Analysis and Design of a Data Network for

Distance Education for the State of Virginia

by

Shikhar Kishore Srivastava

Thesis submitted to the Faculty of the

Virginia Polytechnic Institute and State University

in partial fulfillment of the requirements for the degree of

Master of Science

in

Electrical Engineering

APPROVED:

T. Pratt

Timothy Pratt, Chairman

Wayne A. Scales

Wayne A. Scales

Brian D. Woerner

Brian D. Woerner

June, 1994

Blacksburg, Virginia

Analysis and Design of a Data Network for

Distance Education for the State of Virginia

by

Shikhar Kishore Srivastava

Timothy Pratt, Chairman

The Bradley Department of Electrical Engineering

(ABSTRACT)

A need exists for the State of Virginia to have a data network for its televised distance education program. A combination of a terrestrial and a satellite data network can be utilized for the purpose. The network is analyzed and its strengths and weaknesses are presented. A data protocol has been written to control such a network. Delays and throughput of the network have been calculated. The leased telephone line network can be utilized for transferring data from distant class sites to Blacksburg. Six pages of text for 75 off-campus students can be transferred from distant class sites to Blacksburg using this network, in one hour. When the terrestrial network is used for voice and data communication at the same time, a delay of approximately 30 seconds is introduced between two voice connections. This delay is too high for a distance education network. The satellite data network should be utilized for transferring data from Blacksburg to all distant class sites. A very good 19.2 kbps carrier is available with bit error rate (BER) of 1E-6 or less. A very small aperture terminal (VSAT) network has also been proposed for the purpose.

Acknowledgments

I would like to express my deep appreciation to my advisor, Dr. T. Pratt, for all his encouragement, patience and guidance. I would like to thank Dr. W. Scales and Dr. B. Woerner for their helpful suggestions and valuable time. I would like to acknowledge the students and staff of Satellite Communications Group for their friendship and companionship throughout my stay at Virginia Tech. Without the encouragement and inspiration from my parents I would have never made it this far in my academic career.

TABLE OF CONTENTS

LIST OF FIGURES

List of Tables

CHAPTER 1

Introduction

Virginia Tech has operated a distance learning network since 1983 using a satellite network for transmitting Television (TV) signals from a TV studio in Blacksburg and receiving the TV signal at all the distant class sites. A terrestrial telephone leased line network is used for voice connection between distant class sites and Blacksburg. This thesis examines the network and suggests methods to improve the efficiency and capability of the network.

One point that needs to be discussed before we attempt to improve and analyze the Virginia Tech's distance education network is why distance education is required at all? One of the most important reasons for the use of distance education technologies is economic. Hewlett-Packard Distance Learning System trains the employees at one half the cost of traditional classes [1]. IBM reports [2] that its satellite education network has accounted for a cost avoidance in travel and living expenses in excess of $15 million per year, which represents a return of over 30%.

Distance learning results in economic gains because it eliminates travel and living expenses and at the same time reduces students' time away from their jobs. Fewer instructors are required to reach a much larger student body. Development and delivery of courses is centralized, thereby decreasing duplication of efforts. The above reasons particular affect the corporate world.

On a more social level distance education helps solve problems that arise from geographic isolation. Young people [3] living in rural areas and towns away from big cities and education centers can benefit from this technology. Distance education can provide more choice of courses and instructors compared with a traditional education.

People with severe disabilities will also find this technology useful. They can get instruction in their own home without the need to go to a school or university on a daily basis. Temporarily ill people can also benefit from the service.

Traditional schools and universities may become too expensive for the states to provide. Developing countries may have to spend a major part of their annual budget on education. One cost-cutting alternative which could be employed by governments could be the use of a combination of radio, print and short face-to-face courses. In the U.S.A. it can also serve as an integral part of private education at lower cost which has been seen as an alternative to public education.

A very important advantage of such a technology is for people who work during the day and take classes during evenings. They need to travel less to attend class and they also have the freedom to tape a class lecture and replay it later at their own convenience.

The most important advantage of distance learning technology seems to be that it will bring experts in direct contact with students throughout the world. It will enable the author of a book to talk about it interactively with students throughout the world, or a research scientist to take part in an audio/visual conference with students thousands of miles away across the seas.

This thesis examines the strengths and weaknesses of the present network operated by Virginia Tech. The growing demands on the network and availability of advanced technology suggests that improvements be made in the network. The thesis suggests and analyses improvements for the distance learning network. Its contributions are:

- analysis of the present network for distance learning
- a proposal for improvement in the present network
- a network control protocol for the control of such a network

- analysis of the protocol
- delay and throughput calculations for the network
- proposal of a VSAT network for distance learning.

Chapter 2 covers basic concepts of data communication such as modulation, error correction and detection, retransmission strategies, the ISO reference model and multiaccess communication. Chapter 3 outlines basic concepts and results for delays in data networks, which are always a major factor in the design of a data network. Chapter 4 covers in some detail a multi-access scheme used in the proposed data network in this project: polling. Chapter 5 outlines basic concepts in link design for digital satellite communication. In Chapter 6 the existing network is outlined and strengths and weaknesses of the network analyzed. A protocol has been proposed here that improves the efficiency and capability of the network. Chapter 7 analyses the protocol and calculates the delays and throughput for the proposed data protocol. A VSAT network has been proposed here for data and video transfer between Blacksburg and all distant class sites.

CHAPTER 2

Basic Concepts of Data Communications

This chapter covers principles and theory of communications as applicable to data communications through telephone lines. Data communication concepts have been developed here that will be utilized in the analysis and design of the terrestrial and satellite data network for a distance education network. With even the most efficient techniques of communications it is found that the data are received with some errors at the receiver. The latter part of the chapter discusses how the errors can be detected and corrected at the receiver.

2.1 The Physical Layer [4]

The physical layer has the function of transporting bits or characters from one end of the channel to the other end. This section focuses on channels that connect just two nodes. Channels connecting more than two nodes are discussed in Section 2.7.

A point to point channel can be broadly divided into one of two classes: Digital channel or an Analog channel. A digital channel has discrete symbols as its input and output whereas an analog channel accepts a waveform, which is continuous in both time and amplitude, as input and outputs a continuous waveform. Any physical channel is by nature analog. We may however, choose to use an abstract digital channel model which includes the modulator, physical channel and demodulator. This section discusses analog channels and shows how digital channels can be developed from analog channels.

2.1.1 Linear Time Invariant Filtering [4]

Let $s(t)$ be an analog channel input, a time varying current or voltage waveform, and let $r(t)$ represent the output of the channel. It is found that $r(t)$ is a distorted, delayed and attenuated version of $s(t)$. Distortion occurs because of the effect of filters and other components in the communication systems and the inherent characteristics of the propagation medium. A linear time invariant filter has the following properties [4] :

1. If input $s(t)$ yields output $r(t)$, then for any τ input $s(t-\tau)$ yields output $r(t-\tau)$.
2. If $s(t)$ yields $r(t)$ then for any real number α, $\alpha s(t)$ yields $\alpha r(t)$.
3. If $s_1(t)$ yields $r_1(t)$ and $s_2(t)$ yields $r_2(t)$ then $s_1(t)+s_2(t)$ yields $r_1(t)+r_2(t)$.

