. This drawback can, however, be removed by using dedicated uplink channel or channels for call setup purposes only. The cell site on the receipt of this information, will inform on the dedicated downlink channel all the stations of the occupancy of the channel, i.e., directive to modify the table accordingly. Furthermore, the designated recipient would have to be informed of the presence of a message for it. Thus, the dedicated downlink channel would have to carry an explicit message indicating the availability of a packet for the recipient. This implies that the downlink channel is not being used for an exclusive purpose of transmitting table information, it is also being used for informing the called in party to switch to an allocated channel. This means that complete table information is transmitted from time to time and not continuously. The time elapsed between the complete table information transmission may be considerably longer than the time it takes to transmit the complete table update.

The Table Driven system is subject to some impairments that do not affect the request driven systems described earlier. One obvious problem with this system is the possibility of message collision. This would arise if two users began transmission on the same channel nearly simultaneously, so that neither header was received at the base station in uncorrupted form. The solution to this for any transmitting station to abandon a transmission after a suitable time out without a confirmation on the dedicated downlink channel (which would be of the form of a channel allocation to the designated receiver in a mobile to mobile call). Unfortunately, in this case, both users have wasted the total of switching, header, and time out time and have to start the process anew. In a heavily loaded system; furthermore, they will be likely to collide again on their next attempt at acquiring one of a small number of available channels.

A second type of impairment occurs when the header message is received at the cell site but has not yet been acknowledged on the dedicated downlink channel when another station attempts to transmit on that channel. The second station will abandon its attempt after the confirmation to another user or the time out, but the established user will experience a period of garbled transmission.

Thirdly, due to the finite time for channel switching, and especially in heavily loaded systems, the possibility exists that even a correct decision about the existence of a free channel may be invalidated by the time the switching has been completed. The authenticity of the table contents may also come under doubt if the station fails to update its version of the table due to its location in a shadow zone. This impairment is perhaps common to all the mobile systems.

It is also not clear that this configuration as such can be extended to a mixed voice/data integrated system or any other system where packets of two different priorities exist.

The reason is simply that, in a system which is completely occupied, there is no way to request resources on a priority basis, since there is no way of communicating with the cell site if all the message channels are occupied. One may consider allocating a dedicated uplink channel to alleviate this difficulty at the expense of a reduction in the bandwidth resources. The same dedicated uplink channel could now be used for entry into the league of active users.

Perhaps the most serious problem in the table driven approach is the need for sequential switching. That is, the receiver station cannot begin to switch until the caller has actually acquired the channel. It is possible too that this may be mitigated by continuously switching the channels during idle times so that the transmitter, as nearly as possible, is always ready. Nevertheless, despite all these impairments, the table driven approach can result in considerably higher throughput as the channel resources used up in channel allocations in case polling and ALOHA systems are being utilized in transmission of messages. A logical diagram of successful channel acquisition and a packet transmission is provided in Figure 3.3. Figure 3.4 illustrates the situation where message is aborted due to non-arrival of acknowledgement from the base station within a time out period. If the age of the talk spurt is not too long, the source mobile will try to transmit again on another idle channel.

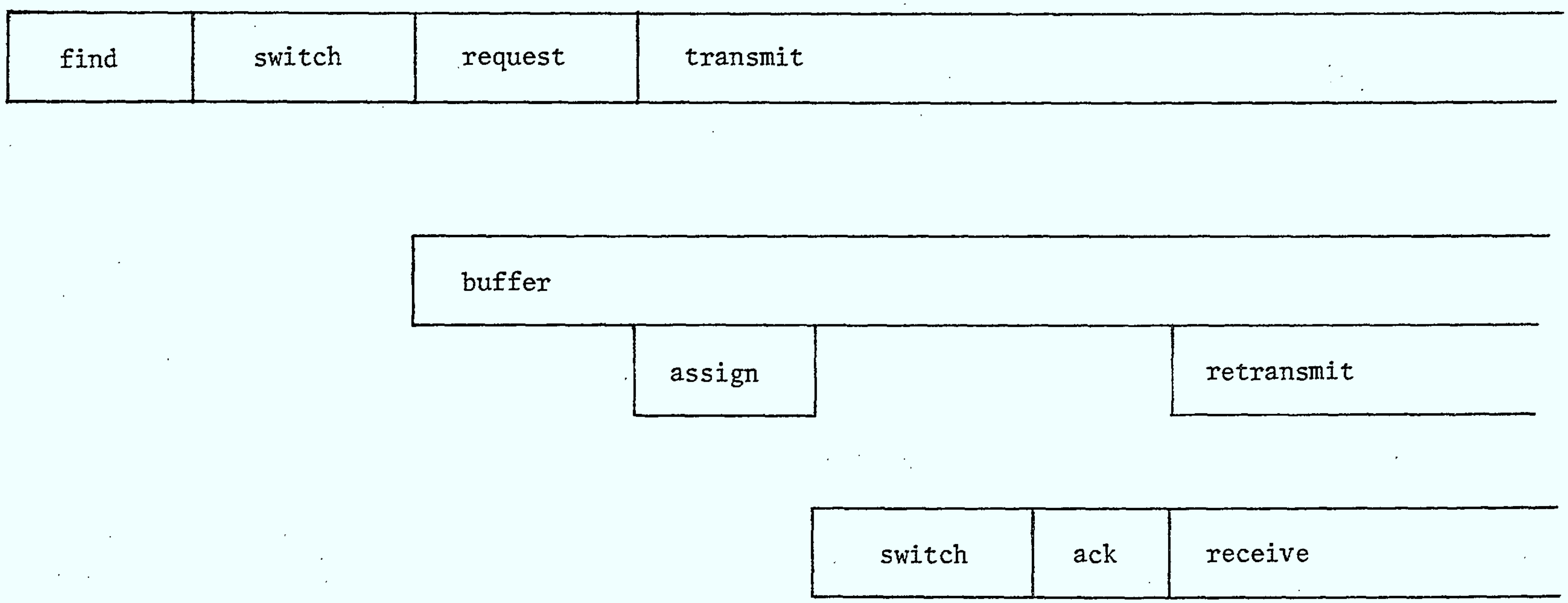


Fig. 3.4: LOGICAL TIMING DIAGRAM FOR TABLE DRIVEN SYSTEMS

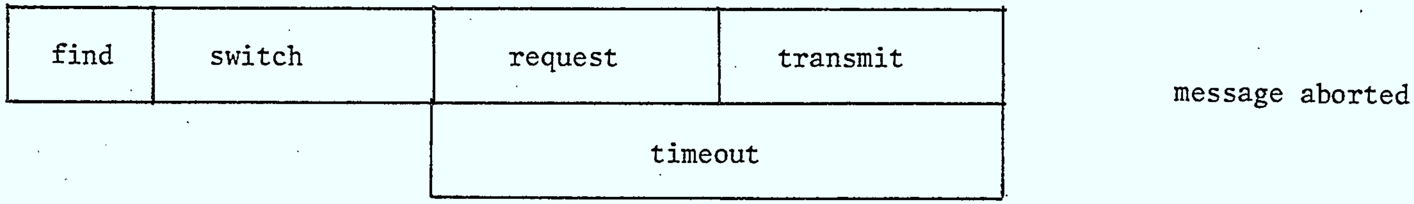


Fig. 3.5: LOGICAL TIMING DIAGRAM FOR TABLE DRIVEN SYSTEM WITH ALLOCATION

3.3.2 Hardware Configuration: The table driven approach has made a tradeoff in hardware, with respect to the request-driven approach of some logic circuitry in the former case compared to a fixed frequency (dedicated uplink channel) transmitter in the latter. The circuitry must receive and decode dedicated downlink channel messages and use them to update its local channel status table, and thereby to make decisions about optional channel selection. A fixed frequency receiver, a fast switching transceiver, and 250 msec of buffering are still required.

The cell site will again need an array of transmitters, receivers and data buffers. The complexity of the computational power at the cell site will be only slightly reduced, if at all, compared to the request methods, owing to a necessity to update the master table and to optimally allocate the time of downlink channel. However, it must be noted that absolutely no assumptions are made regarding the receiver transmitter technology, which is known to exist.

3.4 Summary

From the discussion it follows that systems based on ALOHA request channel and polling concept are very much alike. The polling system has an advantage of reducing the collision due to allocation of time slots and thus seems to have some capacity advantage over ALOHA. Simulation of the system in the next chapter will clear up the picture. On the face of it, the table driven approach seems to have a major problem of message collisions which may cause a greater loss of speech packets resulting in quality degradation. This problem could perhaps be alleviated by adopting a different strategy on system access. Also the issue of the system not being capable of communicating except via voice channel could be resolved by allocating a dedicated uplink channel or channels. This will, however, result in some reduction in the overall capacity of the system.

CHAPTER IV

SIMULATIONS

4.0 Introduction

For meaningful simulation of D-TASI, it will be worthwhile to establish some figures for simulation parameters. These parameter values can then be included in simulation design.

The parameters of the systems are dependent on the particular situation and mobile radio configuration. For example, in mobile cellular system, cell size, user population, nature of destination and reliability may be important parameters. Several questions arise in the category of "nature and destination of the traffic." Firstly, there is the question of whether the majority of traffic is between one mobile station and another, or from the mobile station to the base (or some transmission line network connected to the base station). Preliminary tests show [10] that 60% of the traffic originates from the mobile and majority of it is addressed to base stations. However, in this study mobile to mobile communications, for the reasons that such environment is more demanding upon the time constraints of D-TASI, than would be for mobile to base only. In the latter case, the complexity is reduced by half. Once we demonstrate the feasibility of mobile to mobile communication, it will be safe to assume feasibility of simpler cases of base to mobile and mobile to base.

Secondly, the use of mobile network by dispatch and police services do not represent typical conversational speech, but would constitute mostly silence broken by short and functional talk spurts. It is always technically sound to prove the feasibility of a system for the worst case, then the simpler cases are taken care of automatically. Therefore, naturally occurring speech in a normal conversation is taken as a bench mark for performance. The parameters of natural speech vary widely among speakers and circumstances. However, speech to total time ratio of 0.4 is generally accepted, with talk spurts of 1 to 5 seconds long.

The question as to whether the network is to maintain some data traffic as well as voice must be addressed. The criteria for data traffic are different from those of voice. In data traffic errors cannot be tolerated, but delay is acceptable in order to avoid large instantaneous usage peaks. In case of speech, loss of few bits does not seriously degrade the system performance, but an excessive delay will result in an unacceptable system. This implies that voice/data integrated systems must have two distinct protocols with speech having priority for speed (to avoid delays) and data packets for reliability. This very broad topic is examined only briefly in this paper and much work remains to be done to design protocols to implement these divergent goals simultaneously. There are indications nevertheless that even a heavily loaded TASI network can tolerate some data traffic.

A final consideration which relates directly is the reliability constraint. This refers directly to the quality of the service, i.e., the probability of blocked calls. In the context of TASI, it will be the probability of blocking per call spurt. If a blocking probability of 0.2% is assumed, then probability that a talk spurt fails to reach its destination will indeed be very low.

4.1 Simulation Design Considerations

The simulations for the three systems are based on an event queue and can be classified as event driven simulations. Any change of state scheduled anywhere in the system was recorded as an event, where each event is a record whose fields include a pointer to the next scheduled event.

The occurrence of an event is simulated by popping the first scheduled event from the event queue and all the system parameters affected by such occurrence are changed. In this implementation, each speech stream has a single corresponding event in the event queue. As a result, part of the system update that corresponds to an event occurrence is the scheduling of another event. In these programs, this was accomplished by redefining several of the event fields, notably ones denoting the type of the event and the time of the event. In principle, event driven simulations may have the number of events in an event queue which may change dynamically during execution, with some events causing more or less than one new event to be scheduled right after execution of the current event. This would indeed be necessary in this application for a more complete simulation of the cell site behaviour.

One advantage of having an event queue is the ability to detect interacting events. Our greatest concern is the detection of collisions, a factor which has not been investigated in mobile TASI to date. A beginning of packet event may cause a check for a collision with already scheduled beginning of packet and end of packet events. The possibility that a packet may collide with another packet not yet in the event queue is taken care of by checking for collisions again at the end of the packet. An end of packet event involves a decision between successful or unsuccessful transmission. Using this device of two events per packet subject to collision, an event driven simulation thus handles every possible case.

The program simulated fades by simple expedient of assuming an overall rate of fades for any downlink transmission. This neglects the fact that at any time some stations are more subject to fades than others, which may slightly skew the delay distribution. However, the overall system, traffic should not be greatly affected. Fading and interference on uplink transmission were also neglected. This is reasonable, since antenna size and space diversity at the base station makes loss of packets a relatively rare event which will not have a large effect on overall traffic performance. It is of course necessary that a final protocol specification take these possibilities into account to enable clean and rapid recovery from such unlikely events.

Another aspect of the system which has not been simulated is the call establishment period. That is, the means whereby a user acquires active status and rings the desired destination has not been considered in this simulation. This aspect of the system would be a straightforward extension of the ALOHA and the table method, but somewhat problematic in the polling approach. One solution may be to reserve a special polling slot for call establishment. In any case, the call establishment will form a small percentage of the request traffic.

Finally, the effect of fades on the actual data was not considered, but this is an unavoidable feature of mobile radio regardless of whether TASI is used.

4.2 ALOHA Based System Simulation

Figure 4.1 illustrates the aspects of the ALOHA based system protocol which were simulated. The buffering of the message data at the base station is shown in the flowchart, but was not taken into account in the simulations. The total delay reported in the simulations did take into account the difference in time between the beginning of the message at the originating station and its reception at the destination.

It is seen that if no channel is available, the originator will reissue a request after a short time. In the simulations reported here, this time was taken to be 50 msec. A slightly more sophisticated protocol might cause a queue to be formed at the base station, which might make delayed assignments on a first come first served basis, which would ameliorate congestion on the ALOHA channel.

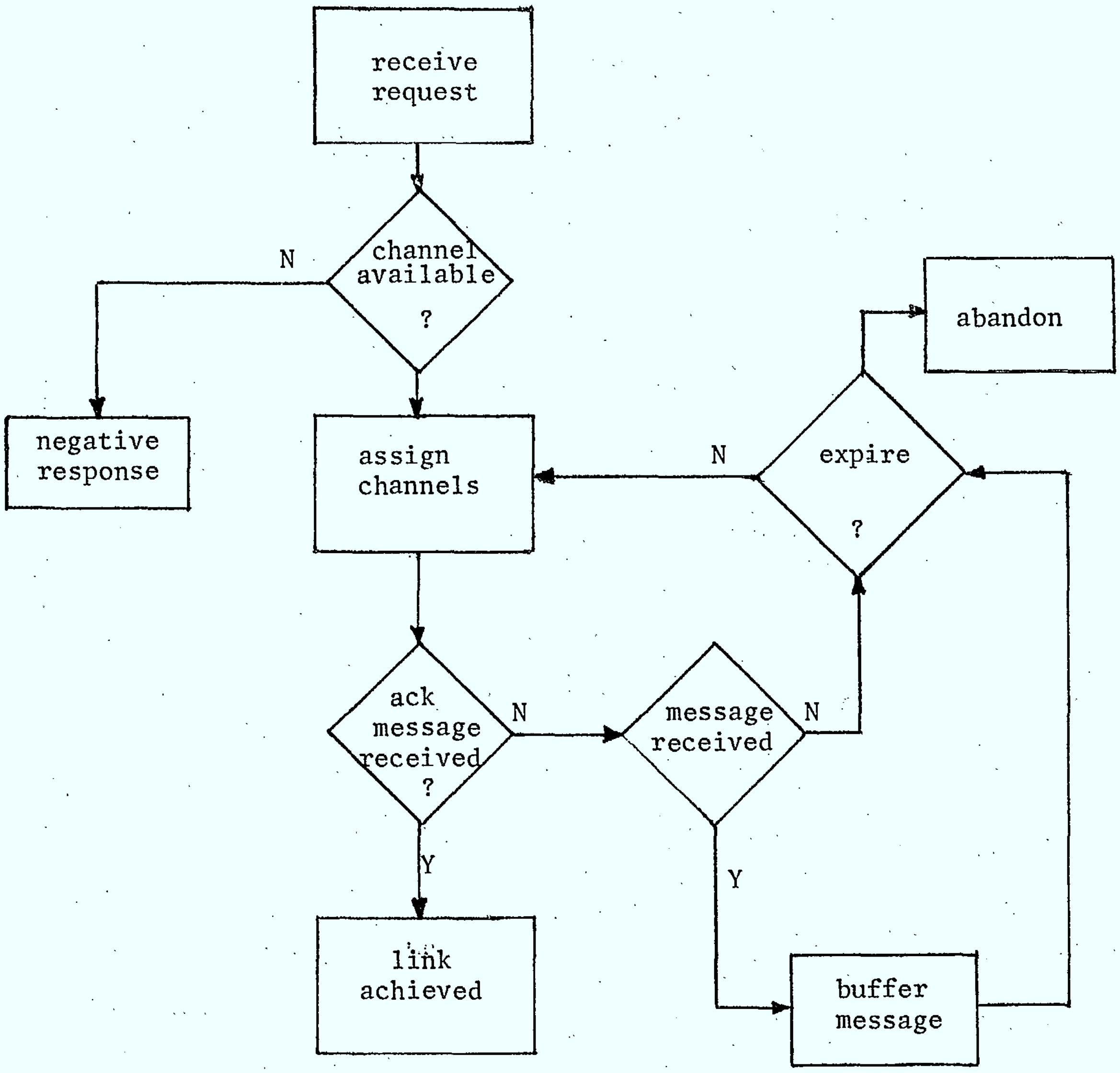
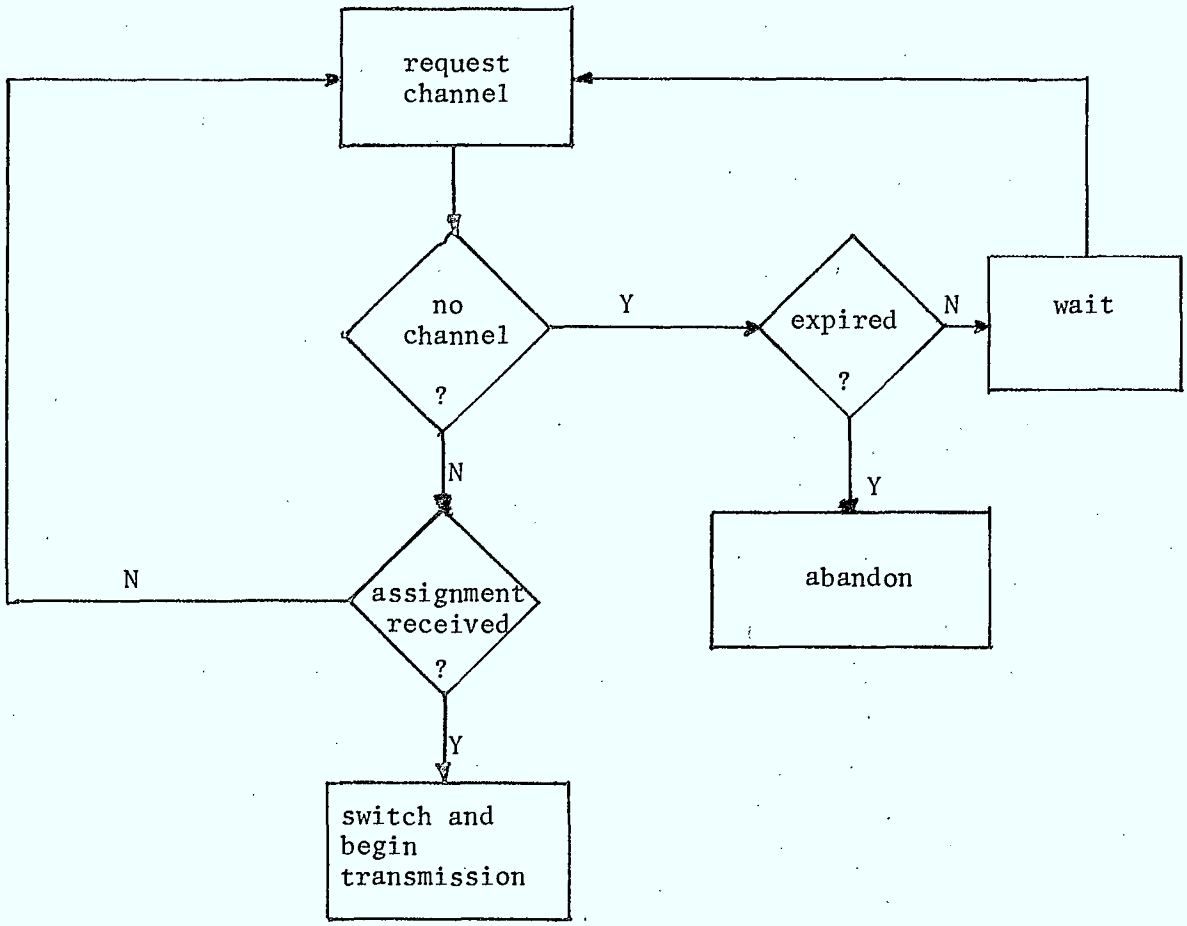


Fig. 4.1(a): SIMULATION FLOWCHART FOR ALOHA NETWORK BASE STATION

originating mobile station flowchart



receiving mobile flowchart

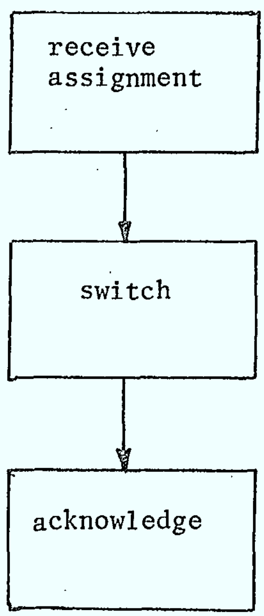


Fig. 4.1(b): SIMULATION FLOW CHART FOR ALOHA NETWORK - ORIGINATING AND RECEIVING MOBILE

Another parameter which was constant over the simulations was the mean fallback after a collision, which was uniformly distributed over (0,100) msec. This must be decided on a pseudo random basis at the transmitting station, in order to avoid repeated collisions due to a fixed retrial period.