Figure 2.1 shows an example of the relation between input and output waveforms for a communication channel with filtering. The figure also shows how to map incoming bits 110100 into an analog waveform for transmission. The mapping scheme is referred to as nonreturn to zero code (NRZ).

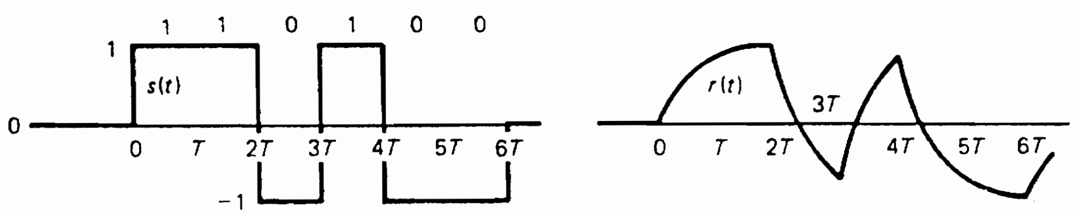
If the data rate is increased, it is seen that the received waveform becomes more distorted. The response of a single pulse lasts longer than one pulse period. The output at a given instant of time t depends on the polarity of several input pulses before time t. This phenomenon is called intersymbol interference (ISI).

2.1.2 Impulse Response Of A Channel [4]

If $h(t)$ is the channel response to an infinitesimally narrow pulse of unit area at time 0 then

(a)

(a) Isolated pulse



(b) NRZ pulse train

Input waveform output waveform

Figure 2.1 Examples of output waveforms and Input waveforms for a baseband communication channel with filtering. Part (a) shows response of one isolated pulse $s(t)$ and part (b) shows response to a sequence of pulses. The input $s(t)$ in part (b) is known as the NRZ code where a stream of bits is mapped into rectangular pulses. (From Bertsekas D. and Gallagher R., "*Data Networks*", Prentice Hall Inc., Engelwood Cliffs, New Jersey, 1990).

$h(t)$ is called the impulse response of the channel. It is also known [4] that

$$r(t) = \int_{-\infty}^{+\infty} s(\tau)h(t-\tau)d\tau \quad . \qquad 2.1$$

The Equation (2.1) indicates that filtering characteristics of the channel can be completely described by the impulse response $h(t)$. If $h(t)$ is known, the output $r(t)$, which is the response to the waveform $s(t)$, can be completely determined.

2.1.3 Frequency Response Of The Channel [4]

Let $s(t)$ be a complex sinusoid $[e^{2\pi ft} = \cos(2\pi ft) + j\sin(2\pi ft)]$. It should be noted here that the physical input waveform is always real but it is convenient analytically to allow $s(t)$ to be a complex function. An input $s(t)$ will yield a response $r(t)$. The physical input waveform would be Re $[s(t)]$ and output waveform would be Re$[r(t)]$. A good analysis of communication systems is given in [5, 6, 7]. Now integrating Equation (2.1) with $s(t)$ as defined above yields

$$r(t) = H(f)e^{2\pi ft} \qquad 2.2$$

where $H(f)$ is the fourier transform of $h(t)$, given by

$$H(f) = \int_{-\infty}^{\infty} h(t)e^{-j2\pi ft}dt \ . \qquad 2.3$$

$H(f)$ is called the frequency response of the channel. It is seen that when $s(t)$ is a complex sinusoid with frequency f, the response $r(t)$ is also a complex sinusoid multiplied by a factor $H(f)$, the frequency response of the channel. It can also be shown [4] that the impulse response $h(t)$ is related to $H(f)$ by the inverse fourier transformation

$$h(t) = \int_{-\infty}^{\infty} H(f) e^{j2\pi ft} dt, \qquad 2.4$$

so that $H(f)$ is the Fourier transform of $h(t)$ and $h(t)$ is the Fourier transform of $H(f)$.

The channel response $r(t)$ to a general function $s(t)$ and frequency function $R(f)$ of $r(t)$ are also related in the same way that $h(t)$ and $H(f)$ are related. For a linear channel it can also be shown that [4]

$$r(t) = \int_{-\infty}^{\infty} H(f) S(f) e^{j2\pi ft} df \qquad 2.5$$

and

$$R(f) = H(f) \times S(f) \qquad 2.6$$

where

$$S(f) = \int_{-\infty}^{\infty} s(t) e^{-j2\pi ft} dt \qquad 2.7$$

$$R(f) = \int_{-\infty}^{\infty} r(t) e^{j2\pi ft} dt \qquad 2.8$$

The frequency functions of $r(t)$ and $s(t)$ are simply related by Equation (2.6).

If $H(f)$=1 over the range of frequencies where $S(f)$ is nonzero and if $S(f)$ is zero everywhere else then $R(f)$ will be equal to $S(f)$, from which we can conclude $r(t)=s(t)$. This indicates that $r(t)$ can be exactly the same as $s(t)$ for the above conditions. Unfortunately this case can never be achieved for the following reasons:

1. It is very rare to have $H(f)=1$ over a desired frequency range
2. Additive noise is present in the channel. The noise gets added to the signal at various points along the propagation path.

2.1.4 The Sampling Theorem [8, 4]

If a waveform $s(t)$ is limited to frequencies at most W (i.e. $S(f)$=0, for $|f|>W$) then assuming that $S(f)$ does not contain an impulse at $f=W$, $s(t)$ can be completely determined by its values each $1/(2W)$ seconds [4]. Mathematically,

$$s(t)=\sum_{i=-\infty}^{\infty} s\left(\frac{i}{2W}\right)\frac{\sin[2\pi W(t-\frac{i}{2W})]}{2\pi W[t-\frac{i}{2W}]} \qquad 2.9$$

In Equation (2.9) i varies from $-\infty$ to ∞ to give sample values of $s(t)$ at every $1/(2W)$ seconds. This result shows that digital data can be mapped into sample values at a spacing of $1/(2W)$ and used to create a waveform of the given sample values that is limited to $|f|\leq W$. If this waveform is then passed through an ideal low pass filter with $H(f)$=1 for $|f|\leq W$ and $H(f)$=0 elsewhere, the received waveform will be identical to the transmitted waveform (in the absence of noise). Thus its samples can be used to recreate the original digital data [9].

Another important result from Nyquist [10] is the technique to avoid intersymbol interference (ISI). He showed that ISI can be avoided if the effective filter $H'(f)$ has an odd symmetry at the band edge (a combination of filter at the transmitter, the channel filter and the filter at the receiver, effectively it becomes the transfer function of a network or system). Mathematically if

$$H'(f+W) = 1 - H'(f-W) for |f| < W$$

$$H'(f) = 0 for |f| > 2W \qquad 2.10$$

the channel response has zero ISI. The sampling theorem is depicted in Figure 2.2.