Table 4.1 shows the performance of the ALOHA system for various ALOHA packet durations and system congestion levels. The uplink packet durations vary over a considerable range. The smaller values may be justified by the use of several, rather than a single, ALOHA channels. Furthermore, the down-link channel may be used to perform some time-slotting, with a result of cutting the effective packet length by nearly half. (A full 50% reduction is unachievable due to the variable round-trip delays between the base station and the various mobiles.) An estimate of the packet duration might be about 12 bits of message, 20 bits of preamble, adding to 32 bits for a total packet length of 2 msec. Taking advantage of slotting, this could mean an effective length of little more than 1 msec.

The table shows, for each combination of traffic level (designated p) in erlangs and system size, both blocking probability and mean packet delay in milliseconds.

ALOHA METHOD (10 ms switching; 1.3 sec talkspurts)

p	uplink packet	n=8 p(b)	n=8 del	n=16 p(b)	n=16 del	n=32 p(b)	n=32 del	n=64 p(b)	n=64 del	n=128 p(b)	n=128 del
.6	0.1 ms	0.41%	.012	0.00%	.008	0.00%	.007	0.00%	.008	0.00%	.008
	0.4	0.45	.013	0.00	.008	0.00	.007	0.00	.008	0.00	.009
	1.6	0.27	.014	0.00	.011	0.00	.013	0.00	.019	0.003	.035
	6.4	0.14	.022	0.00	.024	0.01	.041	-	-	-	-
.65	0.1	0.88	.015	0.02	.008	0.00	.007	0.00	.008	0.00	.009
	0.4	0.67	.017	0.10	.009	0.00	.009	0.00	.010	0.00	.013
	1.6	0.98	.019	0.02	.012	0.00	.014	0.00	.021	0.00	.039
	6.4	1.15	.034	0.06	.026	0.04	.045	-	-	-	-
.7	0.1	1.35	.019	0.19	.011	0.00	.007	0.00	.008	0.00	.009
	0.4	1.66	.023	0.16	.011	0.00	.009	0.00	.010	0.00	.014
	1.6	1.93	.026	0.08	.015	0.00	.014	0.00	.021	0.01	.042
	6.4	1.70	.057	0.14	.031	0.08	.050	-	-	-	-
.75	0.1	-	-	0.20	.015	0.03	.008	0.00	.008	0.00	.009
	0.4	-	-	0.31	.018	0.04	.011	0.00	.011	0.00	.010
	1.6	-	-	0.58	.022	0.02	.016	0.00	.022	0.02	.040
	6.4	-	-	0.58	.040	0.16	.054	-	-	-	-
.8	0.1	-	-	0.93	.023	0.08	.012	0.004	.009	0.00	.009
	0.4	-	-	0.92	.026	0.15	.014	0.03	.012	0.00	.003
	1.6	-	-	0.81	.027	0.17	.021	0.00	.024	0.03	.051
	6.4	-	-	1.41	.050	0.61	.065	-	-	-	-
.85	0.1	-	-	1.65	.032	0.44	.020	0.13	.013	0.01	.010
	0.4	-	-	2.27	.039	0.47	.024	0.05	.015	0.01	.016
	1.6	-	-	1.90	.039	0.71	.032	0.07	.031	0.05	.059
	6.4	-	-	2.23	.060	1.28	.078	-	-	-	-

Table 4.1: Performance of ALOHA SYSTEMS for Various Packet Lengths and Congestion Levels

In this, and all the results which follow, runs which indicate a blocking probability greater than two percent are not shown. The dashes in the tables thus indicate a non-functional configuration. All runs are based on ten minutes of simulated time, representing from 2000 to 50000 completed messages. Despite the rather large population, considerable statistical fluctuation is evident on these tables that occasionally masks the trends being examined, particularly with respect to the blocking probability value. The reason for this fluctuation is that blocking events tend to come in bunches, when the system is fully loaded. A blocking situation may be very short or relatively long, depending on how many users enter into contention immediately after system saturation and before active channels are relinquished.

The table shows that performance nearly double that of non-interpolated assignment (that is 0.4 erlangs) can be achieved with a channel pool as small as 16 channels. Furthermore, as might be expected, performance improves with increasing pool size. An apparent anomaly shows up at longer packet lengths for large systems, where performance breaks down. Figure 4.2 plots the performance of the ALOHA systems for a packet length which illustrates these phenomena.

Table 4.2 shows the results of ALOHA simulations in which the mean talkspurt duration was taken to be 2.0 seconds rather than 1.3 seconds. The ratio of speech to silence was kept at 2 to 3. As expected, ALOHA system performance improves, due to reduced ALOHA channel loading.

Table 4.3 shows the results of ALOHA simulations where the switching time was taken to be uniformly distributed over (0,60) msec, rather than (0,10) msec, as in most of the simulations. Some deterioration in performance is noted.

Table 4.4 shows simulations of ALOHA performance in the presence of fading. Downlink message failure rates as high as 15% were simulated. Increased mean delay is evident, but any increase in blocking probability is not particularly apparent compared to the statistical fluctuations.

Figure 4.2 shows an output from the ALOHA simulation. A switching time of 60 msec was chosen so that its effect would be obvious in the delay histogram. This run simulates a rather heavily loaded system, so the effect of collisions can be seen.

ALOHA METHOD

```
channels:                                 32
users:                                    68

end of simulation at:             5.000000000E+02
statistics cleared at:            2.000000000E+02

mean collisiion fallback:         5.000000075E-02
rerequest after blocking:         5.000000075E-02
age of abandoned jobs:            5.000000000E-01
fade probability:                 5.000000075E-02
mean pause after abandonment:     1.000000000E+00
uplink packet length:             1.000000047E-03
mean channel switch time:         2.999999933E-02
base station timeout:             9.999999776E-03
mean time between talkspurts:     1.950000048E+00
mean transmission duration:       1.299999952E+00

end of simulation at 5.000000000E+02
system time =                     4.221733398E+03
active time =                     3.956291992E+03
number of jobs completed =              3093
number of messages colliding =           245
number of jobs blocked =                3370
number of jobs abandoned =               102
length of observation =           3.000000000E+02

probability of abandonment =  3.192488104E-02
mean delay =  8.582004905E-02

total jobs in system =           9 of which          9 active
total age in system =  2.056896973E+01 of which  2.004095459E+01 active

  0 0.000E+00   53  ***
  1 1.000E-02  186  ***********
  2 2.000E-02  288  *****************
  3 3.000E-02  416  *************************
  4 4.000E-02  538  ********************************
  5 5.000E-02  677  ****************************************
  6 6.000E-02   85  *****
  7 7.000E-02   40  **
  8 8.000E-02   43  ***
  9 9.000E-02   54  ***
 10 1.000E-01   64  ****
```

ALOHA METHOD (10 ms switching; 2 sec talkspurts)

p	uplink packet	n=8 p(b)	n=8 del	n=16 p(b)	n=16 del	n=32 p(b)	n=32 del	n=64 p(b)	n=64 del	n=128 p(b)	n=128 del
.6	0.4 ms	1.02%	.015	0.00%	.008	0.00%	.007	0.00%	.009	0.00%	.011
	1.6	0.77	.012	0.17	.011	0.00	.011	0.00	.015	0.00	.023
	6.4	0.56	.019	0.10	.021	0.00	.029	0.18	.059	-	-
.65	0.4	0.85	.015	0.03	.009	0.02	.008	0.00	.009	0.00	.011
	1.6	0.96	.018	0.10	.011	0.00	.012	0.00	.016	0.00	.025
	6.4	1.37	.025	0.03	.021	0.00	.030	0.42	.067	-	-
.7	0.4	2.80	.028	0.41	.012	0.00	.008	0.00	.009	0.00	.011
	1.6	2.35	.023	0.35	.015	0.03	.013	0.00	.016	0.00	.027
	6.4	3.19	.033	0.47	.028	0.03	.013	0.66	.074	-	-
.75	0.4	-	-	0.50	.015	0.04	.010	0.00	.010	0.00	.011
	1.6	-	-	0.76	.020	0.07	.015	0.00	.018	0.00	.029
	6.4	-	-	1.08	.033	0.18	.037	1.70	.086	-	-
.8	0.4	-	-	2.62	.028	0.19	.013	0.00	.010	0.00	.012
	1.6	-	-	2.04	.034	0.30	.020	0.04	.019	0.00	.031
	6.4	-	-	2.01	.044	0.33	.042	0.82	.017	-	-
.85	0.4	-	-	-	-	1.13	.024	0.15	.014	0.02	.013
	1.6	-	-	-	-	1.23	.031	0.31	.026	0.03	.033
	6.4	-	-	-	-	1.68	.057	-	-	-	-

Table 4.2: Performance of ALOHA System for Various Packet Length and Congestion Levels

ALOHA METHOD (60 ms switching; 1.3 sec talkspurts)

p	uplink packet	n=8		n=16		n=32		n=64		n=128	
		p(b)	del	p(b)	del	p(b)	del	p(b)	del	p(b)	del
.65	0.1msec	0.84%	.050	0.06%	.042	0.00%	.041	0.00%	.041	0.00%	.042
	0.4	1.39	.052	0.06	.043	0.00	.047	0.00	.043	0.00	.047
	1.6	1.61	.054	0.10	.047	0.00	.047	0.00	.047	0.005	.074
.7	0.1	-	-	0.20	.046	0.00	.041	0.00	.041	0.00	.042
	0.4	-	-	0.12	.046	0.00	.043	0.00	.043	0.00	.047
	1.6	-	-	0.22	.050	0.00	.049	0.00	.056	0.01	.078
.75	0.1	-	-	0.72	.051	0.05	.043	0.00	.042	0.00	.042
	0.4	-	-	0.82	.054	0.04	.044	0.005	.044	0.00	.048
	1.6	-	-	0.51	.057	0.10	.051	0.00	.056	0.00	.081
.8	0.1	-	-	1.07	.061	0.18	.047	0.004	.042	0.00	.042
	0.4	-	-	1.67	.061	0.22	.049	0.02	.045	0.00	.048
	1.6	-	-	1.32	.068	0.22	.055	0.02	.059	0.06	.087
.85	0.1	-	-	-	-	0.80	.057	0.06	.047	0.004	.043
	0.4	-	-	-	-	0.80	.059	0.09	.050	0.00	.049
	1.6	-	-	-	-	0.65	.067	0.13	.066	0.14	.095
.9	0.1	-	-	-	-	1.66	.075	0.52	.059	0.13	.049
	0.4	-	-	-	-	2.25	.078	0.58	.063	0.08	.057
	1.6	-	-	-	-	1.48	.082	0.57	.077	0.64	.108

Table 4.3: Performance of ALOHA System for Increased Switching Time

ALOHA METHOD (10 ms switching; 2 sec talkspurts)

p	fade prob.	n=8 p(b)	n=8 del	n=16 p(b)	n=16 del	n=32 p(b)	n=32 del	n=64 p(b)	n=64 del	n=128 p(b)	n=128 del
.65	0	0.98%	.019	0.02%	.012	0.00%	.014	0.00%	.021	0.00%	.039
	0.05	0.50	.022	0.08	.016	0.00	.017	0.00	.023	0.003	.042
	0.10	0.38	.022	0.00	.017	0.00	.019	0.00	.027	0.003	.045
	0.15	0.55	.026	0.06	.022	0.00	.023	0.00	.029	0.01	.049
.7	0	1.93	.026	0.08	.015	0.00	.014	0.00	.021	0.00	.042
	0.05	1.36	.026	0.29	.019	0.00	.017	0.00	.024	0.01	.046
	0.10	1.46	.031	0.19	.022	0.00	.021	0.00	.027	0.01	.049
	0.15	1.75	.033	0.35	.025	0.00	.024	0.00	.030	0.01	.052
.75	0	-	-	0.58	.022	0.02	.016	0.00	.022	0.02	.040
	0.05	-	-	0.31	.023	0.05	.020	0.00	.026	0.02	.050
	0.10	-	-	0.49	.025	0.06	.023	0.01	.028	0.02	.053
	0.15	-	-	0.49	.030	0.12	.027	0.00	.033	0.02	.057
.8	0	-	-	0.81	.027	0.17	.021	0.00	.024	0.03	.051
	0.05	-	-	1.40	.034	0.30	.025	0.01	.028	0.04	.056
	0.10	-	-	1.69	.040	0.18	.028	0.01	.031	0.03	.058
	0.15	-	-	1.86	.042	0.21	.032	0.01	.034	0.04	.063
.85	0	-	-	-	-	0.71	.032	0.07	.031	0.05	.059
	0.05	-	-	-	-	0.55	.035	0.07	.033	0.08	.061
	0.10	-	-	-	-	0.61	.037	0.10	.035	0.07	.064
	0.15	-	-	-	-	0.60	.040	0.10	.040	0.14	.068
.9	0	-	-	-	-	1.53	.043	0.41	.039	0.41	.071
	0.05	-	-	-	-	1.81	.052	0.55	.046	0.30	.072
	0.10	-	-	-	-	1.06	.048	0.55	.048	0.36	.077
	0.15	-	-	-	-	1.91	.056	0.48	.052	0.40	.078

Table 4.4: Performance of ALOHA in Presence of Fading

4.3 Polling Simulation

The logical aspects of the polling simulation were similar to those in the ALOHA approach, so that Figure 4.1 remains valid in this instance. Indeed, the simulation was rather simplified by the absence of any collision events. Figure 4.3 illustrates a typical run, representing circumstances similar to those of Figure 4.2. A rather long polling cycle is illustrated. It is clear that this polling can contribute significantly to the delay experienced.

Table 4.5 shows the results of simulations of a polling network with a 10 millisecond maximum channel switching time, and a 1.3 second mean talkspurt duration. Performance is seen to be similar to that of the ALOHA based system, though not, of course, subject to collisions, as far as the probability of blocking is concerned. However, for networks which are both geographically and numerically large, the expected delay becomes unacceptable. Assuming no guard time and zero pulse width, a 160 microsecond polling interval implies a maximum diameter of about 25 kilometres.

Table 4.6 shows that, as expected, changing the talkspurt distribution while maintaining traffic intensity has no significant effect on the polling system performance. Table 4.7

RES

POLLING METHOD


```
channels:                          32
users:                             68

end of simulation at:          4.000000000E+02
statistics cleared at:         2.000000000E+02

mean collision fallback:       5.000000075E-02
rerequest after blocking:      5.000000075E-02
age of abandoned jobs:         5.000000000E-01
fade probability:              5.000000075E-02
mean pause after abandonment:  1.000000000E+00
uplink packet length:          9.999999747E-05
mean channel switch time:      2.999999933E-02
base station timeout:          9.999999776E-03
mean time between talkspurts:  1.950000048E+00
mean transmission duration:    1.299999952E+00

end of simulation at 4.000000000E+02
system time =                  2.762555664E+03
active time =                  2.601542725E+03
number of jobs completed =          2009
number of jobs blocked =            1590
number of jobs abandoned =            53
length of observation =        2.000000000E+02

probability of abandonment =  2.570320107E-02
mean delay =  8.014581352E-02

total jobs in system =          14 of which         13 active
total age in system =  1.701034546E+01 of which  1.633111572E+01 active

  0 0.000E+00   14  *
  1 1.000E-02   76  *******
  2 2.000E-02  161  ***************
  3 3.000E-02  257  ************************
  4 4.000E-02  309  *****************************
  5 5.000E-02  437  *****************************************
  6 6.000E-02  314  ******************************
  7 7.000E-02   26  **
  8 8.000E-02   21  **
  9 9.000E-02   21  **
 10 1.000E-01   32  ***
 11 1.100E-01   38  ***
```

POLLING METHOD (10 ms switching; 1.3 sec talkspurts)

p	uplink packet	n=8 p(b)	n=8 del	n=16 p(b)	n=16 del	n=32 p(b)	n=32 del	n=64 p(b)	n=64 del	n=128 p(b)	n=128 del
.6	10 usec	0.36%	.011	0.00%	.007	0.00%	.007	0.00%	.009	0.00%	.010
	40	0.45	.012	0.00	.009	0.00	.011	0.00	.014	0.00	.022
	160	0.31	.014	0.00	.014	0.00	.022	0.00	.038	0.00	.068
	640	0.28	.028	0.00	.037	0.00	.068	0.00	.129	0.00	.251
.65	10	0.79	.014	0.08	.009	0.00	.008	0.00	.009	0.00	.010
	40	0.37	.015	0.06	.011	0.01	.011	0.00	.015	0.00	.023
	160	0.62	.022	0.00	.015	0.00	.023	0.00	.040	0.00	.073
	640	0.74	.032	0.06	.041	0.00	.073	0.00	.141	0.00	.273
.7	10	1.55	.021	0.27	.013	0.00	.008	0.00	.009	0.00	.011
	40	1.07	.024	0.10	.012	0.00	.012	0.00	.016	0.00	.025
	160	1.31	.024	0.12	.020	0.04	.025	0.00	.042	0.00	.078
	640	1.41	.039	0.14	.047	0.01	.079	0.00	.149	0.00	.293
.75	10	-	-	0.53	.016	0.15	.011	0.01	.009	0.00	.011
	40	-	-	0.36	.019	0.02	.013	0.00	.016	0.00	.026
	160	-	-	0.58	.027	0.04	.028	0.00	.045	0.00	.084
	640	-	-	0.24	.054	0.00	.084	0.00	.159	0.00	.314
.8	10	-	-	1.32	.025	0.16	.014	0.00	.010	0.00	.012
	40	-	-	1.13	.028	0.15	.018	0.00	.018	0.00	.027
	160	-	-	0.83	.035	0.14	.033	0.01	.048	0.00	.089
	640	-	-	1.17	.064	0.26	.094	0.004	.171	0.00	.334
.85	10	-	-	-	-	0.29	.020	0.04	.012	0.00	.013
	40	-	-	-	-	0.57	.025	0.04	.022	0.00	.029
	160	-	-	-	-	0.33	.040	0.05	.054	0.00	.094
	640	-	-	-	-	0.85	.108	0.08	.184	0.004	.355

Table 4.5: Performance of POLLING System for Various Packet Lengths and Congestion Levels

POLLING METHOD (10 ms switching; 2 sec talkspurts)

p	uplink packet	n=8 p(b)	n=8 del	n=16 p(b)	n=16 del	n=32 p(b)	n=32 del	n=64 p(b)	n=64 del	n=128 p(b)	n=128 del
.6	10 usec	0.62%	.011	0.04%	.008	0.00%	.008	0.00%	.009	0.00%	.010
	40	0.49	.015	0.00	.009	0.00	.010	0.00	.014	0.00	.022
	160	0.33	.015	0.00	.014	0.00	.022	0.00	.037	0.00	.068
.65	10	0.56	.019	0.34	.009	0.00	.008	0.00	.009	0.00	.011
	40	1.98	.021	0.26	.011	0.00	.011	0.00	.015	0.00	.023
	160	1.47	.020	0.00	.016	0.00	.023	0.00	.040	0.00	.073
.7	10	-	-	0.38	.012	0.07	.009	0.00	.009	0.00	.011
	40	-	-	0.18	.012	0.03	.012	0.00	.015	0.00	.024
	160	-	-	0.42	.020	0.03	.025	0.00	.042	0.00	.078
.75	10	-	-	1.58	.022	0.11	.011	0.00	.009	0.00	.014
	40	-	-	0.86	.018	0.11	.014	0.00	.016	0.00	.025
	160	-	-	1.03	.029	0.09	.028	0.00	.045	0.00	.084
.8	10	-	-	-	-	0.53	.015	0.00	.010	0.00	.012
	40	-	-	-	-	0.23	.018	0.03	.018	0.00	.027
	160	-	-	-	-	0.12	.032	0.00	.048	0.00	.088
.85	10	-	-	-	-	1.37	.027	0.15	.014	0.00	.013
	40	-	-	-	-	0.23	.018	0.03	.018	0.004	.029
	160	-	-	-	-	1.41	.051	0.07	.054	0.00	.094

Table 4.6: Performance of POLLING System for Various Packet Lengths and Congestion Levels