2.1.5 Bandpass Channels [4]

For most physical channels, $H(f)$ is a nonzero quantity in some frequency band $f_1 \leq |f| \leq f_2$, where $f_1 > 0$. These channels are called bandpass channels. Many of them also have $H(0)$=0. The impulse response of these channels fluctuates around zero as shown in Figure 2.3. The use of an NRZ waveform for these channels does not produce very desirable results. Most modems for bandpass channels either directly encode digital information into signals with no dc component or else use modulation techniques. Modulation techniques for data communication are discussed briefly in the next section. We introduce here Manchester coding (also known as biphase coding) which directly encodes digital data into signals for transmission. As shown in Fig 2.4, Manchester coding has no dc component but has a transition in the middle of each signaling interval. Timing recovery at the receiver is easier when Manchester encoding is used because there is a transition in every interval. Manchester coding requires larger bandwidth than NRZ coding. Manchester coding is used in the Ethernet system and the corresponding IEEE 802.3 standard.

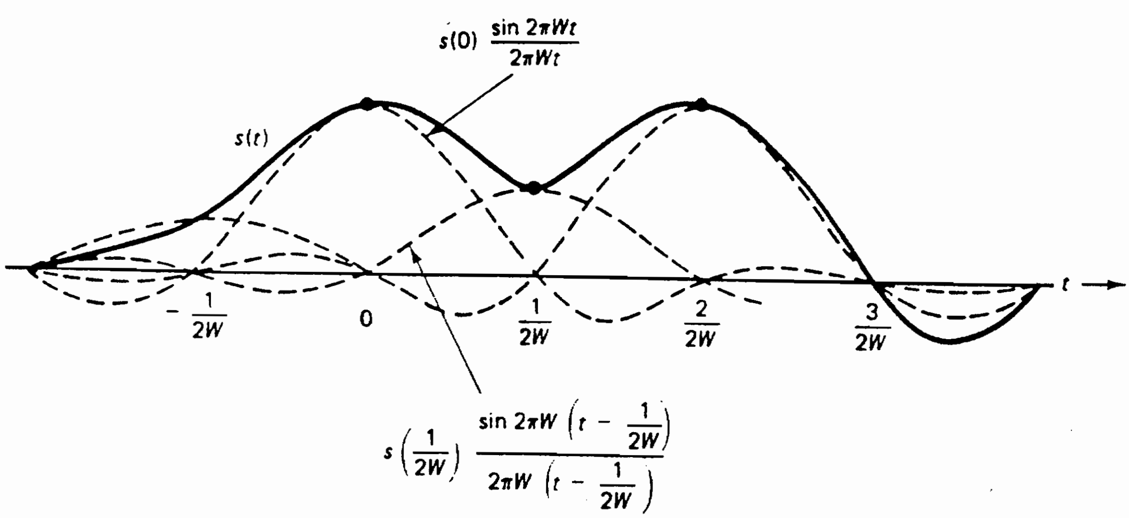


Figure 2.2 Illustration of Sampling Theorem. It depicts a function $s(t)$ that is low pass limited to frequency W. The function is a superposition of sin(x)/(x) functions. One sin(x)/(x) function at each sample instant. The value of sin(x)/(x) function is scaled to the value of $s(t)$ at the sample instant. (From Bertsekas D. and Gallagher R., "*Data Networks*", Prentice Hall Inc., Engelwood Cliffs, New Jersey, 1990).

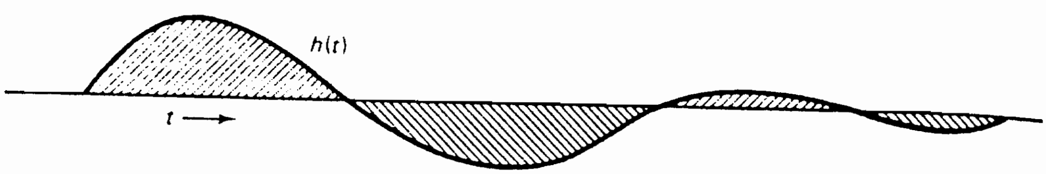


Figure 2.3 Example of an Impulse response $h(t)$ for which $H(f)$ = 0 at f = 0. (From Bertsekas D. and Gallagher R., "*Data Networks*", Prentice Hall Inc., Engelwood Cliffs, New Jersey, 1990).

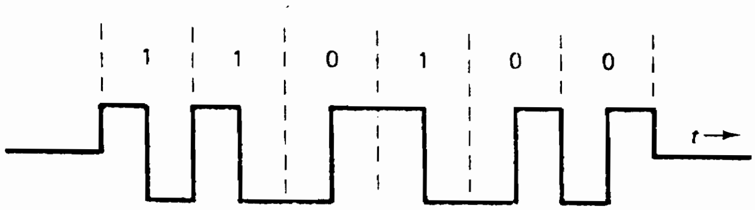


Figure 2.4 Manchester Coding. A binary 1 maps into a positive pulse followed by a negative pulse and a binary 0 maps into a negative pulse followed by a positive pulse. (From Bertsekas D. and Gallagher R., "*Data Networks*" Prentice Hall Inc., Engelwood Cliffs, New Jersey, 1990).

2.1.6 Modulation [4, 5]

Modulation is a process of encoding the source information onto a bandpass signal with a carrier frequency f_c. This bandpass signal is called the modulated signal $s(t)$, and the baseband source signal is called the modulating signal $m(t)$ [5].

2.1.7 Amplitude Modulation [4, 5]

In this section Amplitude Modulation (AM) is introduced and related to Quadrature Amplitude Modulation (QAM) and Phase Shift Keying (PSK).

In amplitude modulation a baseband signal $s(t)$ is generated from the digital data which is then multiplied by a sinusoidal carrier, say $\cos(2\pi f_0 t)$ resulting in a modulated signal $s(t)\cos(2\pi f_0 t)$ known as a Double Sideband Suppressed Carrier (DSB-SC) signal. Frequency representation of this signal is given by [5]

$$\frac{[S(f-f_0)+S(f+f_0)]}{2} \qquad 2.11$$

this is graphically shown in Figure 2.5.