POLLING METHOD (60 ms switching; 1.3 sec talkspurts)

p	uplink packet	n=8 p(b)	n=8 del	n=16 p(b)	n=16 del	n=32 p(b)	n=32 del	n=64 p(b)	n=64 del	n=128 p(b)	n=128 del
.6	10 usec	0.31%	.047	0.00%	.040	0.00%	.041	0.00%	.042	0.00%	.044
	40	0.13	.045	0.00	.042	0.00	.044	0.00	.049	0.00	.055
	160	0.41	.050	0.00	.048	0.00	.055	0.00	.071	0.00	.101
.65	10	0.83	.048	0.06	.042	0.00	.044	0.00	.042	0.00	.044
	40	0.96	.051	0.04	.044	0.00	.044	0.00	.048	0.00	.056
	160	1.10	.053	0.15	.051	0.00	.057	0.00	.073	0.00	.106
.7	10	1.08	.056	0.33	.047	0.01	.042	0.00	.042	0.00	.044
	40	1.31	.054	0.16	.046	0.00	.045	0.00	.049	0.00	.058
	160	1.76	.063	0.20	.054	0.00	.058	0.00	.076	0.00	.111
.75	10	-	-	0.62	.053	0.04	.044	0.00	.043	0.00	.045
	40	-	-	0.93	.056	0.01	.047	0.00	.050	0.00	.059
	160	-	-	0.77	.061	0.03	.062	0.00	.078	0.00	.117
.8	10	-	-	1.39	.060	0.21	.050	0.02	.044	0.00	.045
	40	-	-	1.18	.061	0.17	.054	0.00	.051	0.00	.061
	160	-	-	1.14	.071	0.17	.066	0.01	.082	0.00	.122
.85	10	-	-	-	-	0.50	.056	0.02	.048	0.004	.047
	40	-	-	-	-	0.62	.062	0.08	.056	0.01	.063
	160	-	-	-	-	0.40	.076	0.11	.090	0.00	.127

Table 4.7: Performance of POLLING System for Various Packet Lengths and Congestion Levels

POLLING METHOD (10 ms switching; 1.3 sec talkspurts)

p	fade prob.	n=8		n=16		n=32		n=64		n=128	
		p(b)	del	p(b)	del	p(b)	del	p(b)	del	p(b)	del
.65	0	0.37%	.015	0.06%	.011	0.01%	.011	0.00%	.015	0.00%	.023
	0.05	0.50	.018	0.04	.014	0.00	.014	0.00	.118	0.00	.026
	0.10	0.63	.022	0.08	.016	0.00	.017	0.00	.021	0.00	.029
	0.15	1.06	.028	0.02	.019	0.00	.020	0.00	.024	0.00	.032
.7	0	1.07	.024	0.10	.012	0.00	.012	0.00	.016	0.00	.025
	0.05	1.46	.025	0.14	.015	0.00	.014	0.00	.019	0.00	.028
	0.10	0.96	.028	0.10	.018	0.01	.018	0.00	.021	0.00	.031
	0.15	1.81	.032	0.20	.022	0.00	.020	0.00	.025	0.00	.034
.75	0	-	-	0.36	.014	0.02	.013	0.00	.016	0.00	.026
	0.05	-	-	0.42	.020	0.05	.017	0.00	.019	0.00	.029
	0.10	-	-	0.57	.023	0.01	.020	0.00	.023	0.00	.032
	0.15	-	-	0.30	.028	0.01	.024	0.005	.025	0.00	.035
.8	0	-	-	1.13	.028	0.15	.018	0.00	.018	0.00	.027
	0.05	-	-	1.22	.028	0.13	.019	0.00	.021	0.00	.030
	0.10	-	-	1.01	.033	0.07	.023	0.00	.024	0.00	.033
	0.15	-	-	0.88	.038	0.21	.026	0.01	.027	0.00	.036
.85	0	-	-	-	-	0.57	.025	0.04	.012	0.00	.029
	0.05	-	-	-	-	0.42	.031	0.08	.026	0.00	.032
	0.10	-	-	-	-	0.38	.033	0.09	.028	0.002	.035
	0.15	-	-	-	-	0.48	.033	0.09	.028	0.00	.038
.9	0	-	-	-	-	1.34	.040	0.38	.034	0.01	.032
	0.05	-	-	-	-	1.19	.041	0.21	.031	0.02	.036
	0.10	-	-	-	-	1.81	.049	0.44	.039	0.04	.040
	0.15	-	-	-	-	1.18	.049	0.30	.039	0.04	.043

Table 4.8: Performance of POLLING Network in the Presence of Fading

shows the performance with switching time uniformly distributed on (0,60) milliseconds, rather than on (0,10) as in most runs. Nothing surprising emerges. Blocking probabilities are similar, and delay times increase an average of 25 milliseconds. This compounds the problems of the larger systems.

Finally, Table 4.8 shows the performance of the polling network in the presence of fading. A 40 microsecond polling slot is assumed. Performance is seen to be robust in the presence of a loss of downlink command messages at a rate of 15%.

4.4 Table-Driven System Simulation

Figure 4.4 shows a flowchart of the process simulated by the table-driven network program. While the decision process illustrated seems simple, much complexity is masked in the channel acquisition block. In order to fit the simulation in the limited memory of the PDP-11, and to speed execution as well, the actual maintenance of system maps at all mobile stations was not simulated. Rather, a decision was made about the age of the local map at the time of detection of a speech burst. The map was then set to be the correct map at the time of last successful refresh.

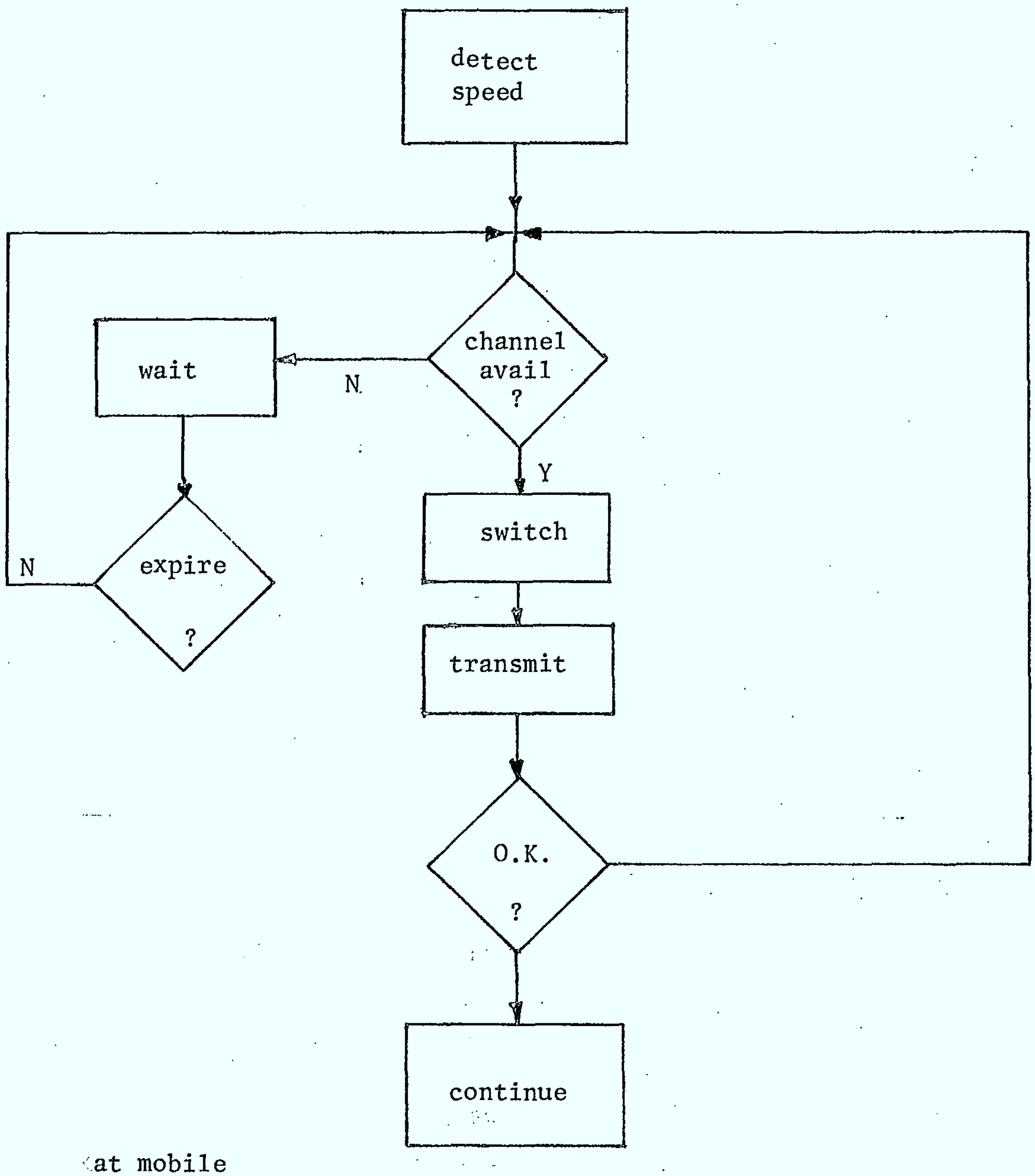
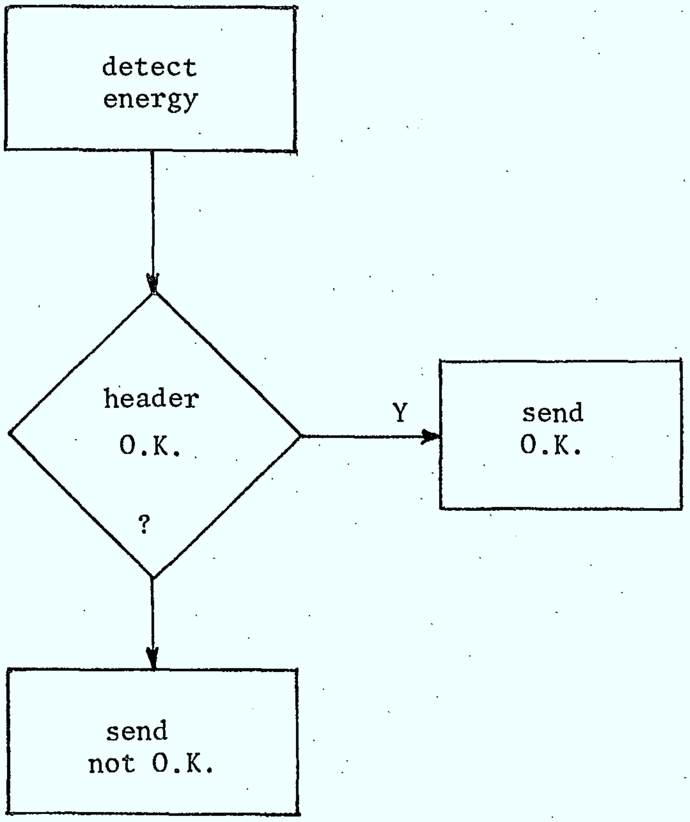


Fig. 4.4(a): FLOW CHART FOR TABLE DRIVEN SIMULATION



at base

Fig. 4.4(b): SIMULATION FLOW CHART FOR TABLE DRIVEN NETWORK

Also, no additional checks were simulated during the switching interval, although in practice this would probably be a good idea if no interference due to the channel switching process were anticipated. The expectation was that this simplification would penalize the performance of systems with a long switching time, but this turned out not to be the case.

As simulated, the table-driven system shows some instability. In particular, the "header length" is seen in Table 4.9 to be a rather critical performance parameter. This is defined as the time between physical acquisition of a channel and the confirmation of the assignment by the base station. A long header greatly increases the chance that the headers will be subject to collision, and that hence neither user will have the channel allocated. Thus both will continue to search a possibly limited set of free channels, with a high probability of colliding again.

Despite this problem, it is seen that quite acceptable performance can be achieved relative to non-interpolated assignment.

Table 4.10 shows the performance of the same set of systems with talkspurt durations averaging 2.0 seconds rather than 1.3. There is a surprising indication of deteriorated

TABLE METHOD (10 ms switching; 1.3 sec talkspurts)

p	header length	n=8 p(b)	n=8 del	n=16 p(b)	n=16 del	n=32 p(b)	n=32 del	n=64 p(b)	n=64 del	n=128 p(b)	n=128 del
.6	1 msec	2.49%	.024	0.27%	.016	0.01%	.011	0.00%	.011	0.00%	.011
	4	3.10	.031	0.23	.019	0.02	.015	0.04	.015	0.03	.015
	10	5.13	.041	3.26	.032	3.34	.029	3.76	.030	4.47	.029
.65	1	-	-	0.66	.018	0.00	.012	0.00	.011	0.00	.011
	4	-	-	0.82	.023	0.01	.016	0.01	.015	0.05	.015
.7	1	-	-	-	-	0.15	.016	0.05	.012	0.00	.011
	4	-	-	-	-	0.19	.019	0.19	.016	0.16	.016
.75	1	-	-	-	-	0.32	.021	0.03	.013	0.00	.011
	4	-	-	-	-	0.63	.023	0.80	.018	0.45	.017
.8	1	-	-	-	-	0.88	.030	0.18	.017	0.02	.012
	4	-	-	-	-	2.69	.033	1.39	.021	-	-
.85	1	-	-	-	-	-	-	0.50	.026	0.04	.015
	4	-	-	-	-	-	-	6.58	.028	-	-

Table 4.9: Performance of Table Driven System for Various Packet Lengths and Congestion Levels

TABLE METHOD (60 ms switching; 1.3 sec talkspurts)

p	header length	n=8 p(b)	n=8 del	n=16 p(b)	n=16 del	n=32 p(b)	n=32 del	n=64 p(b)	n=64 del	n=128 p(b)	n=128 del
.6	1 msec	3.67%	.075	0.36%	.068	0.01%	.063	0.00%	.062	0.00%	.062
	4	3.27	.077	0.55	.071	0.05	.066	0.00	.066	0.00	.066
	16	3.27	.095	0.82	.087	0.25	.084	0.32	.084	0.00	.084
.65	1	-	-	0.72	.073	0.12	.065	0.00	.063	0.00	.063
	4	-	-	1.12	.076	0.08	.069	0.01	.067	0.00	.067
	16	-	-	1.25	.092	0.48	.089	0.67	.088	0.84	.088
.7	1	-	-	1.87	.078	0.26	.068	0.01	.065	0.00	.065
	4	-	-	1.66	.082	0.17	.071	0.01	.070	0.00	.069
	16	-	-	2.72	.101	1.05	.093	1.81	.095	2.21	.094
.75	1	-	-	-	-	0.63	.074	0.10	.068	0.01	.066
	4	-	-	-	-	1.02	.081	0.09	.073	0.01	.072
	16	-	-	-	-	-	-	-	-	-	-
.8	1	-	-	-	-	1.44	.085	0.27	.074	0.03	.070
	4	-	-	-	-	1.80	.092	0.30	.081	0.04	.078
.85	1	-	-	-	-	-	-	0.95	.084	0.11	.077
	4	-	-	-	-	-	-	1.38	.093	0.54	.088

Table 4.10: Performance of Table Driven System for Various Packet Lengths and Congestion Levels

performance at smaller network sizes. The larger networks perform better than in the case of more frequent switching, as expected.

Table 4.11 is also surprising, in that no deterioration is shown with respect to blocking probability in a system with a much higher switching time. However, closer examination of these runs has revealed that the collision probability is much higher in the cases where faster switching is possible. Thus a "chain reaction" type of situation is set up, in which channel acquisition problems proliferate among all users, while in a system with a slower switching time, abandonment is much more likely.

Table 4.12 shows the relationship between table-driven system performance and fading probability. The most noticeable thing is the large fluctuation from run to run, with relatively little correspondence between fading probability and blocking rate. Indeed, in one case, the simulation "blew up" for a case of zero fade (an exit trap was provided for simulations where blocking exceeded 25%) while results were obtained with higher fade values. This situation was not repeated for an identical run with a different random number seed. In that case the result was comparable to the others in the cell.

TABLE METHOD (10 ms switching; 2 sec talkspurts)

p	header length	n=8 p(b)	n=8 del	n=16 p(b)	n=16 del	n=32 p(b)	n=32 del	n=64 p(b)	n=64 del	n=128 p(b)	n=128 del
.6	1 msec	4.13%	.023	0.35	.014	0.02	.012	0.00	.011	0.00	.011
	4	4.91	.027	0.49	.018	0.07	.014	0.00	.014	0.00	.014
	16	5.54	.042	3.41	.034	2.16	.029	2.72	.029	-	-
.65	1	-	-	1.22	.020	0.10	.012	0.00	.011	0.00	.011
	4	-	-	1.84	.024	0.06	.024	0.00	.014	0.00	.015
.7	1	-	-	-	-	0.30	.016	0.00	.012	0.00	.011
	4	-	-	-	-	0.39	.019	0.00	.015	0.00	.015
.75	1	-	-	-	-	0.97	.022	0.11	.013	0.00	.012
	4	-	-	-	-	1.03	.026	0.17	.017	0.09	.015
.8	1	-	-	-	-	-	-	0.34	.018	0.05	.013
	4	-	-	-	-	-	-	0.98	.022	1.12	.017

Table 4.11: Performance of Table Driven System for Various Packet Lengths and Congestion Levels

TABLE METHOD (10 ms switching; 1.3 sec talkspurts)

p	fade	n=8		n=16		n=32		n=64		n=128	
	prob.	p(b)	del	p(b)	del	p(b)	del	p(b)	del	p(b)	del
.6	0	3.10	.031	0.23	.019	0.02	.015	0.04	.015	0.03	.015
	0.05	2.67	.027	0.34	.019	0.00	.015	0.00	.015	0.02	.015
	0.10	2.48	.030	0.32	.021	0.00	.016	0.00	.016	0.05	.029
	0.15	2.40	.029	0.58	.020	0.01	.017	0.02	.016	0.01	.016
.65	0	-	-	0.82	.023	0.01	.016	0.01	.015	0.05	.015
	0.05	-	-	0.88	.021	0.05	.017	0.03	.016	0.04	.015
	0.10	-	-	0.92	.025	0.15	.018	0.02	.016	0.03	.016
	0.15	-	-	1.20	.026	0.03	.017	0.02	.017	0.03	.017
.7	0	-	-	-	-	0.19	.019	0.19	.016	0.16	.016
	0.05	-	-	-	-	0.25	.019	0.15	.017	0.47	.017
	0.10	-	-	-	-	0.34	.021	0.04	.017	0.04	.016
	0.15	-	-	-	-	0.30	.022	0.07	.018	0.13	.018
.75	0	-	-	-	-	0.63	.023	0.80	.018	0.45	.017
	0.05	-	-	-	-	0.87	.025	0.53	.018	0.52	.019
	0.10	-	-	-	-	0.88	.027	0.49	.018	0.19	.017
	0.15	-	-	-	-	0.73	.026	0.42	.020	0.51	.019
.8	0	-	-	-	-	2.69	.033	1.39	.021	-	-
	0.05	-	-	-	-	2.19	.033	1.85	.022	2.00	.020
	0.10	-	-	-	-	1.87	.034	2.09	.022	0.54	.018
	0.15	-	-	-	-	1.90	.035	1.49	.023	1.34	.021

Table 4.12: Performance of Table Driven System in Presence of Fading

The conclusion is that the behaviour of the table-driven system is both complicated and unstable. If such a system is to be reliably workable, a more complete protocol will have to be specified and implemented.

4.5 Comparisons and Conclusions

Figures 4.5, 4.6 and 4.7 present a graphic comparison of the three approaches studied, as far as blocking probability is concerned at least. The performances of the two request-driven methods are seen to be very similar in this regard, and considerably better than that of the table-driven method.

The decision between these approaches rests on technological factors and geographic ones. Large cell sizes work strongly against the polling method in terms of delay. This can be ameliorated, of course, by the addition of extra polling channels, so that each user would make requests on only one of a number of polling cycles. Similarly, in situations where the ALOHA channel is the principal performance limitation, extra ALOHA channels may be added. Of course, both of these approaches will lessen the bandwidth advantage which is the main purpose of TASI.

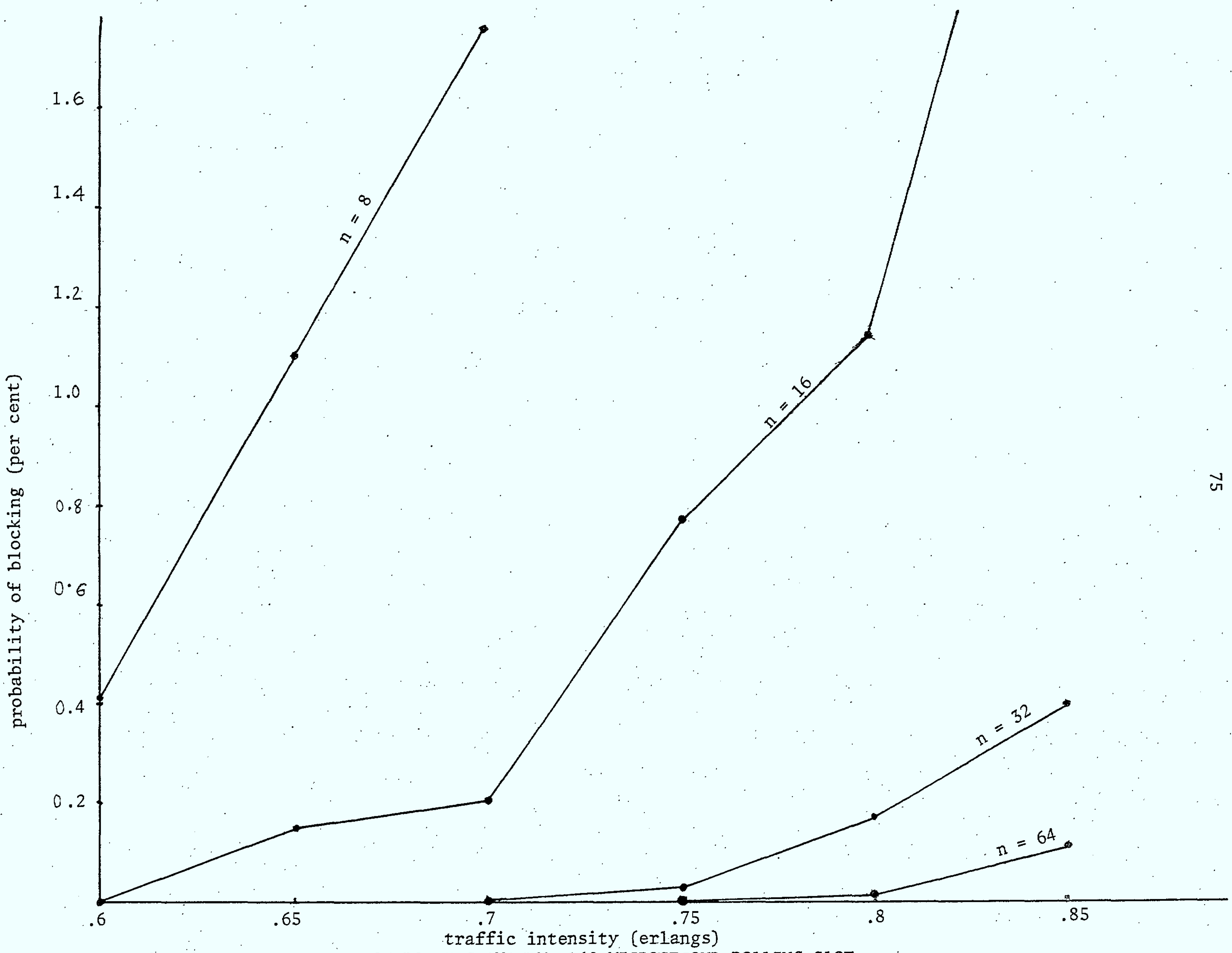


Fig. 4.6: PERFORMANCE OF POLLING SYSTEM, 160 MICROSECOND POLLING SLOT, 60.0 MILLISECOND MAXIMUM SWITCHING TIME

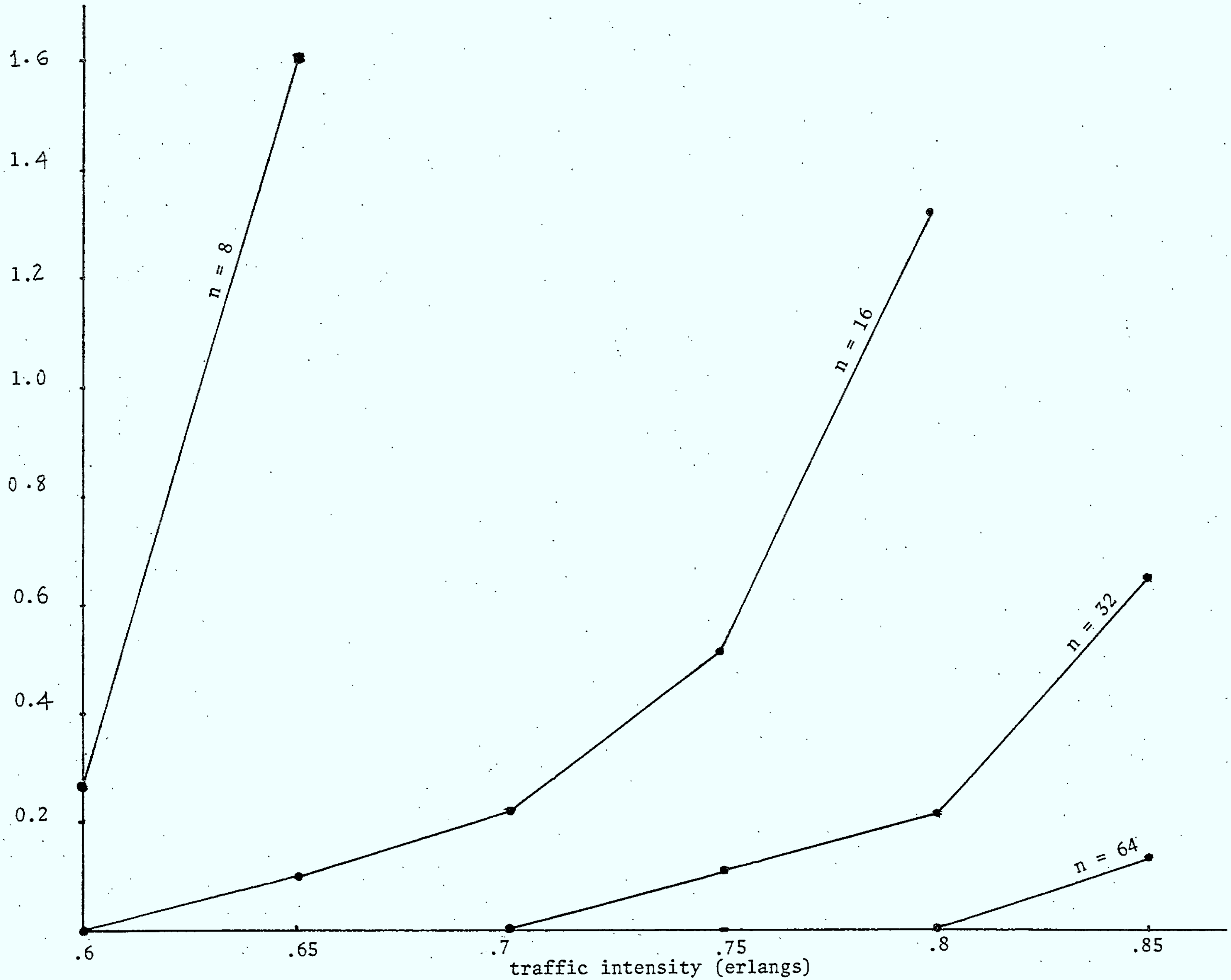


Fig. 4.5: PERFORMANCE OF ALOHA SYSTEM WITH 1.6 MILLISECOND REQUEST PACKETS
60.0 MILLISECOND MAXIMUM SWITCH TIME

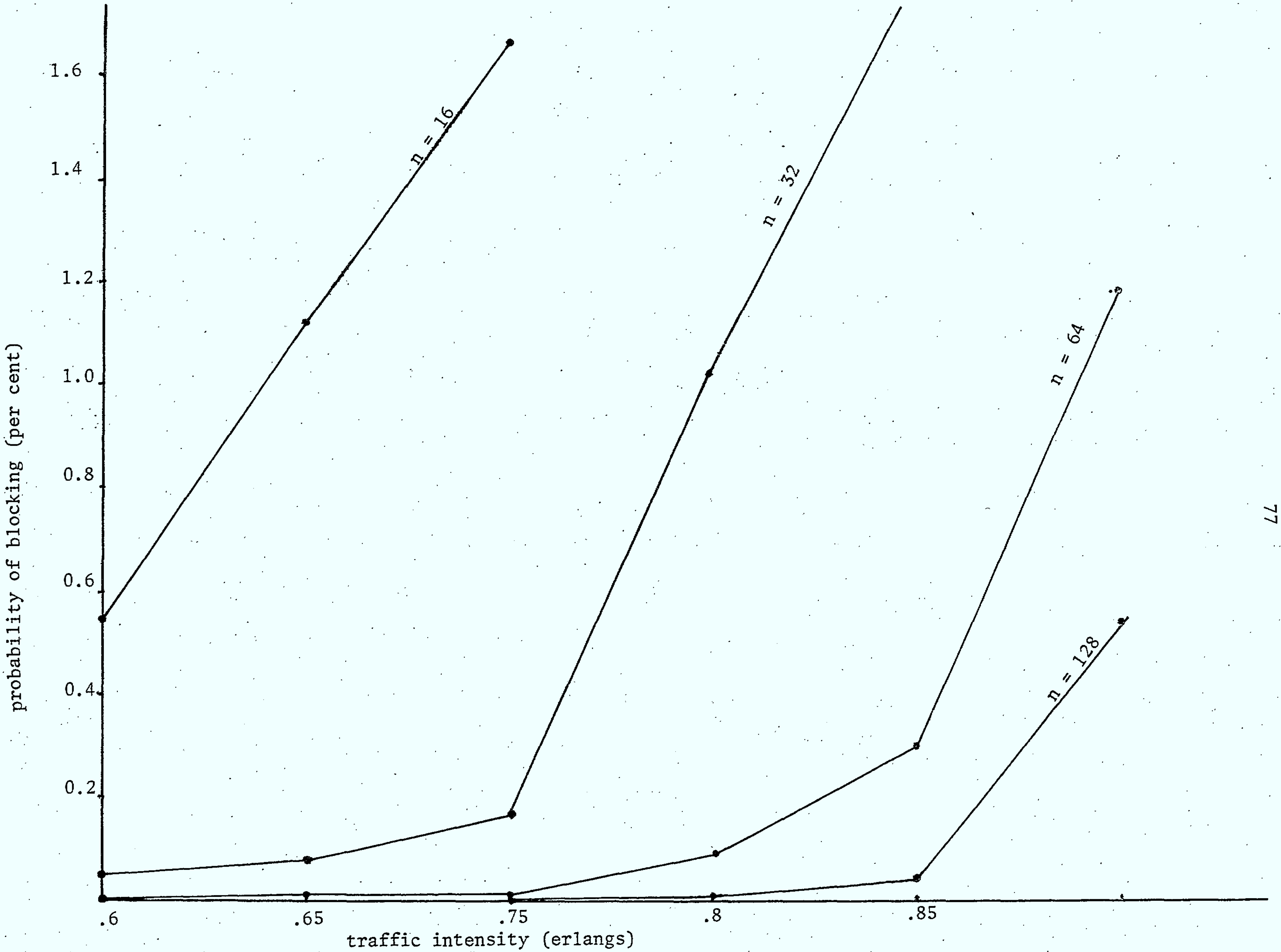


Fig. 4.7: PERFORMANCE OF TABLE DRIVEN SYSTEM 4 msec HEADER TIME
60 MILLISECOND MAXIMUM SWITCHING TIME

The technological considerations for the ALOHA method require development of very fast non-coherent demodulators, minimizing the required preamble. For the polling method, the requirement is for very short, easily detectable pulses which do not require too much bandwidth. Optimal performance of the two methods is comparable.

The principal conclusion of this study is that TASI in the mobile radio environment seems feasible, and that work should continue. The present work could be expanded to include

1) simulation of queueing on the dedicated downlink channel;

2) simulation of varying fade probabilities for various users;

3) call establishment;

4) base-to-mobile and mobile-to-base communication;

5) mixed voice and data packets with data integrity checks, and

6) a more sophisticated table-driven approach.

CHAPTER V

ISSUES RELATING TO DATA COMMUNICATIONS

5.1 Introduction

The network configurations proposed in the previous chapters can be greatly enhanced by including data communication facilities. This can be accomplished by simply reserving a small number of channels for data traffic. Of course the data and voice can be fully integrated on all channels, but that would complicate the design considerably. The data traffic generated by the mobile units is expected to consist of short and infrequent messages, hence one or two radio channels can support a large number of data terminals. Moreover, the bursty nature of the data traffic suggests the use of random access schemes. Slotted ALOHA is difficult to implement in the mobile environment because of the large differential delay between terminals at various locations, while CSMA techniques are more complicated and probably unreliable under fading conditions. This suggests that the simple ALOHA scheme is the most suitable technique. In the AMPS system ALOHA was modified in order to reduce the probability of collision between packets. The modified version can be described as a quasi-CSMA, where the base station sends a control bit periodically to indicate the status of the random access channel ("1" for busy and "0" for idle); the mobile is required to monitor three consecutive control bits and take a majority voting decision as to transmit or wait. Here, we shall restrict the discussion to the simple ALOHA scheme.

Although packet radio is feasible with today's technology, a great deal of work remains to be undertaken before it becomes a realistic practical alternative to message or circuit switching. This is particularly true for mobile systems where frequent signal fading and high error probability severely constrain the selection of packet length, modulation method and access protocol.

In this chapter, we address a number of outstanding problems in designing packet radio systems. It should be noted that these problems affect the control channel throughput of the TASI system discussed previously as well as the mobile data communication in general. An expression for the packet error probability under fading conditions is derived in Section 5.2. In Section 5.3 the channel throughput of ALOHA/ARQ protocol for fading channel is determined. A discussion of the effect of synchronization time on the channel throughput is given in Section 5.4. Then, the chapter is concluded in Section 5.5.

5.2 Packet Error Rate

In a typical land mobile system the envelope of the received signal at the mobile is known to have a Rayleigh probability distribution with a time variation proportional to the product of the vehicular speed and the carrier frequency. At the base station the received signal has different statistical characteristics. The signal is composed of few specular reflections from the buildings in the immediate vicinity of the mobile unit, plus a large number of weak reflections from the clutter surrounding the transmitter. The combined waves can be viewed as a Ricean process and it may be expressed as:

$$S(t) = A(t) \cos [2\pi f_c t + \theta(t)] + n(t) \qquad (1)$$

where $A(t)$ and $\theta(t)$ are the random amplitude and phase of the strong reflections and $n(t)$ is, by the central limit theory, a Gaussian process.

The envelope of the received signal will then vary with time and occasionally drops below a certain decision threshold level given rise of error bursts. The envelope variation at the base station is usually much slower than that at the mobile unit, and a high degree of correlation between the signal levels seen by consecutive bits in a packet does exist. Under these conditions, the assumption of independent errors seems to be invalid. In fact, since the average signal to noise ratio is relatively high, most of the errors occur during the fading intervals, and they tend to occur in bursts. Therefore, one would expect an errored packet to contain more than one error. Then, the packet error rate is upper bounded by:

$$P(E) \leq [1-(1-P_e)^L] \qquad (2)$$

where P_e is the average bit error rate and L is the packet length in bits. The bound represents the worst case of independent errors.

A simple model which can be used to estimate the packet error rate for a given bit error rate under the assumption that errors are not independent is a stationary first-order Markov chain. The model is a first-order approximation to real fading channels, but it represents a considerable improvement over the assumption of independence between transmission.

We consider a sequence of identically distributed random variables X_1, X_2, ..., X_i, ... each of which can take on just two values, "1" if a bit is in error and "0" if the bit is correct. The average bit error rate is:

$$P_e = \text{Prob. } [X_i=1] = 1 - \text{Prob. } [X_i=0] = 1 = q \qquad (3)$$

For a Markovian relation, the errors of consecutive bits are related by:

$$\beta = \text{Prob. } [X_i=1 \mid X_{i-1}=1] \tag{4}$$

where β is constant to be determined by the envelope correlation coefficient and the particular modulation system being used. Under these assumptions, the packet error rate, P(E), is related to the bit error rate by:

$$P(E) = 1 - (1-P_e)\left[1-(1-\beta)\frac{P_e}{1-P_e}\right]^{L-1} \tag{5}$$

The two extreme cases are:

(i) Independent error: $\beta = P_e$

$$P(E) = 1 - (1-P_e)^L \quad \text{(non-fading)}$$

and

(ii) complete correlation: $\beta = 1$

$$P(E) = P_e$$

The factor $Q = 1 - P(E)$ is bounded by:

$$(1-P_e)^L \leq Q \leq (1-P_e) \tag{6}$$

Figure 5.1 shows the relation between a and β for PSK and DPSK modulation systems at SNR = 30 dB. In this figure β was the independent parameter and the calculations were carried according to Eqn. (5). The figure indicates that Q remains fairly constant for values of β below 0.1, then it increases rapidly for larger values of β until it reaches the upper bound of $(1-P_e)$.

Unfortunately, the observable parameters of the system (i.e., SNR, speed, carrier frequency, etc.) are not enough to evaluate the parameter β; what is needed is the joint probability density of consecutive envelope samples, which is very difficult to obtain. For Rayleigh fading channels, however, a simplified model for estimating β was obtained. The model is based on creating two concurrent two-state Markov processes, one for the signal level above or below a certain threshold, and the other for the bit decision (correct or erroneous). A complete description of the model is given below.

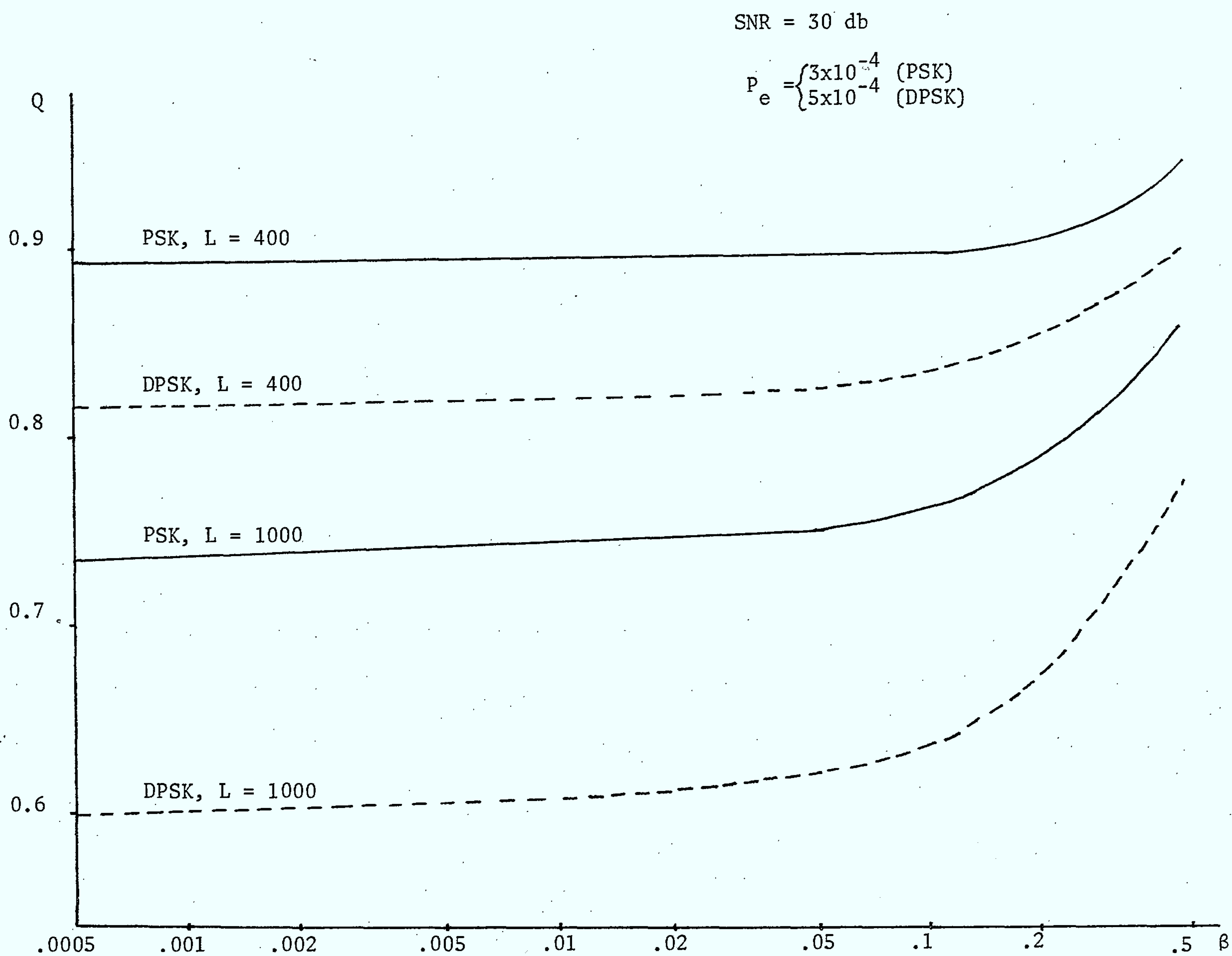


Fig. 5.1: PROBABILITY OF CORRECT PACKET VS. THE ERROR CORRELATION PARAMETER β

In a rayleigh fading channel the probability density function of the signal power (envelope squared) is the square Chi density, i.e.,

$$P(\gamma) = \frac{1}{\gamma_0} e^{-\gamma/\gamma_0} \quad ; \quad \gamma > 0 \tag{7}$$

where γ is the instantaneous power level and γ_0 is the local average power level. The fraction of time the signal stays below a certain threshold level γ_T ; (average fading interval) is:

$$T_b = \langle t_b \rangle = \frac{e^{\gamma_T/\gamma_0} - 1}{f_D \sqrt{2\pi\gamma_T/\gamma_0}} \tag{8}$$

and the average non-fade interval is

$$T_g = \langle t_g \rangle = \frac{1}{f_D \sqrt{2\pi\gamma_T/\gamma_0}} \tag{9}$$

where f_D = Doppler frequency = $\frac{f_0 \cdot \nu}{C}$ (10)

where f_0 is the carrier frequency, ν is the vehicular speed and C is the speed of light.