It should be noted here that $S(f)$ is the Fourier transform of $s(t)$. At the receiver, the modulated signal is again multiplied by $\cos(2\pi f_0 t)$ resulting in the received signal $r(t)$ given by

$$r(t) = s(t)\cos^2(2\pi f_0 t)$$

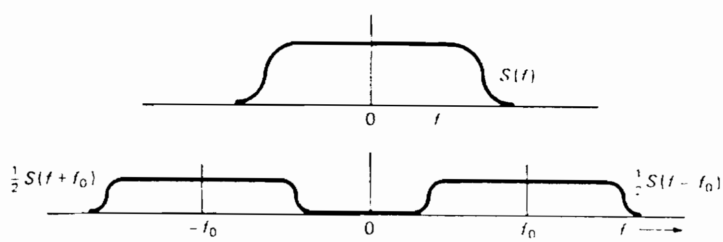


Figure 2.5 DSBSC Amplitude Modulation. Frequency spectrum in A.M. is shown for $s(t)$. (From Bertsekas D. and Gallagher R., "*Data Networks*", Prentice Hall Inc., Engelwood Cliffs, New Jersey, 1990).

$$= \frac{s(t)}{2} + \frac{s(t)\cos(4\pi f_0 t)}{2}. \qquad 2.12$$

This signal is then passed through a low pass filter to filter out the high frequency contents from the signal. The output of the filter is $s(t)/2$, which is then converted back to the digital data. One weakness of DSB-SC AM is the need for a coherent local oscillator at the receiver to demodulate the DSB-SC signal.

An improvement over DSB-SC AM is the technique called Quadrature Amplitude Modulation (QAM). QAM can be used to send twice as many bits as can be sent by the Amplitude Modulation technique. In QAM, the bits from the data source are mapped into two baseband signals $s_1(t)$ and $s_2(t)$, then $s_1(t)$ is multiplied by cos $(2\pi f_0 t)$ and $s_2(t)$ is multiplied by sin $(2\pi f_0 t)$. The sum of these products forms the transmitted QAM signal.

QAM is used widely in high data rate modems for voice grade telephone lines. The channels have a useful bandwidth from about 500 to 2900 Hz and a carrier frequency of about 1700 Hz. The $s_1(t)$ and $s_2(t)$ are waveforms limited to 1200 Hz each.

Another important consideration in communication systems is how to map digital information into waveforms for modulation. The basic idea is to map one bit into a sample of $s_1(t)$, mapping 1 into +1 and 0 into -1, and similarly map a second bit into a sample of $s_2(t)$. This is shown in Figure 2.6. Similarly for any given integer k, one can map k bits into two amplitudes. Each of the 2^k combinations of k bits maps into a different amplitude pair. This set of amplitude pairs in a mapping is called a signal constellation. The first two constellations in Figure 2.6 are also called Phase Shift Keying (PSK). As only the phase of the carrier is changed, amplitude remains constant during the process of modulation. Voice grade modems can sustain 4800 bps (with k=2 and bandwidth W=2400 Hz) and 9600 bps (with k=4 and W=2400 Hz). Superior modulation techniques have made much higher data

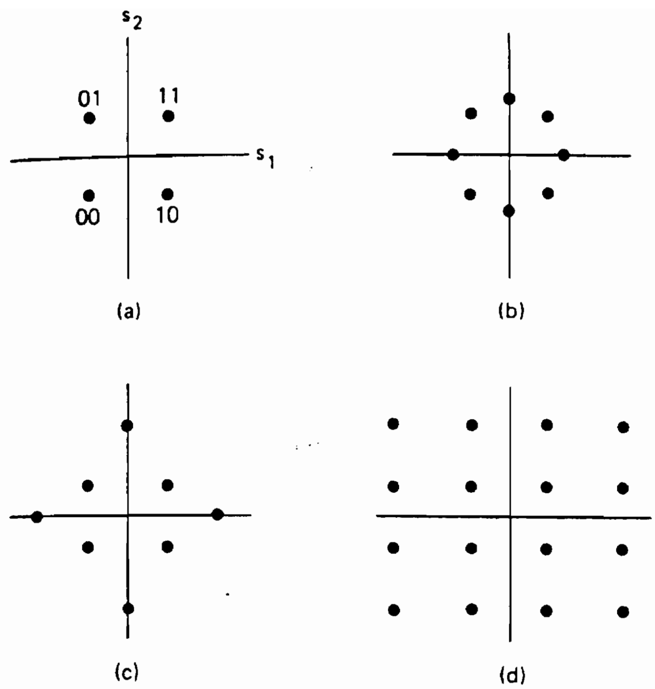


Figure 2.6 Signal Constellation for QAM. Part (a) maps two binary digits into one sample (quadrature phase shift keying). Part (b) (8-PSK) and (c) (8 QAM) each three binary digits, and part (d) (16 QAM) maps four binary digits into one sample. (From Bertsekas D. and Gallagher R., "*Data Networks*", Prentice Hall Inc., Engelwood Cliffs, New Jersey, 1990).

rates possible. Table 2.1 lists modem characteristics.

Simply increasing the value of k does not increase the channel capacity indefinitely. The work of C. E. Shannon published in 1948 [12] showed that the maximum reliable data rate sustained by a channel subject to bandwidth constraint and noise is given by

$$C = W\log\left(1+\frac{S}{N_0W}\right) \qquad 2.13$$

where the symbols have following definition and units.

Symbol	Definition	Units
C	Capacity	bits/sec
W	Bandwidth	Hz.
S	Allowable signal power	watt
N_0W	Noise power in W	watt

Shannon did not show how such data rates could be achieved. Coded modulation methods (Ungerboeck codes) are beginning to approach this limit.

Table 2.1 [11]

Modem Characteristics

Date	Rec.	Rate(b/s)	R(b/s/Hz)	W(Hz)	Modulation
1968	V.26	2400	2	1200	4-phase
1972	V.27	4800	3	1600	8-phase
1976	V.29	9600	4	2400	16-QAM
1986	V.32	9600	4	2400	2D TCM
1986	V.33	14400	6	2400	2D TCM

2.2 Errors [4, 10, 13]

This section briefly discusses why errors occur during transmission and what can be done to detect or correct errors at the receiver.

One reason for errors on telephone lines or on satellite paths is the presence of thermal noise. The electrons in the telephone lines' copper wire or in a radio receiver move randomly at high speed because of their thermal energy, producing a broad spectrum of noise.

More errors occur during data transmission on telephone lines because of impulse noise [13]. The noise pulses or spikes on the line typically last for about 10 msec. On a 9600 bps line that results in errors in 96 transmitted bits.

Another major source of errors is the fact that any transmission channel is basically a non-linear time variant channel. This also means that amplitude, propagation speed and phase of signals are all dependent on frequency. Cross-talk between two telephone lines can result in errors. Microwave links are subject to fading, migrating birds, and similar phenomena. For some modulation type like AM and PSK if the receiver loses synchronization, errors occur in bursts. All the above phenomena introduce errors in bursts. Burst errors are much more common than single errors on telephone lines.

2.2.1 Error Correcting And Detecting Codes

Two basic strategies have been developed to overcome the problem of errors occurring in data communication systems.

The first strategy is the use of Error-Correcting codes which enable the receiver to correct any errors that occur in the data during communication. The second strategy is the use of error detecting codes in conjunction with retransmission. These codes only enable the receiver to detect any errors that occur in the data during communication. After an error has been detected the receiver can ask the transmitter for retransmission of the data. The error correcting and detecting codes are discussed in [4, 13, 14].