The Pdf of T_b and T_g are not known, but previous work has indicated that under high signal-to-noise ratio they can both be approximated as exponential densities. If this assumption is acceptable, then the Markovian model is a valid approximation. The model consists of two random processes: The first process, (Y(n), is the transition between two states denoted by G (signal above threshold) and B (signal below threshold). The second process, X(n), is the sequence of error (E) and correct (C) bit decisions (see Figure 5.2). We assume that:

$$P[X(n) = E \mid Y(n) = G] = P_1 \qquad (11)$$

$$P[X(n) = E \mid Y(n) = B] = P_2 \qquad (12)$$

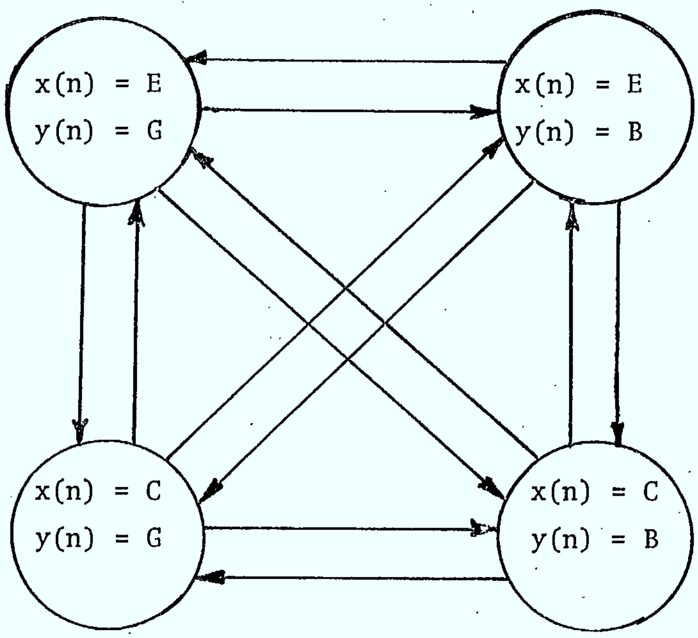


Fig. 5.2: TWO 2-STATE MARKOVIAN MODEL

where P_1 and P_2 are constants. Now, the parameter β can be described as:

$$= P[x(n) = E \mid x(n-1) = E] \qquad (13)$$

Figure 5.3 shows a tree diagram for two arbitrary time indeces (n-1) and (n) including all possible transitions of the two concurrent Markovian processes. In the diagram we define four new probabilities.

$$P_3 = P[y(n) = G \mid x(n) = E] \qquad (14)$$

$$P_4 = P[y(n) = B \mid x(n) = E] \qquad (15)$$

$$\mu = P[y(n) = B \mid y(n-1) = G] \qquad (16)$$

$$\mu = P[y(n) = G \mid y(n-1) = B] \qquad (17)$$

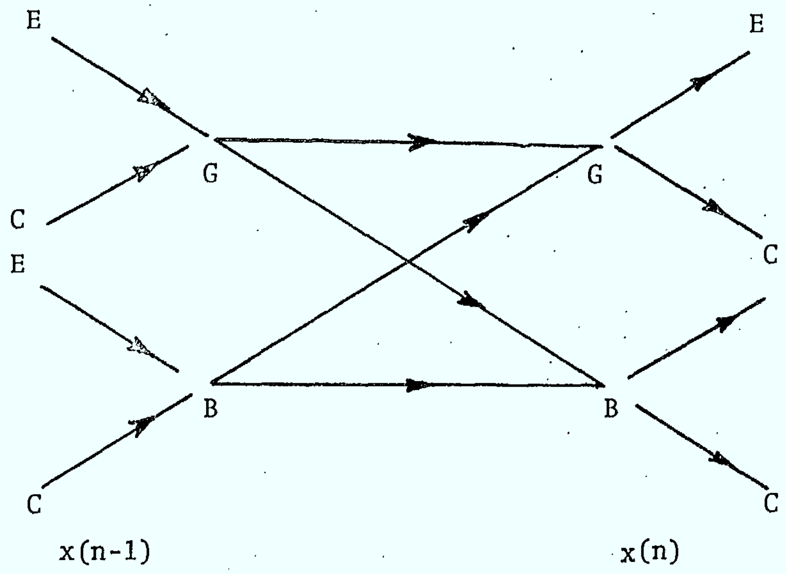


Fig. 5.3: ALL POSSIBLE TRANSITIONS BETWEEN INDEX (n-1) AND INDEX (n)

Counting the four possible ways of obtaining two consecutive errors, the parameter β can now be expressed as:

$$\beta = P_4(1-\eta)\, P_2 + P_4\eta P_1 + P_3\mu P_2 + P_3\,(1-\mu)\, P_1$$

For a given power threshold level γ_t, the six probabilities P_1, P_2, P_3, P_4, μ and η can be calculated from the following relations:

$$P_1 = \frac{P_e - P_{eb}}{1 - P_b} \tag{19}$$

$$P_2 = \frac{P_{eb}}{P_b} \tag{20}$$

$$P_3 = \frac{P_e - P_{eb}}{P_e} \tag{21}$$

$$P_4 = \frac{P_{eb}}{P_e} \tag{22}$$

$$\mu = Pbg \tag{23}$$

$$\eta = Pbg \cdot \frac{1 - P_b}{P_e} \tag{24}$$

where:

P_e = probability of bit error

$$= \int_0^{\infty} P_0(\gamma)\, f(\gamma)\, d\gamma \tag{25}$$

P_{eb} = joint probability of error and fading

$$= \int_0^{\gamma_T} P_e(\gamma)\, f(\gamma)\, d\gamma \tag{26}$$

P_b = probability of fading

$$= \frac{T_b}{T_b + T_g} \tag{27}$$

P_{bg} = transition probability from G (above thresold) to B (below threshold) during a symbol time interval T

$$= \quad 1-e^{-T/T_g} \tag{28}$$

where $P_0(\gamma)$ is the probability of error of a non-fading channel for the power level γ.

Based on the previous discussion, we conclude that the parameters involved in calculating γ are the following:

- The vehicular speed, ν
- The carrier frequency, f_0
- The threshold power level, γ_T
- An expression for the probability of error over non-fading channel for the modulation technique used, $P_0(\gamma)$
- The symbol duration T (bit rate)

and

- The probability distribution of the received signal, $f(\gamma)$

The procedure described above may be lengthy and may require some numerical integration but it is straightforward. There is only one problem with this method; it requires a careful selection of the threshold level γ_T. Notice that the threshold level divides the entire range of signal levels into two distinct ranges. An error probability is then assigned to each range, and the value of β will depend on the ratio between these two probabilities and the probability of transition from one range to the other. If the value of γ_T is too high, the transition between the two ranges becomes highly probable and the value of β will decrease. On the other hand, a very small value of γ_T will make the probability of fading too small, and since the probability of error is fixed, most of the errors will appear to occur in non-fading intervals and again β will be small. Figure 5.4 shows the dependence of β on γ_T. The calculation of β in this figure was carried out assuming a DPSK modulation which has the following simple expression of error probability.

$$P_0(\gamma)\ (\mathrm{DPSK}) = \frac{1}{2} e^{-(\mathrm{SNR})} \tag{29}$$

One way to remove the dependency of β on γ_T is to integrate over all possible values of β ; i.e.,

$$\beta = \int_0^{\infty} \beta(\gamma_T)\, f(\gamma_T)\, d\gamma_T \tag{30}$$

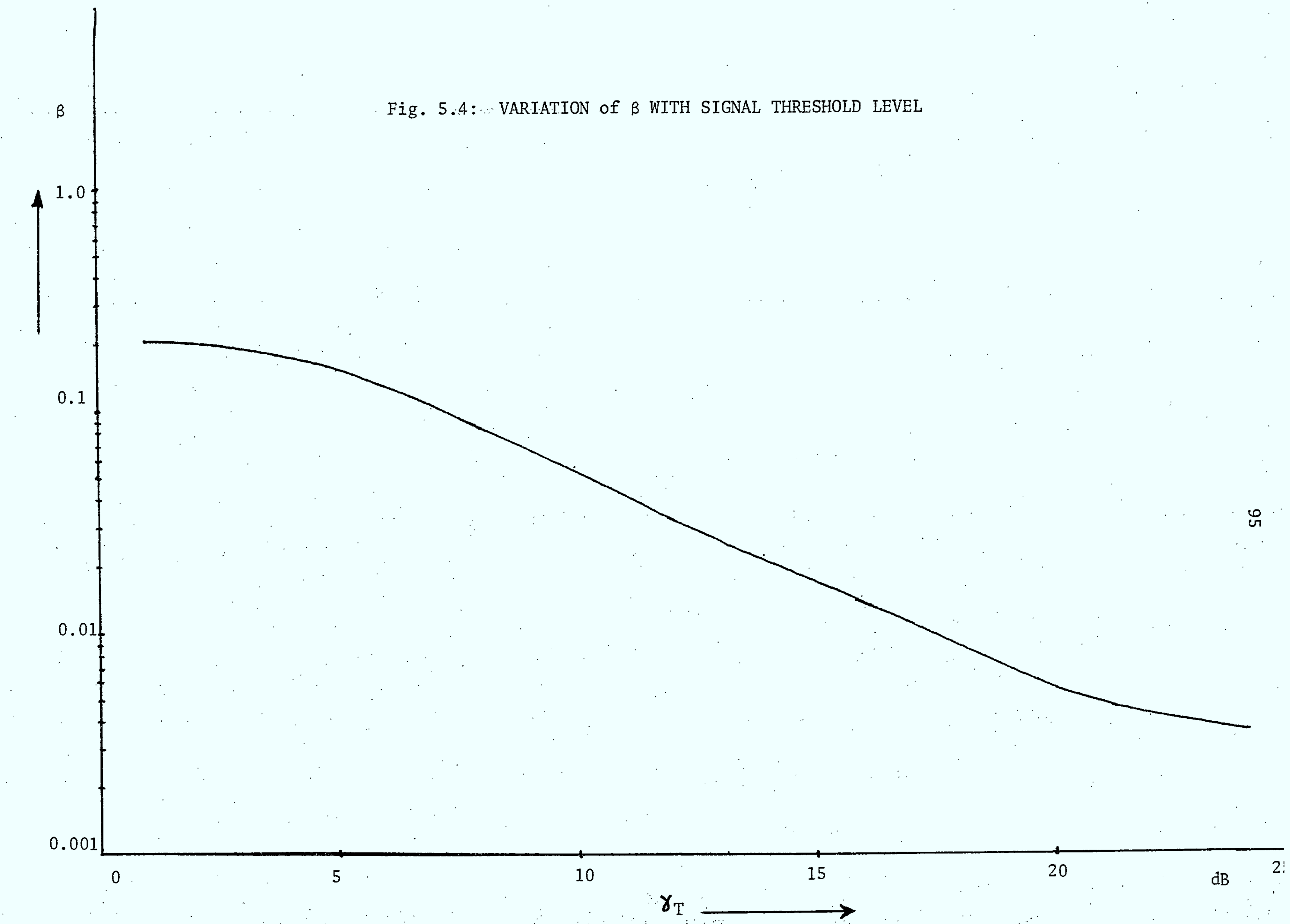


Fig. 5.4: VARIATION of β WITH SIGNAL THRESHOLD LEVEL

where $f(\gamma_T)$ is the Rayleigh distribution given in Eqn. (7). Now, the calculation of β is many order more tedious and time consuming.

A simple approximation, which can be very accurate provided that the fading envelope is much slower than the bit rate, is to assume that the signal level remains nearly constant over two consecutive bits. Then, the probability of two errors in a row is:

$$P_2 = \int_0^\infty P_0^2(\gamma)\, f(\gamma)\, d\gamma \tag{31}$$

and the parameter β can be approximated as:

$$\beta = \frac{P_2}{P_e}$$

Eqn. (31) is a valid approximation as long as the bit rate R is much greater than the Doppler frequency F_D.

For DPSK P_2 and P_e are calculated as follows:

$$P_2 = \int_0^\infty \left(\frac{1}{2}\, e^{-\gamma/N}\right)^2 \circ \frac{1}{\gamma_0}\, e^{-\gamma/\gamma_0}\, d\gamma \tag{32}$$

$$= \frac{1}{4[2\frac{\gamma_0}{N} + 1]} = \frac{1}{4[2(SNR)_{av}+1]}$$

$$\text{and} \quad P_e = \int_0^{\infty} \frac{1}{2} e^{-\gamma/N} \cdot \frac{1}{\gamma_0} e^{-\gamma/\gamma_0} \, d\gamma$$

$$= \frac{1}{2[(SNR)_{av}+1]} \tag{33}$$

and then;

$$\beta_{DPSK} \cong \frac{(SNR)_{av}+1}{4(SNR)_{av}+2} \tag{34}$$

Note that the limit of β_{DPSK} as SNR $\rightarrow \infty$ is .25 and for very small values of SNR $\beta \rightarrow .5$; i.e.,

$$.5 \leq \beta_{DPSK} \leq .25 \tag{35}$$

For other modulation methods Eqn. (31) may have to be evaluated numerically. However, since all modulation methods behave in more or less the same way under fading conditions, then β_{DPSK} can be taken as an approximation of the actual value of β.

Figure 5.5 illustrates the relation between the bit error rate and packet error rate for various values of β.

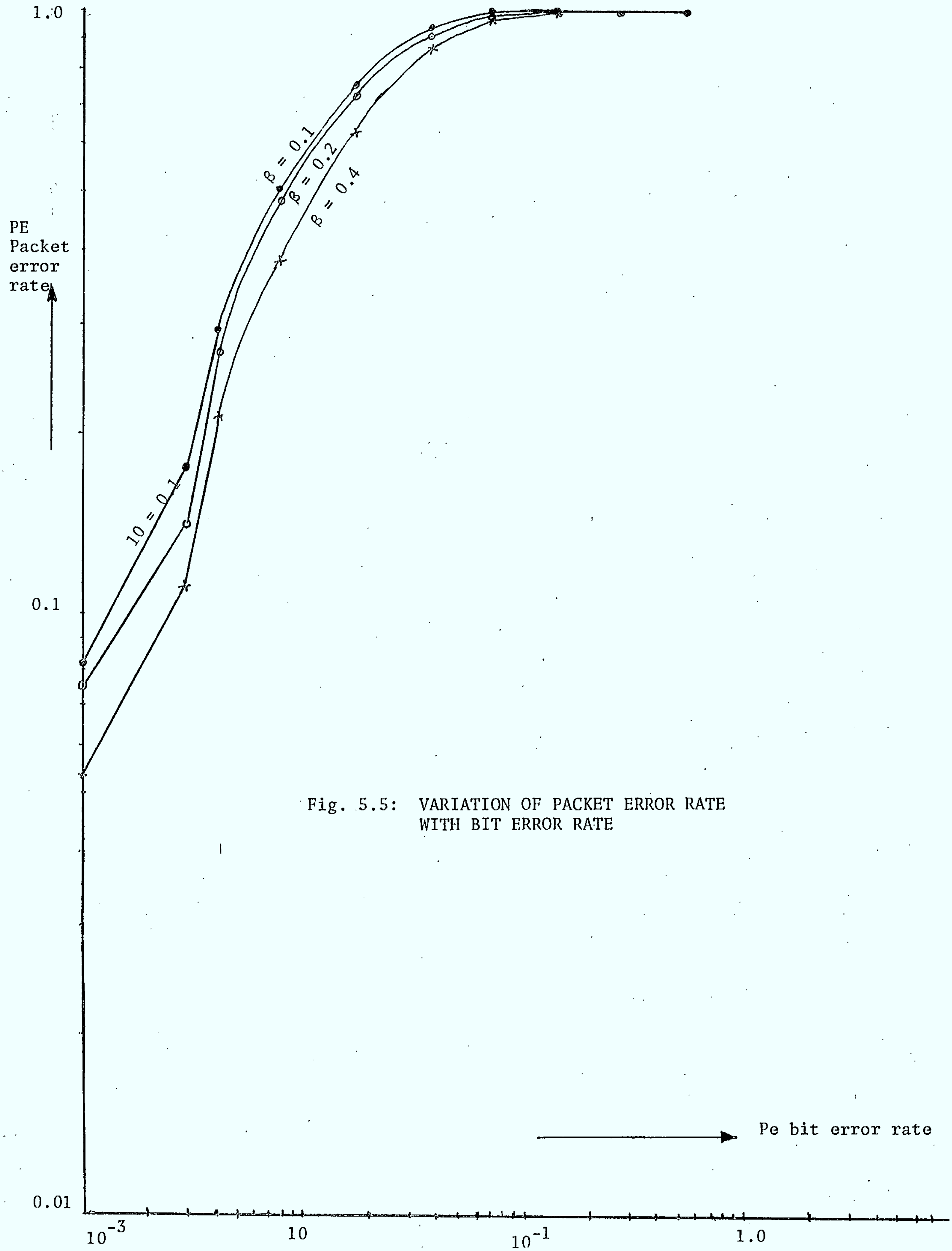


Fig. 5.5: VARIATION OF PACKET ERROR RATE WITH BIT ERROR RATE

5.3 Channel Throughput of ALOHA/ARQ Protocol

The basic structure of a packet radio network consists of a central station and a number of radio terminals scattered over a geographically limited area. The radio terminals share a single radio channel to reach the central station on a random access basis. We assume a pure ALOHA access scheme, whereby a ready terminal blindly transmits its packet and waits for an acknowledgement. If the packet is received correctly, the base station immediately sends a short acknowledgement packet. If no acknowledgement is received within A seconds (A is constant), the terminal reschedules the same packet for retransmission X seconds later, where X is an uniformly distributed random delay with a mean value of X. We assume that all terminals form (collectively) a Poisson source with an average packet generation rate of λ packets/sec. There are two sources of error in the channel: (1) Errors due to noise, and (2) Errors due to packet collision, and since these two mechanisms are independent, then the probability of packet retransmission can be written as:

$$P_r = P(C) + P(E) - P(C)P(E) \tag{35}$$

where P(C) is the probability of packet collision and P(E) is the packet error rate due to noise and fading.

Figure 5.6 illustrates the life cycle of an arbitrary packet. The total packet delay from the instant the packet arrives to the network until it is received correctly and acknowledged can be expressed as:

$$d = (1+R)(T+A + \sum_{i=1}^{R} X_i \qquad (36)$$

where:

T = Packet transmission time, sec.

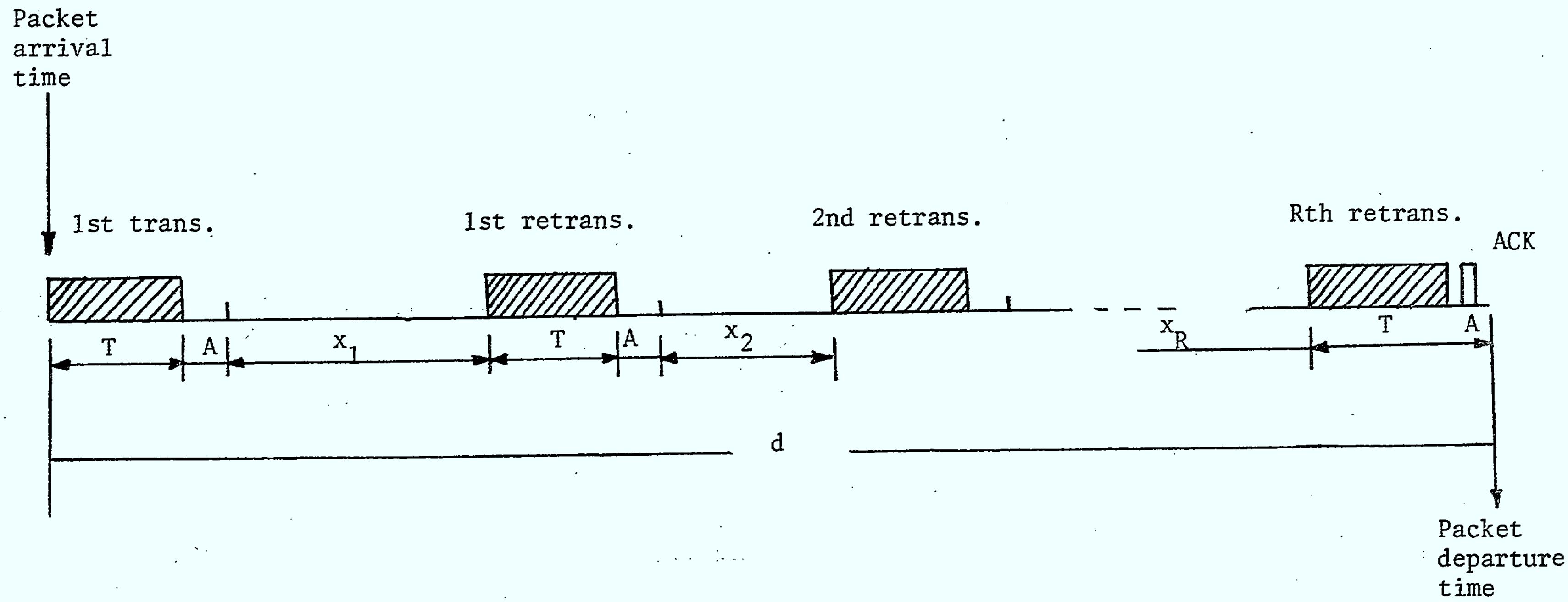
A = Acknowledgement delay, sec.

R = Number of retransmission (random)

X_i = The intentional retransmission delay associated with the i^{th} retransmission trial, sec.

The details of the packet transmission time components, T, and the acknowledgement delay components, A, are shown in Figure 5.7, where

$$T = \tau_S + \frac{L}{r_b} \qquad (37)$$



T = Packet transmission time

A = Acknowledgement delay

x_i = Retransmission delay ; $i = 1, 2, .., R$

d = Total delay

Fig. 5.6: LIFE CYCLE OF AN ARBITRARY PACKET

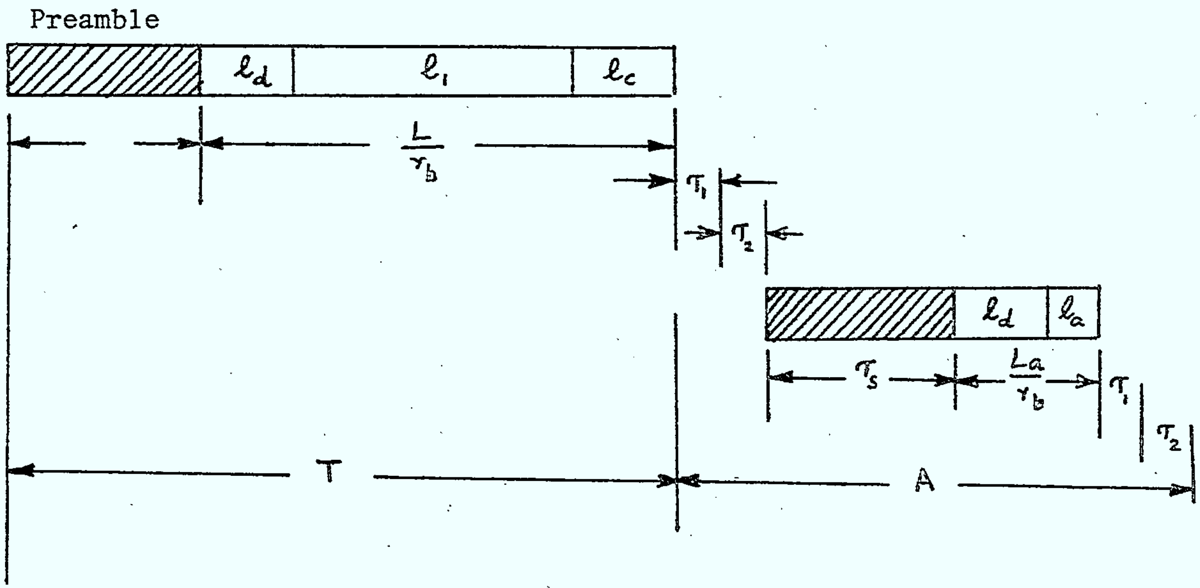


Fig. 5.7: DETAILS OF THE PACKET TRANSMISSION TIME COMPONENTS AND THE ACK. DELAY COMPONENTS

ℓ_d = address

ℓ_i = information bits

ℓ_c = check sum

ℓ_a = ack. bits

$L = \ell_d + \ell_i + \ell_c$

$L_a = \ell_d + \ell_a$

τ_1 = propagation delay

τ_2 = processing delay

τ_s = synch. delay

$$A = \tau_S + \frac{L_a}{r_D} + 2\,(\tau_1 + \tau_2) \qquad (38)$$

and $$L = \ell_d + \ell_i + \ell_c \qquad (39)$$

$$L_a = \ell_d + \ell_a \qquad (40)$$

where:

τ_s = bit synchronization time, sec.

r_b = bit rate, bits/sec.

ℓ_d = addressing bits

ℓ_c = check bits

ℓ_i = information bits

ℓ_a = acknowledgement bits

τ_1 = one way propagation delay, sec.

τ_2 = processing delay, sec.

The retransmission delay, X, is assumed to be uniformly distributed with an average of mT seconds where $m >> 1$.

For a given traffic level, the average number of packet retransmission is $\bar{R}$, which is assumed to be independent of the retransmission delays X_i's.

Then, the average packet delay is

$$\bar{d} = (1+\bar{R})(T+A) + \bar{R}\ \bar{X} \tag{41}$$

where $\bar{X}$ = mT = the mean retransmission delay. Substituting (37) and (38) in (41) gives

$$\bar{d} = k_1 + k_2\ \tau_s \tag{42}$$

$$\text{where } k_1 = (1+\bar{R})\left[\left(\frac{L+L_a}{r_b}\right) + 2(\tau_1+\tau_2)\right] + \frac{mL}{r_b}\bar{R} \tag{43}$$

$$\text{and } k_2 = 2(1+\bar{R}) + m\bar{R}. \tag{44}$$

The average number of packet retransmission R will now be calculated for fading channels.

The probability of packet collision, P(C), is a function of the channel traffic load, λ_c, which is the sum of the source traffic λ and the retransmission traffic λ_r. The exact probability distribution function of the channel traffic is difficult to calculate, since it constitutes two dependent random processes, the source packet generation which is assumed to be Poisson, and the packet retransmission process which is difficult to evaluate analytically. But, if we assume that the random delay X is relatively large and the traffic load is stable, then the distribution of the channel traffic process can be approximated by a Poisson process with an average rate of $\lambda_c = \lambda + \lambda_r$. In this case the probability of packet collision will be

$$P(C) = 1 - \exp[-2\lambda_c T] \qquad (45)$$

λ_c and $\overline{R}$ are related by

$$\lambda_c = \lambda \cdot (1+\overline{R}) \qquad (46)$$

Substituting Eqns. (45) and (46) in (35) we get:

$$P_r = 1 - Q \exp[-2\lambda\overline{T}(1+R)] \qquad (47)$$

where $Q = 1-P(E)$ (48)

Under steady state conditions, the probability of packet retransmission, P_r, is constant; hence the number of transmissions per packet (1+R) has a geometric distribution, i.e.,

$$P[(R+1) = k] = P_r^{k-1}(1-P_r) \qquad (49)$$

It follows that the average number of packet retransmissions is

$$\overline{R} = \frac{P_r}{1-P_r} \tag{50}$$

Using Eqns. (47) and (50) we get:

$$\overline{R} = \frac{1-Q\ \exp[-2\lambda T\ (1+\bar{R})]}{Q\ \exp[-2\lambda T\ (1+\bar{R})]} \tag{51}$$

A numerical solution can be easily obtained for Eqn. (51) which can then be substituted in Eqns. (42) to obtain the average packet delay for various system parameters.

Rearranging Eqn. (51) provides an expression for the channel throughput s = λT in terms of R and Q.

$$S = \lambda T = \frac{\ell n\ [Q(1+\bar{R})]}{2(1+\bar{R})} \tag{52}$$

Now, we have all the relations needed to predict the behaviour of ALOHA/ARQ mobile data communication channels. We shall summarize that part by giving a design example.

Design Example:

The ultimate performance criteria in any multiple access scheme are: The average delay and channel throughput. The final design decision usually involves a trade-off between these two parameters. Let us find the throughput delay relation for an ALOHA/ARQ mobile channel having the following parameters:

(1) Carrier frequency = f_0 = 850 MHZ.

(2) Av. vehicular speed = ν = 30 km/hr.

(3) Modulation method : DPSK

(4) Av. signal-to-noise ratio = $(SNR)_{av}$ = 20 dB

(5) Bit rate = R = $\frac{1}{T}$ = 16 kbps

(6) Bit synchronization time = τ_S = 3 msec (48 bits)

(7) Av. propagation delay (one way) = τ_1 = .03 msec

(8) Processing delay = $\tau_2 \simeq 0$

(9) Addressing bits = ℓ_d = 16 bits

(10) Check bits = ℓ_c = 16 bits

(11) Information bits = ℓ_i = 160 bits

(12) Acknowledgement bits = ℓ_a = 8 bits

(13) Normalized retransmission delay = $\frac{x}{T}$ = m = 20

For DPSK, the burstness parameter β can be calculated from Eqn. (34):

$$\beta \cong \frac{(SNR)_{av} + 1}{4(SNR)_{av} + 2} \simeq \frac{100 + 1}{400 + 2} = \underline{\underline{.2512}}$$

and the average bit error rate is (Eqn. 33)

$$P_e = \frac{1}{2(SNR)_{av} + 2} = \underline{\underline{4.95 \times 10^{-3}}}$$

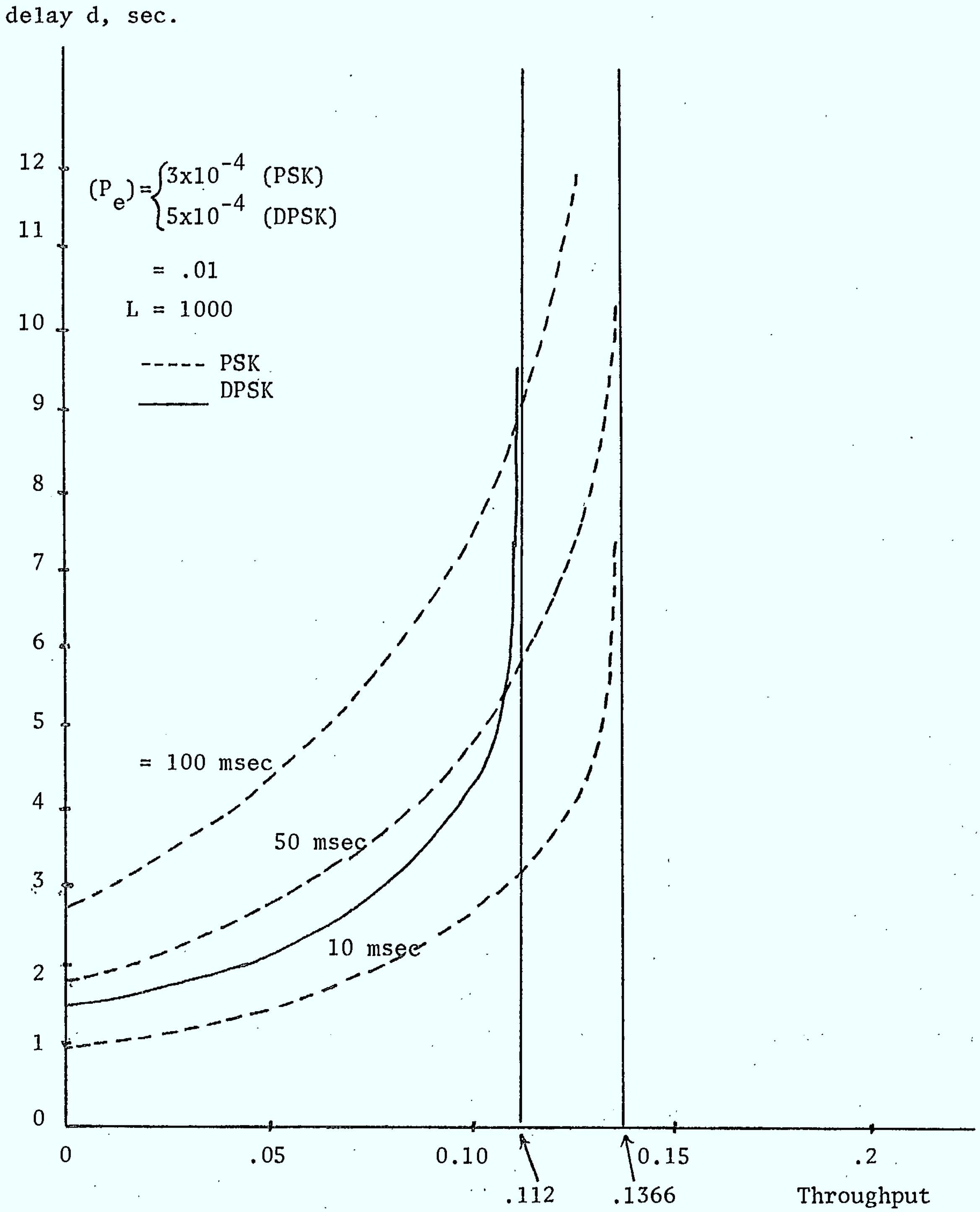


Fig. 5.8: DELAY-THROUGHPUT CHARACTERISTIC FOR A MARKOVIAN CHANNEL MODEL (= .01)

Then the packet error rate P(E) is (Eqn. 5)

$$P(E) = 1 - (1-P_e)[1-(1-\beta)\frac{P_e}{1-P_e}]^{L-1} = \underline{.512}$$

and $Q = 1-P(E) = \underline{.488}$

Using Eqns. (42-44) the average delay $\bar{d}$ can be written as a function of $\bar{R}$ as:

$$d = .01956 + .31956\,\bar{R}$$

and the channel throughput is (Eqn. 51)

$$s = \frac{\ln(.512\,(1+\bar{R}))}{2\,(1+\bar{R})}$$

The relation between s and d is shown in Fig. 5.8. The maximum throughput is S_{max} .0942 with a corresponding average delay per packet of approximately 1.36 sec. At that point the average number of retransmission per packet is nearly equal to 4.

5.4 Effect of Synchronization Time on Channel Throughput

In the upper portion of the UHF band (around 850 MHz) the number of fades per unit time is high and the duration of non-fade intervals is relatively short. Then, in order to minimize the probability of packet retransmission, the pack length must be short. Under these conditions it is important that the packet header (synch. bits and addresses) be as short as possible in order to maintain a reasonable transmission efficiency. This gives rise to one of the outstanding problems in designing packet radio systems, namely, finding the relative performance of coherent and noncoherent systems in terms of the overall packet delay and channel throughput. Coherent systems are characterized by long synchronization times and low error probability. These two characteristics affect the packet delay in opposite ways. The long synchronization time reduces the transmission efficiency and increases the delay, while the low probability of error reduces the probability of packet retrans-mimssion, hence reducing the average packet delay. On the other hand, noncoherent systems exhibit higher bit error rates but have much shorter synchronization time than that of the coherent systems. A meaningful comparison between the two systems can only be made on the basis of their relative performance in terms of the packet delay and channel throughput.

The analysis given in the previous section gives some insight into the problem. In particular, Eqn. (42) expresses the average packet delay as a linear function of the synchronization time τ_s . The coefficients k_1 and k_2 are also functions of the selection of coherent or noncoherent systems. In general, k_1 and k_2 are larger for noncoherent systems because of the larger values of R. But, since R is a function of many of the system parameters (bit rate, packet length, ... etc.), one would expect that for some sets of parameters coherent systems are superior to noncoherent systems, while for other sets the opposite is true.

To examine this point, the channel throughput was calculated for two widely used systems: DPSK which represents the noncoherent class, and PSK which represents the coherent class. A sample of the results is shown in Fig. 5.9. In this figure the synchronization time of DPSK is assumed to be zero; then the throughput vs. τ_s is a family of horizontal straight lines. These lines are not shown in the figure; only the intersection points between these lines and the corresponding PSK curves are marked and connected with the broken line. To the left of the broken line PSK provides higher throughput (for the same delay) than DPSK. To the right of the line the opposite is true. The figure also indicates that larger

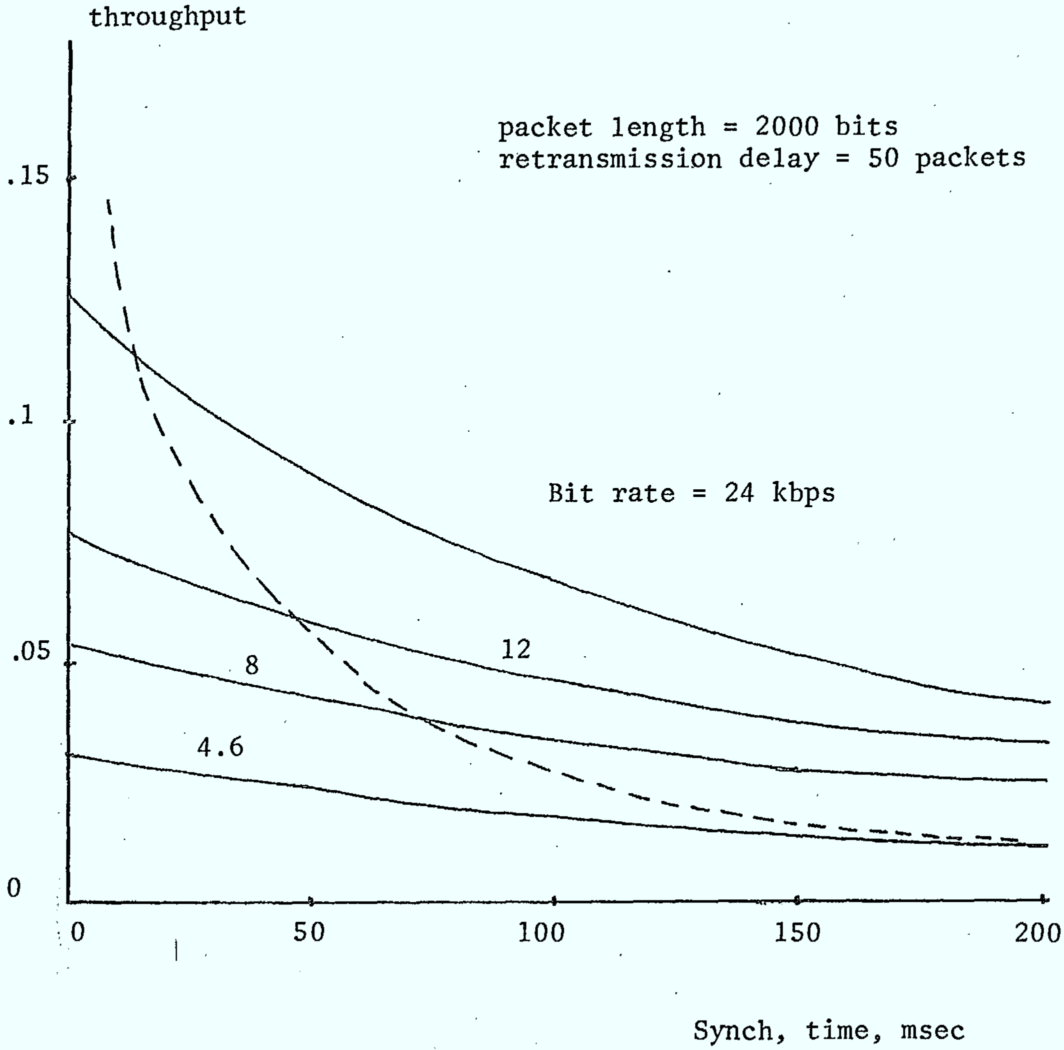


Fig. 5.9: THROUGHPUT VS. SYNCH. TIME FOR COHERENT PSK. FOR VARIOUS BIT RATES

synchronization times can be tolerated for lower bit roles. When the bit rate is high, the transmission efficiency of the coherent system drops very quickly with increased synchronization delays.

There is a basic conclusion that can be drawn from Fig. 5.9, that is for a fixed delay requirement coherent systems provide higher throughput than noncoherent systems provided that the synchronization time is kept very small. Once the synchronization time exceeds a certain critical value the noncoherent systems become more efficient. The critical synchronization time becomes smaller as the bit rate is increased.

CHAPTER VI

SUMMARY AND CONCLUSIONS

6.1 Summary

The need to increase the spectral utilization has been stressed and various methods of achieving this goal have been discussed. Some of the steps already taken by the regulatory activities resulted in proposals of novel systems such as cellular radio, a step further in this direction being time assigned sharing of voice communication channels by more than one user. These systems commonly known as TASI are well known in Switched Telephone Networks [STN]. It has been noted that adaptation of TASI to mobile network confronts complexities unknown to TASI for STN, these obstacles being channel fading and unavailability of permanent wireline connections as is the case in STN. Therefore for adaptation of packetized voice in the land mobile systems, a radio connection (channel) must be demanded before any packet is transmitted. The desirability of contiguous speech at the receiving end results in severe time constraints.

Three schemes have been proposed for packetized voice over the land mobile channel. Two of these, the ALOHA and the POLLING schemes, are modified versions of similar schemes already in use over telephone networks. A provision is obviously made to demand a channel before transmission of voice packets. Simulations of these two schemes resulted in similar results, with POLLING scheme having a slight edge over the ALOHA scheme. However, system parameters used in simulation may seem to rely on a rapid progress in technology in the near future. This could be doubtful. The conclusions are optimistic in the sense that, given that the technology delivers the desired advancements in the system design, the two schemes are feasible and a considerable increase in channel throughput could be obtained. It is though not clear whether the technological goal could be achieved. The performance of the POLLING system is dependent on the geographical layout.