2.2.2 Single Parity Checks

A single parity check is an error detecting code where a single bit, called a parity check bit is added to a string of data bits. This parity check bit has the value 1 if the number of 1's in the bit string is odd (or even) and has the value 0 if the number of 1's in the bit string is even (or odd). In the ASCII character code, characters are mapped into strings of seven bits and then a parity check is appended as an eighth bit. In general it can be easily deduced that any odd number of errors are detected and any even number of errors are undetected.

This scheme is not very reliable for two reasons [13]. The first is that many modems map several bits into a single sample of the physical channel input, and an error in the reception of one such sample usually results in many subsequent bits in error. The second reason is the frequent occurrence of burst errors, the reasons for which have already been outlined in this section. An example of the scheme is shown in Figure 2.7. Single parity checks detect odd number of errors and correct zero error.

2.2.3 Horizontal And Vertical Parity Checks

This is another approach for error detection. The idea is to arrange data in an array and provide a parity bit for each column and each row. Figure 2.8 gives an example of this scheme. The parity check in the lower right corner can be viewed as a parity check on the row parity checks, on the column parity checks, or on the data array. This scheme can detect an even number of errors that are confined to a

s_1	s_2	s_3	s_4	s_5	s_6	s_7	c
1	0	1	1	0	0	0	1

Figure 2.7 Single Parity Check. Bit denoted c is modulo sum of s_1 to s_k. Number of bits in code word, k, is shown to be 8 here. (From Bertsekas D. and Gallagher R., "*Data Networks*", Prentice Hall Inc., Engelwood Cliffs, New Jersey, 1990).

1	0	0	1	0	1	0	1	
0	1	1	1	0	1	0	0	
1	1	1	0	0	0	1	0	Horizontal checks
1	0	0	0	1	1	1	0	
0	0	1	1	0	0	1	1	
1	0	1	1	1	1	1	0	

Vertical checks

(a)

1	0	0	1	0	1	0	1
0	1	1	1	0	1	0	0
1	1	1	0	0	0	1	0
1	0	0	0	1	1	1	0
0	0	1	1	0	0	1	1
1	0	1	1	1	1	1	0

(b)

Figure 2.8 Horizontal and Vertical Parity Checks. Horizontal parity check checks its rows and column parity check checks its column. If the bits circled in (b) are changed at the same time, all the parity checks are still satisfied. (From Bertsekas D. and Gallagher R., "*Data Networks*", Prentice Hall Inc., Engelwood Cliffs, New Jersey, 1990).

single row or a single column. Unfortunately, any pattern of four errors confined to two rows and two columns is undetectable as shown in Figure 2.8. This scheme can be used to transmit ASCII characters with row parity bits the same as the last bit of the ASCII character. The column parity checks can be computed by software or hardware. Even this scheme can fail to detect many errors in data transfer because most errors occur as burst errors. This technique can also be used to correct any single error by locating error in column and row.

2.2.4 Parity Check Codes

The idea behind horizontal and vertical parity checks can be extended to parity check codes. The strategy generates parity checks on various subsets of the bits (the rows and columns in Figure 2.8). The transformation from the string of data bits to the string of data bits and parity checks is called a parity check code or linear code [4]. An example is shown in Figure 2.9.

This section briefly discusses how the effectiveness of a code for error detection can be measured. Three parameters, as defined below, are frequently used for this purpose. They are:

(1) The minimum distance of the code (also known as Hamming distance of the code) is defined as the smallest number of errors that can convert one codeword (an encoded bit string which includes both data and parity checks bits) into another. The minimum distance of single parity check codes and horizontal and vertical parity check codes is 2 and 4 respectively.

(2) The Burst detecting capability is defined as the largest integer B of burst errors that can be detected by a code. The burst detecting capability of the single parity check codes is 1 and horizontal and vertical parity checks is 1 plus the length of a row if rows are sent one after the other in a given array of data. For parity check codes let k be the length of the data string and L be the number of parity checks resulting in a frame of $k + L$ bits. It is shown [4] that the probability of an undetected error is equal to 2^{-L}.

s_1	s_2	s_3	c_1	c_2	c_3	c_4
1	0	0	1	1	1	0
0	1	0	0	1	1	1
0	0	1	1	1	0	1
1	1	0	1	0	0	1
1	0	1	0	0	1	1
1	1	1	0	1	0	0
0	0	0	0	0	0	0
0	1	1	1	0	1	0

$$c_1 = s_1 + s_3$$
$$c_2 = s_1 + s_2 + s_3$$
$$c_3 = s_1 + s_2$$
$$c_4 = s_2 + s_3$$

Figure 2.9 A parity check code. Code words are shown on the left. The right equations give the rules for generating the parity checks. (From Bertsekas D. and Gallagher R., "*Data Networks*", Prentice Hall Inc., Engelwood Cliffs, New Jersey, 1990).

2.2.5 Cyclic Redundancy Checks [4, 13]

Cyclic Redundancy Checks (CRC) are error detecting codes. In other words CRC do not correct the errors in the data received at the receiver but only indicate at the site of the receiver that an error has occurred during transmission. The receiver can then ask for a retransmission of the data from the transmit end of the link. Retransmission strategies are discussed in the next section. Error detecting codes also turn out to be more efficient in terms of utilization of the data network.

Cyclic redundancy codes, also known as polynomial codes, are the most commonly used parity check codes for error detection. Let L be the length of check bits and K be the length of data bits. Here we define the polynomial

$$S(D) = S_{k-1}D^{k-1} + S_{k-2}D^{k-2} + \ldots\ldots + S_0 \qquad 2.14$$

where $S_{k-1}, S_{k-2}, \ldots\ldots, S_1, S_0$ are data bits. It should be noted that $S_{k-1}, S_{k-2}, S_1, S_0$ etc. are either 0 or 1. The powers of the indeterminate D indicate the sequence of transmission. High order terms are assumed to be transmitted first. CRC (parity check bits) are defined to be represented in the following manner:

$$C(D) = C_{L-1}D^{L-1} + \ldots\ldots C_1 D + C_0 \qquad 2.15$$

The entire set of bits that are to be transmitted can be represented as

$$X(D) = S(D)D^L + C(D) \quad \text{or} \qquad 2.16$$

$$X(D) = S_{k-1}D^{L+K-1} + \ldots\ldots S_0 D^L + C_{L-1}D^{L-1} + \ldots\ldots + C_0 \qquad .2.17$$

We also define a generator polynomial here as

$$g(D) = D^L + g_{L-1}D^{L-1} + \ldots\ldots + g_1 D + 1 \qquad 2.18$$

The CRC polynomial $C(D)$ is calculated from the information polynomial $S(D)$ by the following operation

$$C(D) = Remainder\,[\frac{S(D)D^L}{g(D)}] \qquad 2.19$$

In the above operation, two remarks need to be made:

1. All coefficients are binary and

2. Arithmetic on coefficients is performed modulo 2. In modulo 2 operation (1+1)=0 and (0-1)=1. Subtraction and addition are the same in modulo 2 arithmetic. In practice the CRC polynomial is calculated by VLSI chips.