The third scheme, TABLE DRIVEN system, does not depend on ambitious technological advancement, but is performance-wise inferior to the other two schemes. For instance, for traffic intensity of 0.8 Erlangs and 64 channels, the blocking probability is 0.011%, 0.015% and 0.09% for ALOHA, POLLING and TABLE DRIVEN systems respectively. For all the three systems considerable improvement in system performance is seen when the number of available channels is increased.

The nature of the mobile channel complicates the signal design. The packet error rate is dependent on the outage probability of the channel, which affects the channel throughput as well as degrades the control aspects of the channel. An expression for packet error rate is derived and this throughput of ALOHA/ARQ protocol for fading channel is determined.

For packetized voice/data, the overhead required synchronous system has shown to have considerable impact on the throughput of the channel. The decision whether to use synchronous or asynchronous systems depends critically on the synchronization time required. Again this is a technology dependent aspect of the system. If synchronization time is very small, then it is shown that synchronous systems will have higher throughput. For longer synchronization time, the asynchronous system is recommended.

6.2 Recommendation for Future Work:

The feasibility of packetized voice over the mobile channel has been established. It is recommended that further studies are required to study some variations of TABLE DRIVEN systems. These variations could be in the following aspects of the system:

i) modes to get access into the system;
ii) creation of channel groups to reduce the probability of a channel access by two different parties; and
iii) to devise some water-tight methods of frequent update of table of free channel at the receiver.

ALOHA and POLLING methods require further studies in reducing the severe time constraints on the system. It is also not known how the system will perform if realistic system parameters such as channel switching times are used.

References

[1] K. Bullington and J.M. Fraser, "Engineering Aspects of TASI," B.S.T.J., Vol. 38, pp 353-364, March 1959.

[2] S.J. Companella, "Digital Speech Interpolation," COMSAT Technical Review, Vol. 6, No. 1, pp 127-158, Spring 1976.

[3] E. Lynghounis, I. Poretti, G. Monti, "Speech Interpolation in Digital Transmission Systems," IEEE Trans. on Communications, Vol. COM-22, pp 1179-1189, Sept. 1974.

[4] D. Seitzer, H. Gerhäuser, G. Langenbucher, "A Comparative Study of High Quality Digital Speech Interpolation Systems," Proc. ICC, pp 50.2.1 - 50.2.5, June 1978.

[5] J. Soumagne, "A High Quality Digital Speech Interpolation System," Proc. ICC, pp 37.6.1 - 37.6.6, June 1979.

[6] R.L. Easton, P.T. Hutchison, R.W. Kolot, R.C. Mondello, and R.W. Muise, "TASI-E Communications System," Proc. ICC, pp 49.3.1 - 49.3.5, June 1981.

[7] G.H. Dysart, M.J. Krone, J.E. Fielding, "Integrated Voice/Data Private Network Planning."

[8] M.S. Nakhla and D.A. Black, "Analysis of a TASI System Employing Speech Storage," Proc. ICC, pp 49.2.1 - 49.2.8, June 1981.

[9] J.H. Rieser, H.G. Suyderhoud, and Y. Yatsuzuka, "Design Consideration for Digital Speech Interpolation," Proc. ICC, pp 49.4.1 - 49.4.7, June 1981.

5.5 Tamed Frequency Modulation

The Tamed Frequency Modulation (TFM) is a member of a class of modulation called "Spectrally Efficient Constant Envelope Modulation." This class of modulation is ideal for digital radio applications since the spectra of the modulated signals are very compact around the center frequency with a very little out-of-band radiation. The TFM in particular combines this important advantage with the relative ease of demodulation and good error performance.

A complete TFM modem was designed and built at Carleton over the last two years. The parameters of that modem were selected so that it can be used to transmit high speed packet data or digital voice. It is designed to operate at 16 kbps over 30 khz RF channel spacing with an out-of-band radiation level that satisfy the DOC radio regulation.

Within the context of this work the TFM can be looked at as a part of the transmission facilities that facilitate the development of voice/data integrated radio network.

In this section we shall first describe the modem, and then turn our attention to the question of synchronization time both for the carrier frequency and phase and for the timing signal, Results of some laboratory measurements are included at the end of the section.

5.5.1 A Description of Carleton TFM Modem

A block diagram of the modultion is shown in Fig. 5.10. This is what is known as the quadrature (or I-Q) implementation strategy. In the figure the vector $\underline{x}$ represents a sequence of random binary data, which is fed simultaneously into two quadrature baseband waveform generators. These generators are basically ROMs and logic circuits. Several waveforms are stored in the ROMs, and, depending on the pattern of the input data and the current state of the modulated signal, one of these waveforms will be selected by the logic circuit and read out to the next stage of the modulator. A stable carrier (set at 455 khz) is used in a quadrature form to generate the I and Q parts of the modulated signals which are then assumed to produce the final signal $s(t, \underline{\alpha})$. A more detailed figure of the waveform generators is shown in Fig. 5.11, and the measured frequency spectrum of the I-channel baseband signal (which is similar to the RF spectrum) is shown in Fig. 5.12.

A block diagram of the TFM receiver is shown in Fig. 5.13. Again, this is what is known as a quadrature receiver which consists of two quadrature branches. The received signal is fed simultaneously to these two branches as well as the carrier and clock (timing) recovery circuits. The carrier recovery circuit regenerates coherent quadrature carriers cos $(w_0 t)$

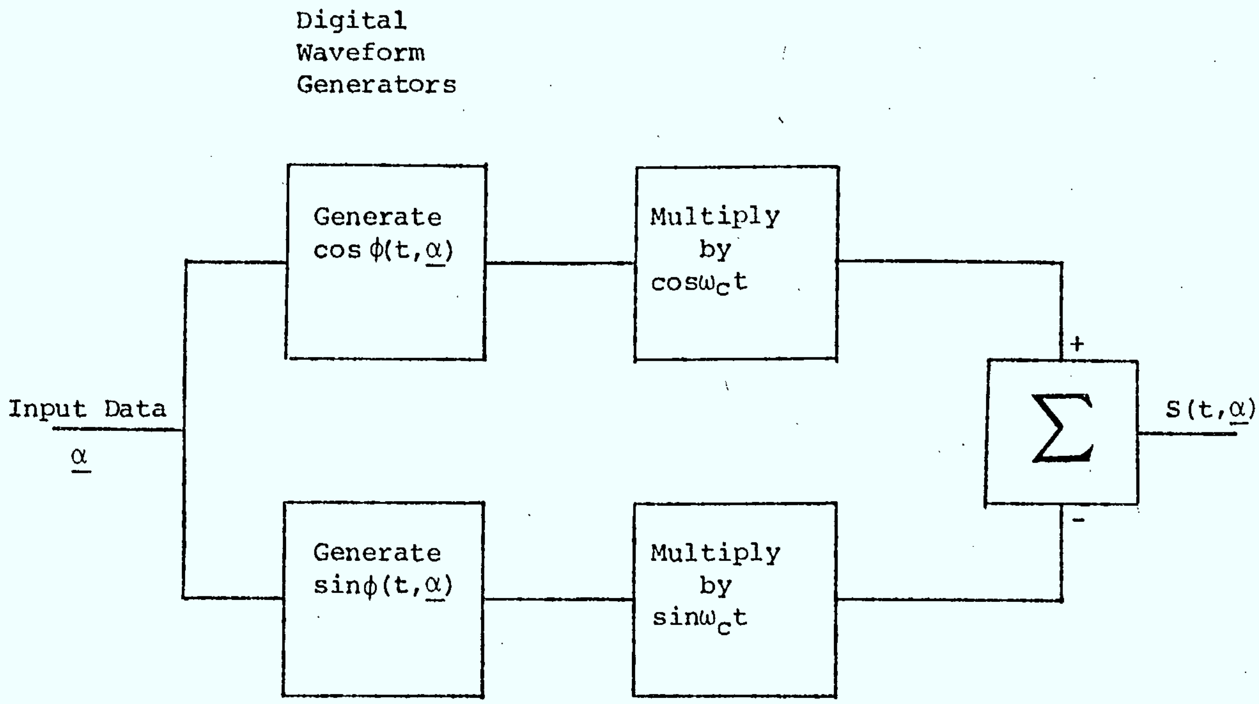


Fig 5.10 A Functional Diagram of the I-Q Implementation for the TFM Modulator

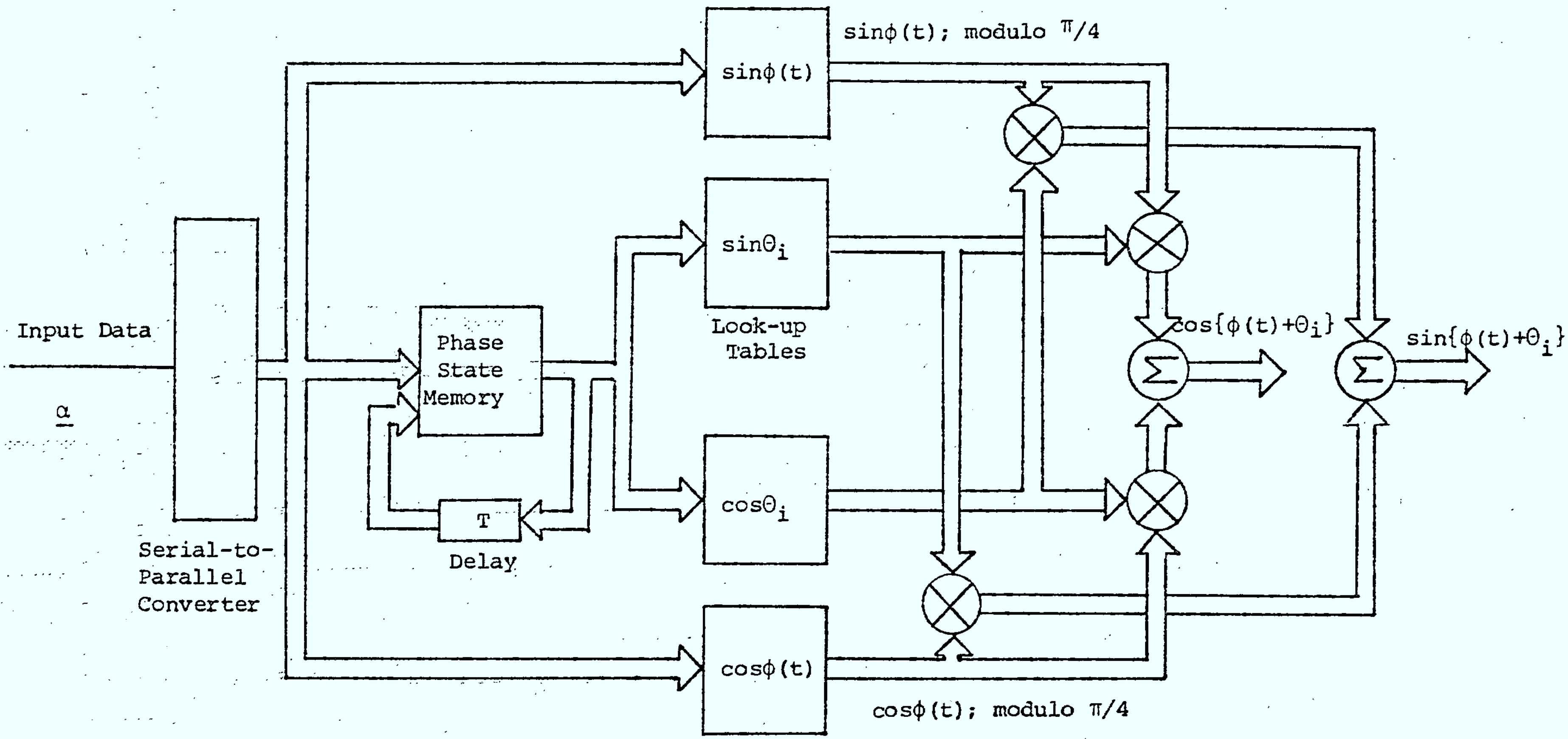


Fig 5.11 A Realization for the SIN and COS Waveform Generators.

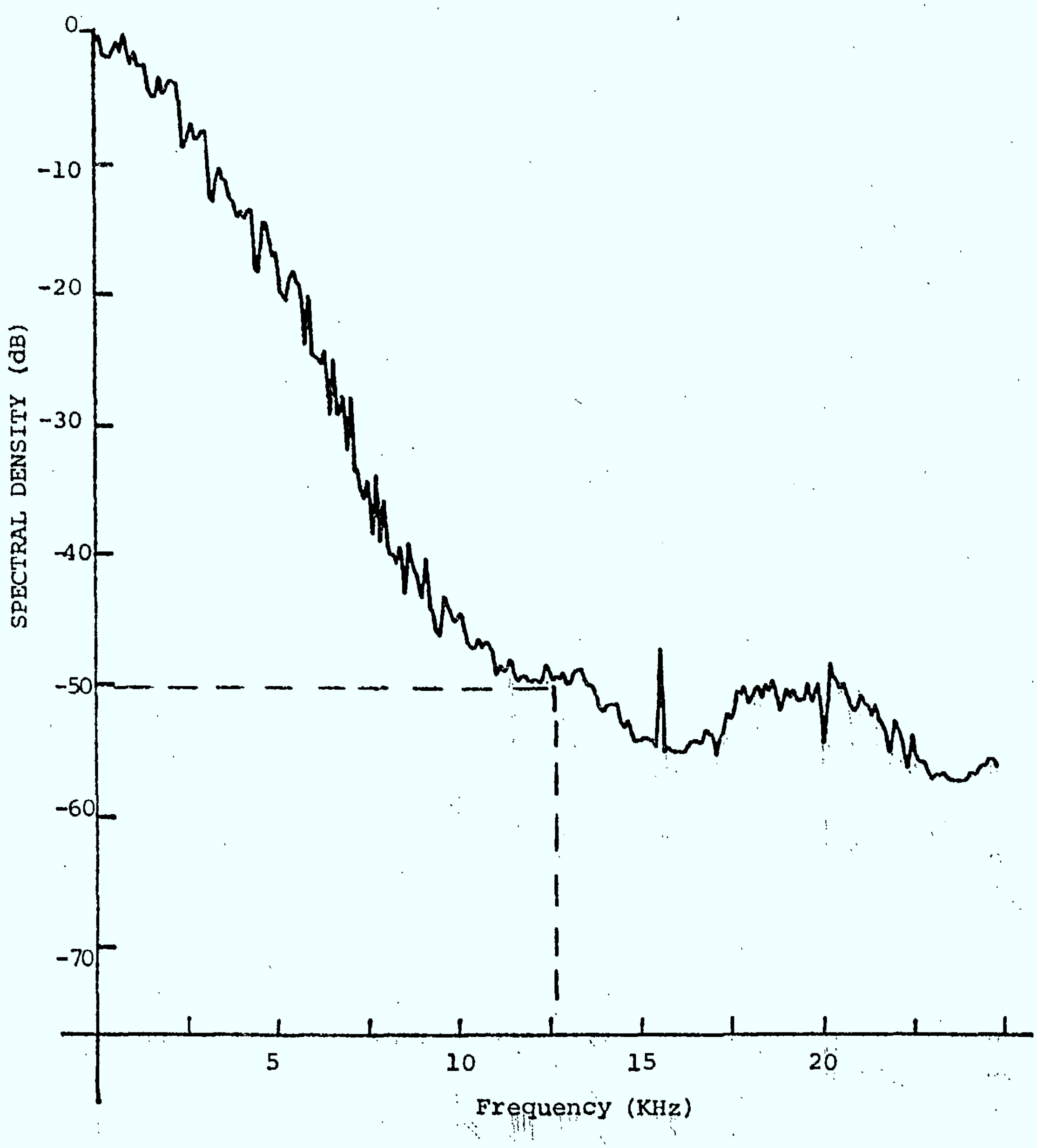


Fig 5.12 Measured Spectral Density of the I Channel Baseband Signal

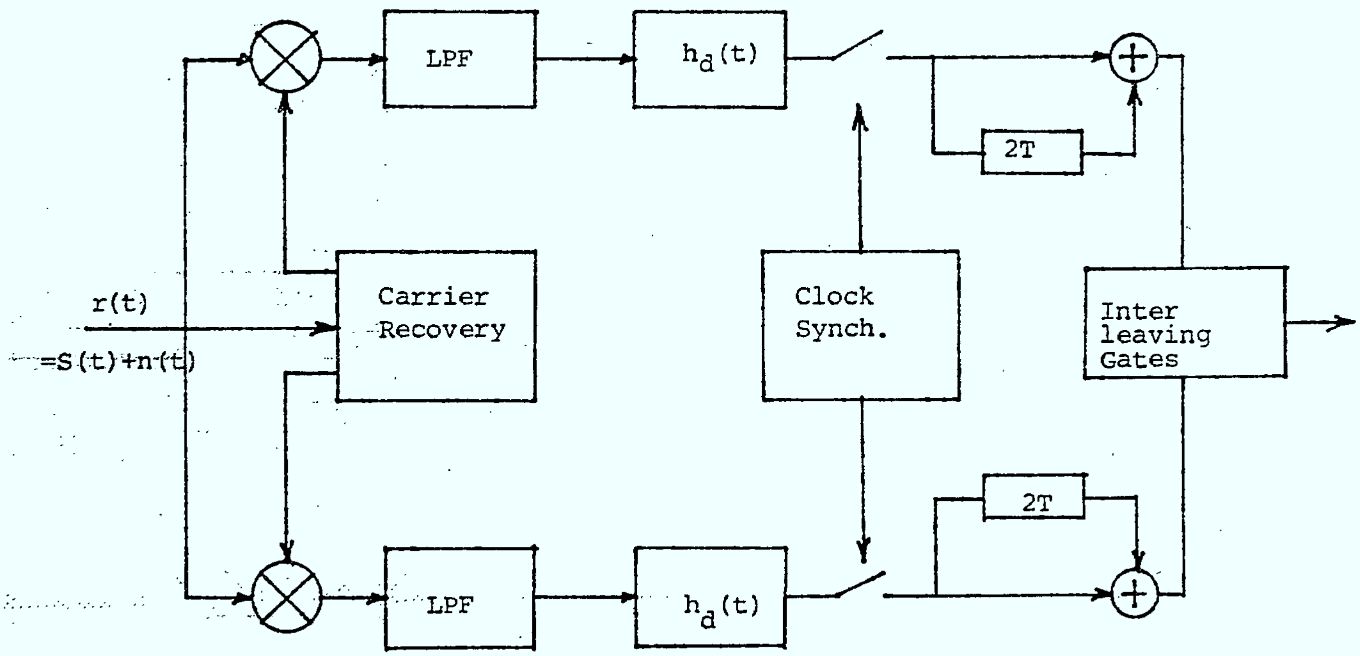


Fig 5.13 A Quadrature Coherent Receiver Structure for CPM with modulation Index ½. (The System must satisfy equ 6.11)

and sin $(w_o t)$. These carriers are multiplied by the incoming signal and low pass filtered to produce the two quadrature baseband signals. The filter $h_d(r)$ has an impulse response which is optimized to provide a near optimal shaping of the baseband spectrum such that the error performance of the system remain near optimal. Sampling and decoding processes follow the baseband spectral shaping. Again, the timing signal used in sampling the baseband signals is coherently recovered from the received signal.

The performance of the system depends critically on the accuracy and speed of the carrier and clock recovery circuits. Small deviation from perfect coherency in both circuits results in rapid degradation of the system error rate performance. The measured and theoretical error performance of the modem is shown in Fig. 5.14.

5.5.2 Carrier Acquisition Time

Successful transmission of information through a phase coherent communications system requires, by definition, a receiver capable of determining the phase and frequency of the received signal with as little error as possible. The acquisition of the phase and frequency of the received carrier takes time. During that time the communication link is unreliable, so that time is a necessary overhead on the system.

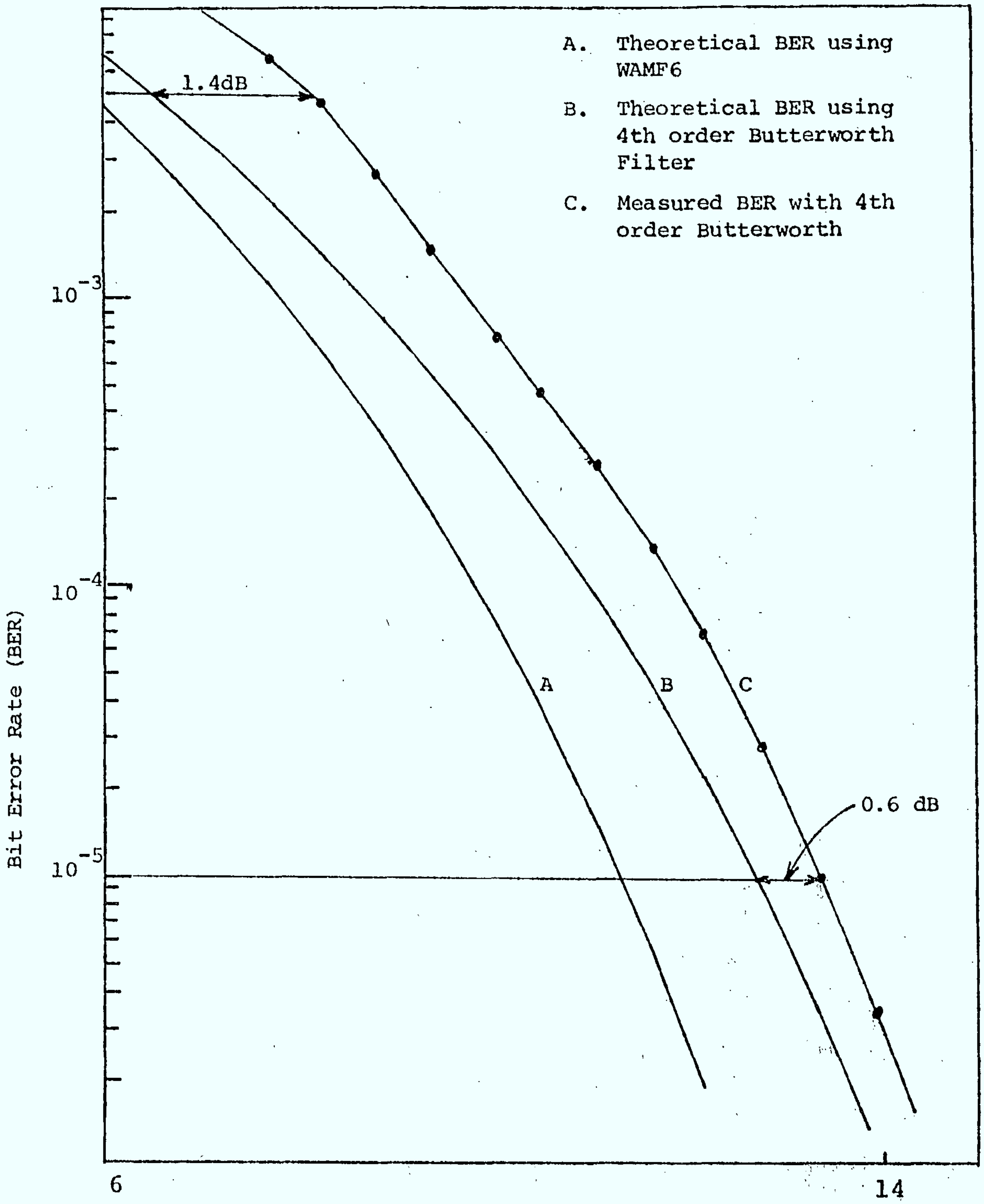


Fig (6.14): Measured System Bit Error Rate as a Function of E_B/No

For packet transmission mode, as it was discussed in previous sections, this time should be minimized if the channel throughput is to be kept high.

The carrier synchronization mechanism employed in Carleton TFM modem consists of a Costas loop enhanced with an Automatic Frequency Control (AFC) circuit as shown in Fig. 5.15. The acquisition time of this loop depends mainly on the loop filter and the AFC filter, and the theoretical determination of its achial value is difficult since it involves solving a complex differential equation. Instead of the theoretical solution, the acquisition time was measured in the laboratory by applying a frequency step at the demodulator input and monitoring the VCO control signal. The loop time is shown in Fig. 5.16. When a modulated signal is applied to the receiver, the measurement becomes harder due to the continuous fluctuation of the VCO control voltage. However, one may estimate that the setting time in that case may increase from 6 msec to probably (10-15) msec, which represents the time required to acquire the carrier phase on a packet-to-packet basis. We should note here that the acquisition time is a function of the signal-to-noise ratio as well as the initial frequency error. In general the acquisition performance improves at higher signal-to-noise ratio and smaller frequency deviation from the nominal carrier frequency.

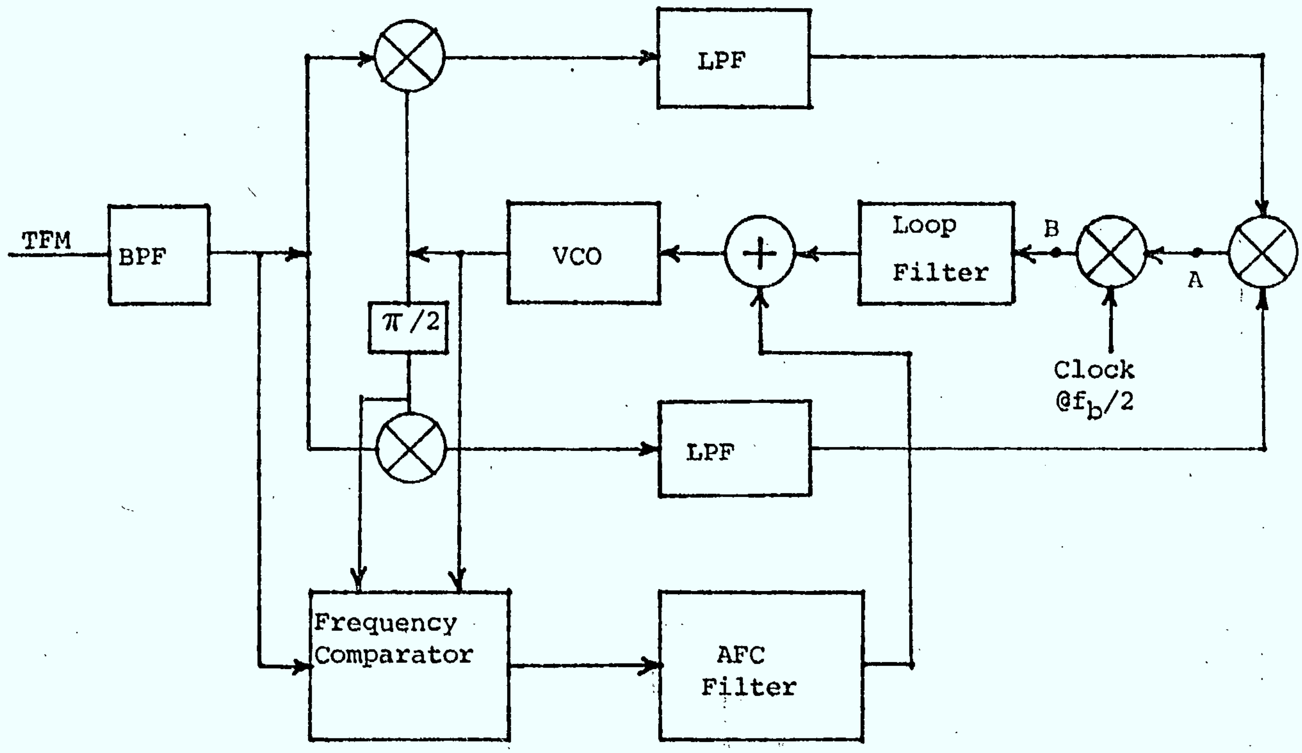


Fig 5.15 Carrier Recovery Loop, Combined with an AFC Path to Aid Acquisition

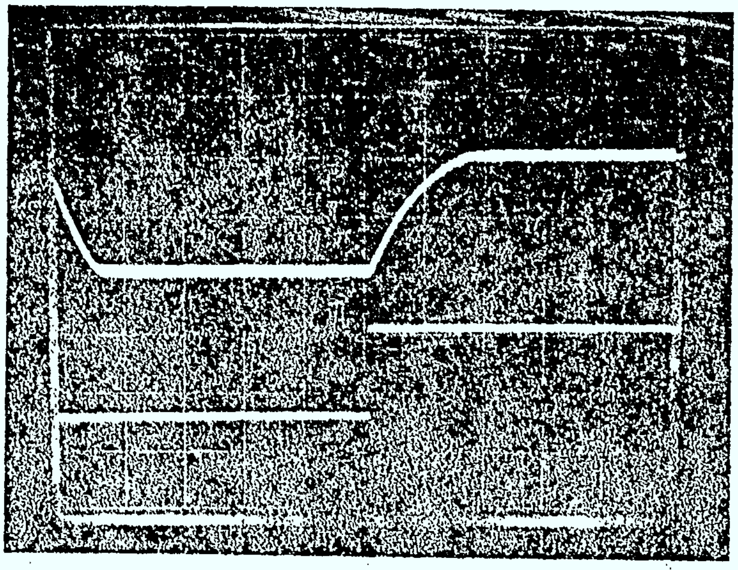


Fig. 5.16 Loop response to a 20KHz frequency step.
The IF signal is a single frequency. The bottom trace: the Signal generator tuning voltage.
The top trace: the VCO Control Voltage.
Horizontal scale: 4m.s/div.

5.5.3 Clock Recovery Circuit

The circuit used to recover the timing information in Carleton TFM modem is shown in a block form in Fig. 5.17. The loop performance was measured using the setup shown in Fig. 5.18. The setup tests the response of the clock recovery circuit to a step in the phase and frequency. The results of the test are shown in Figs. 5.19 - 5.22.

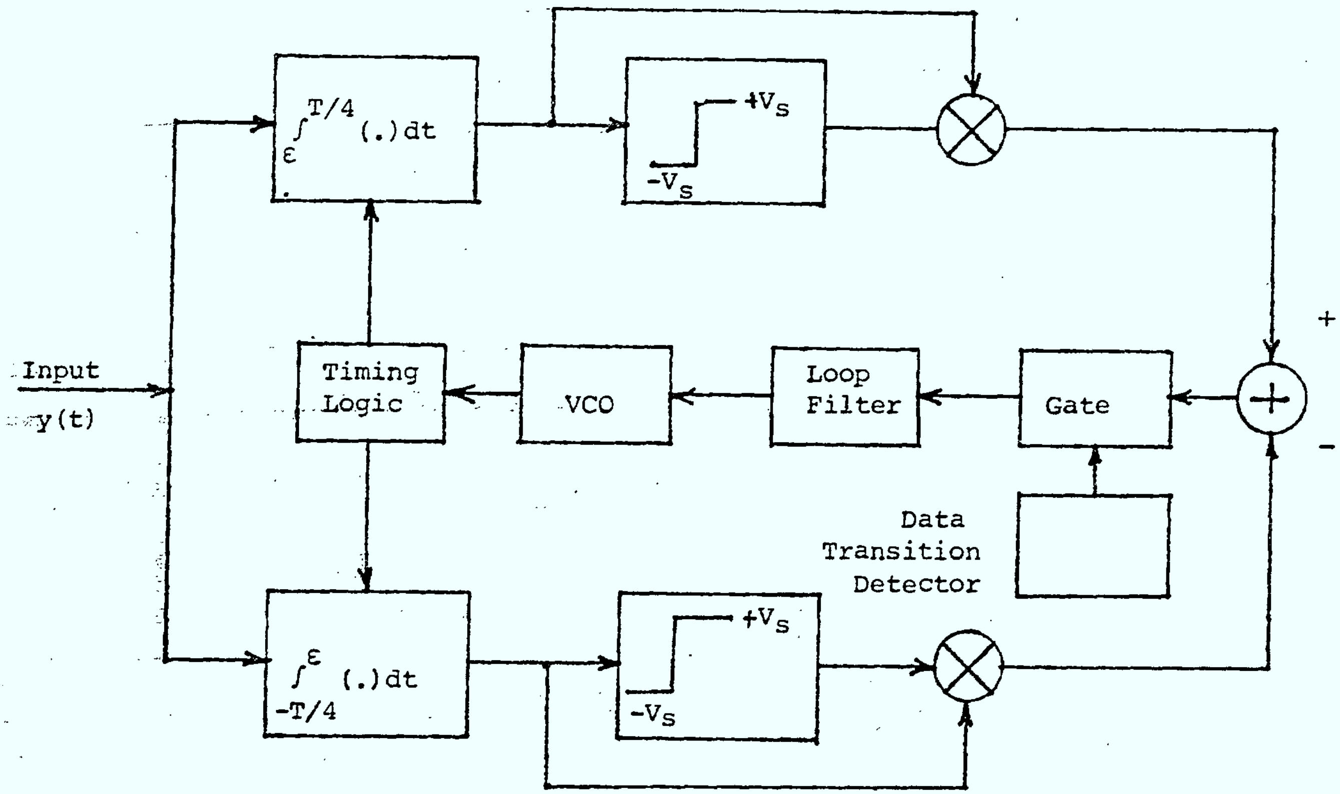


Fig. 5.17 Block Diagram of the Proposed Clock Synchronizer

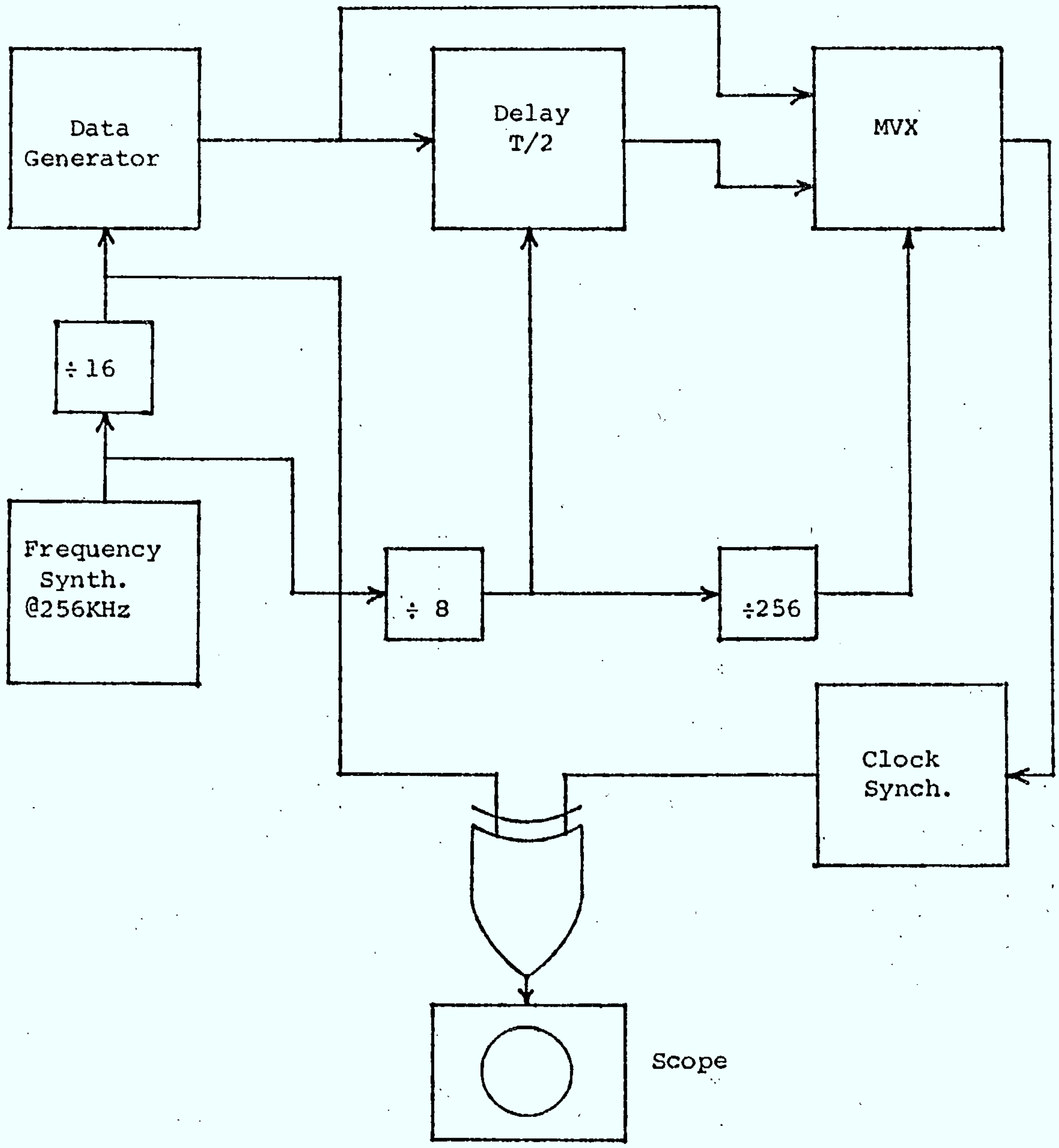


Fig 5.18 Experimental Set-up for the Measurement of the Loop Response to a step in Phase.

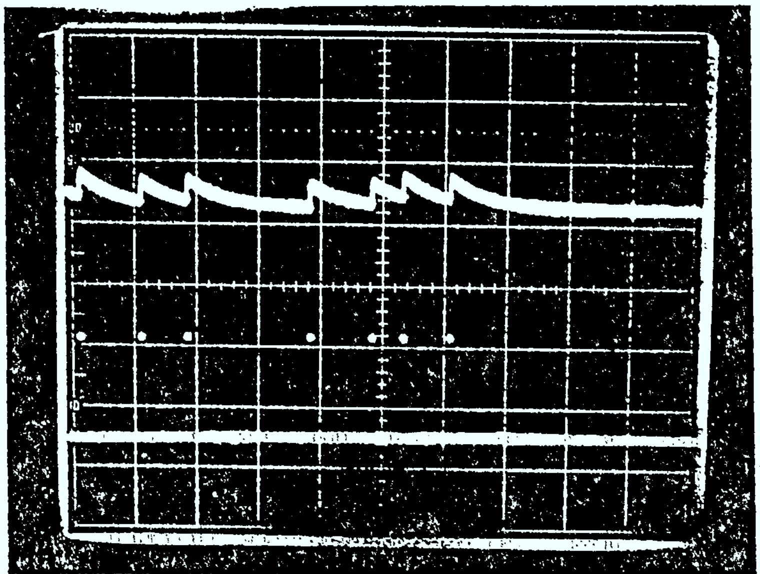


Fig. 5.19 Steady State Loop Phase Error Due to a Frequency Step Equal to 0.195% of the Free Running Frequency.

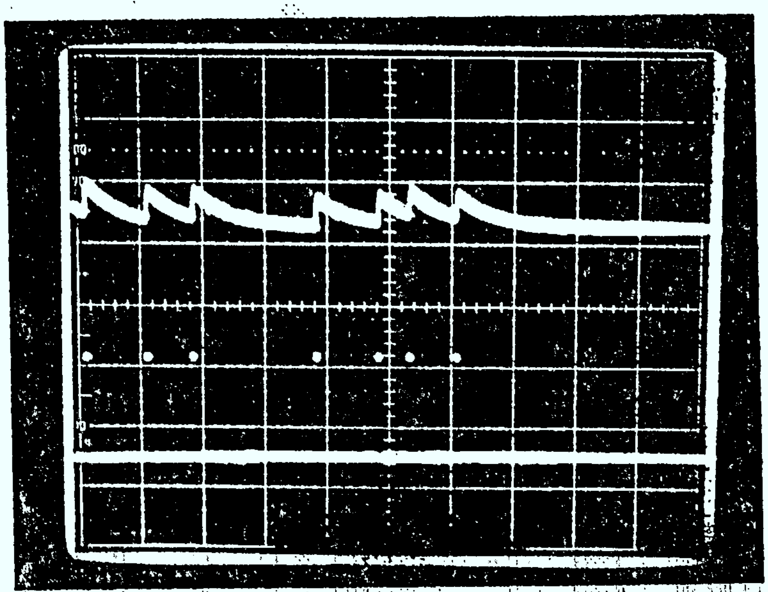


Fig. 5.20 Steady State Phase Offset Due to a Frequency Step Equal to 0.27% of the Free Running Frequency.

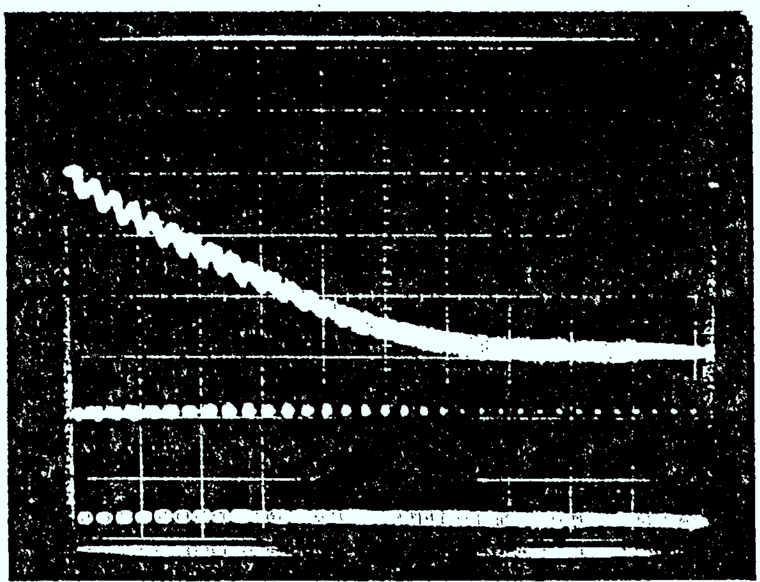


Fig. 5.22. Loop Response to a 180° Phase Step.
Note that the loop requires 20 data transitions to achieve lock.