$X(D)$ is the transmitted polynomial. Suppose $Y(D)$ is the received polynomial at the end of the communication link. If there are errors in the received data $Y(D)$ will differ from $X(D)$ by a polynomial $e(D)$ where, in modulo 2 operation

$$Y(D) = X(D) + e(D) \qquad 2.20$$

The scheme works because at the receiver

$$\text{Remainder}\left[\frac{y(D)}{g(D)}\right] = \text{Remainder}\left[\frac{e(D)}{g(D)}\right] \qquad 2.21$$

can be calculated. If no errors occur, $e(D)$=0 and the above operation will give the remainder equal to 0.

It can be shown [4] that any CRC with generator polynomial taken equal to the product of a special polynomial called a primitive polynomial and $(D+1)$ has a minimum distance of 4, a burst detecting capability of at least L and a probability of failing to detect errors in completely random strings of 2^{-L}, where L is the number of check bits.

2.3 Retransmission Strategies

When the error detecting codes indicate an error in the received data, the receiver asks for retransmission of the data. This section discusses basic principles of retransmission strategies.

For simplicity it is assumed that all the data packets (data to be transmitted is divided into smaller set of bits called packets, control bits are added in beginning of the packet, called its header, and in the end, called its trailer. The header, packet and trailer make a frame.) with errors are detected by the receiver to be error packets, which can then be retransmitted. There is one more problem that has to be taken care of by the retransmission strategies. This is of data being lost over the transmission channel. There is a possibility that data is lost over the transmission channel and the receiver module should have some procedure to know which data packets have been lost so that they could be retransmitted by the transmitter. In this section three retransmission protocols are discussed:

1. Stop-and- Wait ARQ
2. Go-Back-n ARQ
3. Selective Repeat ARQ

2.3.1 Stop-and-Wait ARQ

The most basic of retransmission protocols is the stop-and-wait ARQ (Automatic Repeat Request). When a packet is received at the receiver, the receiver checks for errors and depending on the received packet the receiver sends an acknowledgment packet (called an ack) when the packet is received error free. If there is an error in the received packet the receiver sends a negative acknowledgment (called a nak). The ack and nak packets also have CRC because errors can occur in the reverse path (from receiver to transmitter) as well.

At this point let us assume that data transmission is taking place between two nodes, Node A and Node B. Node A is sending data packets to node B and node B is sending ack or nak packets back to node A. A good discussion of retransmission strategies is given in [4, 15]

If an error free packet is received at B, B will send an ack packet back to A. If A receives an error free ack packet it can proceed to send the next packet to node B. On the contrary A can receive a nak packet if B receives an error packet or B sends an ack packet which gets transformed into a unrecognizable or a nak packet due to errors in the transmission channel. In the case when A receives a nak packet, A sends the old packet again to node B. If either the packet from A to B or the ack from B to A is lost, A times out and resends the old packet again. The above strategy does not work in cases when A times out and resends the old packet again due to abnormally delayed ack packets from node B. In this case node B receives the same packet twice. If these packets are error free, B has no way to find out that they are the same packets and one has to be discarded. Node B cannot discard one packet even by comparing them because as far as Node B is concerned two different packets could be identical.

To avoid the above and related problems two changes are made in this basic strategy:

1.) Sequence numbers are used in the header of the frame (one set of the data bits sent from node A or node B) to distinguish between successive packets.

2.) The receiver instead of sending ack or nak sends the number of the next packet awaited. Figure 2.10 illustrates the use of stop-and-wait ARQ scheme.

Stop-and-wait ARQ does not utilize the capability of the data network very efficiently. There is a lot of waiting done by nodes at both ends of the communication link in this type of protocol. The nodes should be utilized to do something else while they are waiting for acks or naks or data frames from the node at the other end of the communication link.

2.3.2 Go-Back-n ARQ [4, 13, 15]

Go-back-n ARQ is an improvement over the stop-and-wait ARQ discussed in the previous section. The basic difference between Go-back-n ARQ and stop-and-wait ARQ is that the transmitter can send a predefined number of frames without receiving an acknowledgment from the receiver. The frames, as in stop-and-wait ARQ, are numbered sequentially and the sequence number (say SN) is sent in the header of the frame containing the packet. In a Go-Back-n ARQ protocol a node can send packets up to SN equal to i+n-1 where i is the largest number of the acknowledged packet from the receiver. The receiver accepts packets only in correct order and sends request numbers (say RN) in the header of the frame back to node A. If i is the largest request number RN which has been received by A, then A at that instant of time is only allowed to send a "window" of n packets from i to i+n-1. As i increases due to received subsequent requests from node B the window slides upward. That is why these protocols are often called sliding window ARQ protocols in the literature.

Figure 2.11 illustrates the operation of a Go Back 7 ARQ protocol. It is shown in [4] that the sequence numbers SN and request numbers RN can be sent modulo m with m greater than n as long as frames do not get out of order on the links.

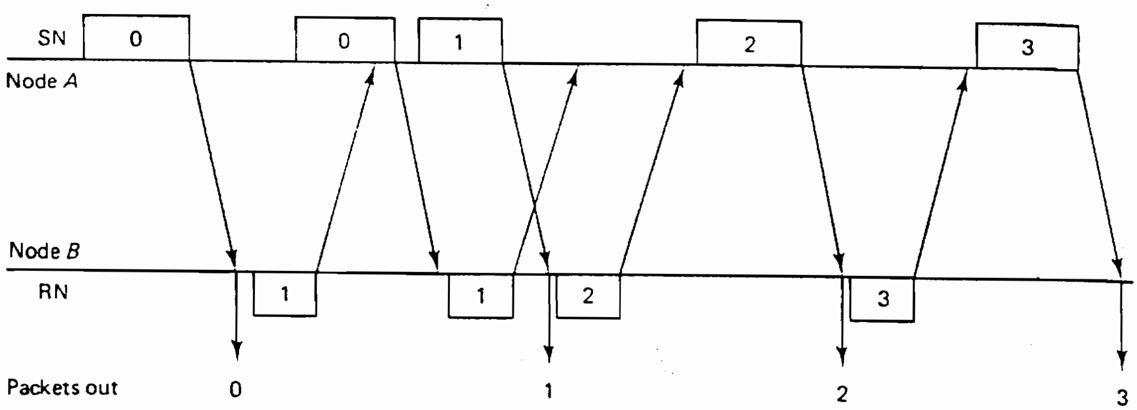


Figure 2.10 Use of sequence and request numbers for stop and wait transmissions from node A to B. Numbers in the rectangles show the sequence numbers at node A and request numbers at node B. (From Bertsekas D. and Gallagher R., "*Data Networks*", Prentice Hall Inc., Engelwood Cliffs, New Jersey, 1990).

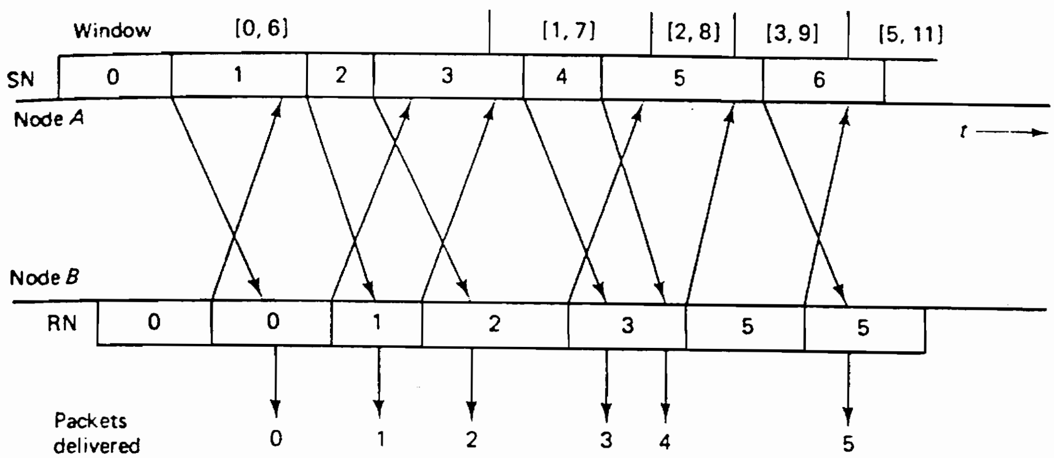


Figure 2.11 Go-Back-7 protocol for node A to node B. The numbers in the brackets constitute the sliding window. (From Bertsekas D. and Gallagher R., "*Data Networks*", Prentice Hall Inc., Engelwood Cliffs, New Jersey, 1990).

2.3.3 Selective Repeat ARQ

The third protocol which is an improvement over Go-Back-n ARQ protocol is the selective repeat ARQ protocol. In this protocol the receiver is allowed to receive all the error-free frames even if they arrive out of order. In a Go-Back-n protocol, in case of reception of an error frame, at least one round trip delay worth of frames has to be retransmitted in the case of a single error detected at the receiver. The

selective repeat ARQ protocol can have appreciably better efficiency than the protocols that have been discussed before especially in the two following cases:

1. On communication links where small single error probabilities per frame are difficult to achieve or maintain.
2. On communication links where the number of frames transmitted in a round trip delay time is very large (e.g. high speed links and satellite links).

If p is the probability of frame error, the expected number η of the packets delivered to B per frame from A to B is bounded by

$$\eta \leq 1 - p \qquad 2.22$$

2.4 The ISO Reference Model [17]

The International Standards Organization has proposed the reference model of Open Systems Interconnection (OSI) for international standardization. This section deals with the proposal in some detail. The proposal has seven layers, as shown in Figure 2.12. Each layer is discussed briefly below.

1. Physical Layer: Physical Layer is concerned with a reliable transmission of bits (1 and 0) over a communication channel. Decisions have to be made regarding bit mapping into signals, pulse timings, characterization of communication channels etc.

2. The Data Link Layer: The data link layer provides the Network layer, the layer above the data link layer, with an error free data communication channel. The data link layer's job is to get bits from the physical layer which could have some errors, then to detect and correct the errors, if any.

3. The Network layer: The network layer controls the operation of the subnet. A subnet basically consists of communication computers and transmission lines. It takes care of the routing of data packets, the units of information communicated between the computers. The network layer has to ensure that the data packets are received at the correct destination.

4. The Transport Layer: The transport layer divides the information into data packets and at the receiver site reconstitutes the packet in correct order into the original form of information in the most efficient way.

5. The Session Layer: The session layer takes care of the set up of a session, its maintenance, its successful closure and billing.

6. The Presentation Layer: The presentation layer performs functions like data compression, data encryption or file conversion to circumvent problems arising from incompatible file formats used by different computers.

7. The Application Layer: The application layer is highly dependent on the individual user. For

example Banking and Airline reservations will have different application layer functions. An application layer software may decide how to divide problems into smaller processes and utilize the computing resources and how to present information to the human user.

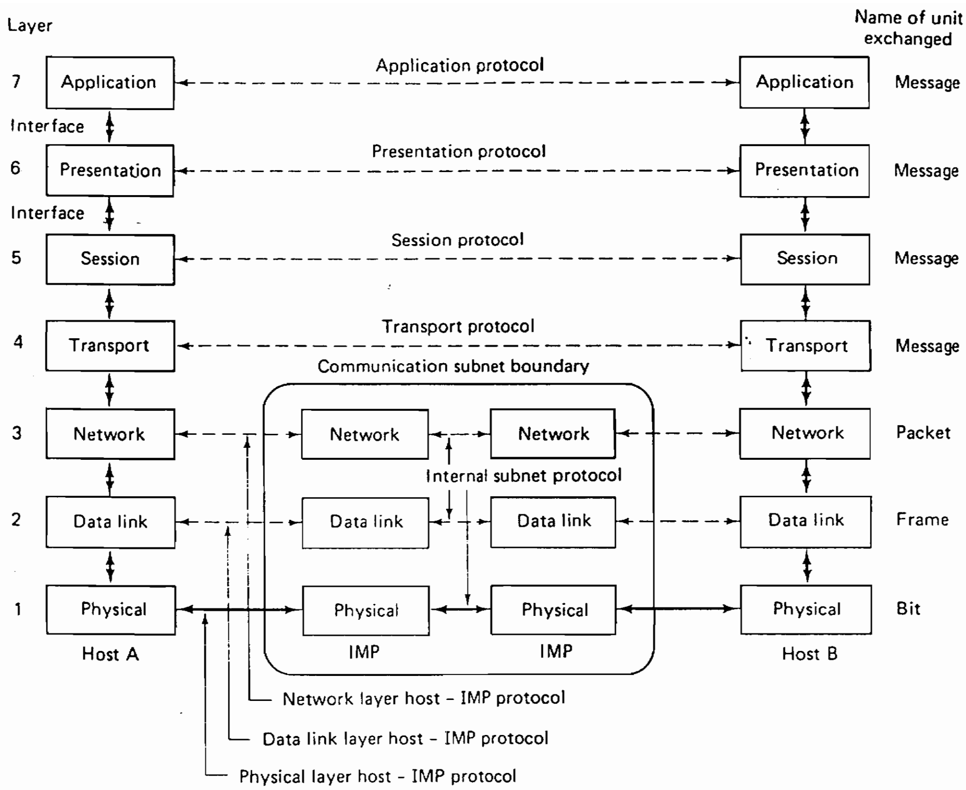


Figure 2.12 Seven Layer OSI network architecture. Each layer communicates virtually with the next higher layer. (From Tannenbaum, A.S., "*Computer Networks*", 2nd. ed., Engelwood Cliffs, N.J., Prentice-Hall, 1988).

2.5 X.25 Network layer standard [16, 17]

The X.25 network layer standard is widely used to connect users' data terminal to public networks. A public network is similar to the public telephone system and often shares the public telephone system capabilities. The public network is owned and run by government or private companies. The X.25 Network layer standard was developed by the International Consultative Committee on Telegraphy and Telephony (CCITT). Three key protocols (set of rules) have been standardized.

- The physical layer protocol, called X.21
- Data Link Layer called LAPB
- Network Layer Protocol

These three protocols are collectively called the X.25 standard. In this section all the three protocols are briefly discussed.

2.5.1. X.21 Physical Layer Protocol

X.21 is a digital signaling interface. The CCITT assumes that sometime in the near future, carriers will connect all the subscribers through digital lines extending to the subscribers premises. In this protocol a customers' computer, or any terminal where data originates, is called a DTE (Data Terminal Equipment). The carrier's equipment which exchanges signals with the DTE is called a DCE (Data Circuit terminating Equipment). X.21 defines eight wires connecting a DTE to a DCE. The physical connection has 15 pins but only 8 have been defined in X.21. Signals lines are shown in Figure 2.13. These lines have following functions:

- T and C lines used by the DTE to transmit data and control information respectively.

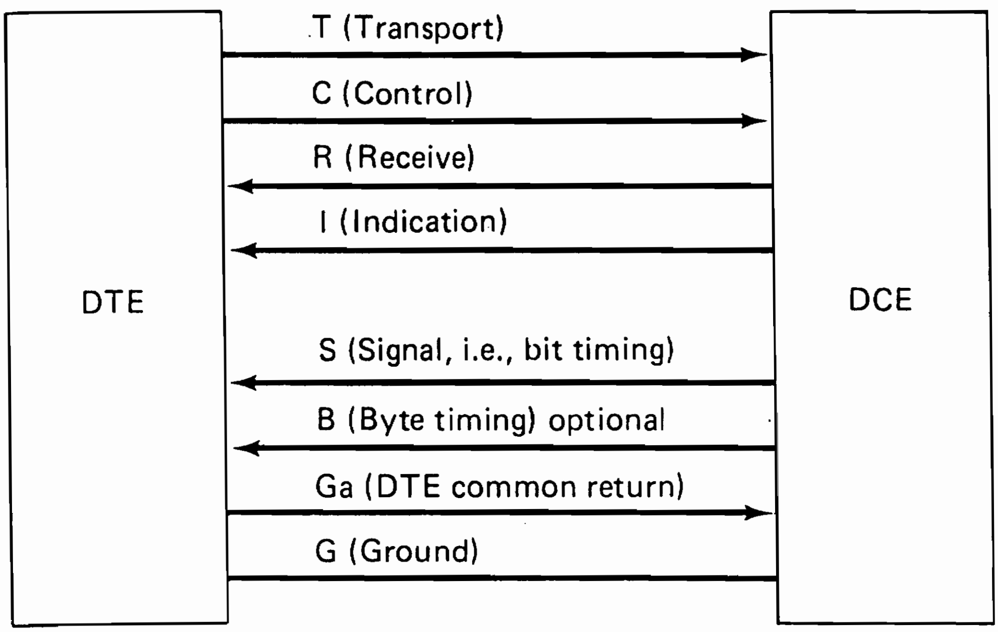


Figure 2.13 Signals lines in the X.21 protocol. (From Tannenbaum, A.S., "*Computer Networks*", 2nd. ed., Engelwood Cliffs, N.J., Prentice-Hall, 1988).

- R and I used by the DCE to transmit data and control information respectively.
- S line contains timing information for the DTE which consists of a signal stream emitted by the DCE.
- B line groups the bits into 8 bit frames.

Figure 2.14 is an example of how the DTE places the call to a distant DCE and how the connection is finally terminated. Initially the four signaling lines are all 1. CCITT calls a 1 as off and a 0 as on. To initiate a call the DTE sets T to 0 and C to ON. When the DCE is ready to accept the call, it transmits the ASCII "+" character on the R line. When the DTE recognizes + on the R line it sends the ASCII address of the distant DTE on the T line, one bit at a time. DCE then sends call progress signals to inform the DTE of the result of the call. These signals consist of two digit numbers, the first of which indicates the general class of the result and the second provides detailed information. The general classes include information like: call put through, try again, call failed, short term network congestion and long term network congestion. If the call placement is successful, DCE sets I to ON to indicate that Data Transfer can take place now. At this point either DTE can send data because a full-duplex digital connection has been established between the two DTE's. A full duplex connection means that the communication can take place simultaneously in both directions. At this point any DTE can switch its C line to OFF indicating that it has no more data to send, but it should be ready to receive data from the other DTE. In the example of Figure 2.14, the original DTE turns its I line to OFF. When the distant DTE also switches its I line to OFF, the DCE at the originating side sets R to 1. Finally, the DTE sets T to 1 to put the interface back in the original idle state.

2.5.2 Data Link Layer Standard

Data Link Layer standard is basically designed to deal with transmission errors on the telephone line between the users' equipment and the public network. The DLC has a number of

Step	C	I	Event in telephone analogy	DTE sends on T	DCE sends on R
0	Off	Off	No connection-line idle	T = 1	R = 1
1	On	Off	DTE picks up phone	T = 0	
2	On	Off	DCE gives dial tone		R = "+ + + . . . +"
3	On	Off	DTE dials phone number	T = address	
4	On	Off	Remote phone rings		R = call progress
5	On	On	Remote phone picked up		R = 1
6	On	On	Conversation	T = data	R = data
7	Off	On	DTE says goodbye	T = 0	
8	Off	Off	DCE says goodbye		R = 0
9	Off	Off	DCE hangs up		R = 1
10	Off	Off	DTE hangs up	T = 1	

Figure 2.14 Typical X.21 usage. (From Tannenbaum, A.S., "*Computer Networks*", 2nd. ed., Engelwood Cliffs, N.J., Prentice-Hall, 1988).

variations but they all have some basic similarities. Hence in this section we will deal with the basic structure of the DLC standard protocols. It should be mentioned here that DLC protocols like HDLC (High Level Data Link Control), SDLC (Synchronous Data Link Control), ADCCP (Advanced Data Communication Control Procedure) and LAP (Link Access Procedure) are all similar protocols with slight differences which sometimes make them incompatible. CCITT modified LAP to LAPB and made LAPB part of the X.25 network interface standard.

All these protocols are bit oriented as opposed to character oriented which means that these protocols will have frames containing an arbitrary number of bits. These protocols use the frame structure as shown in Figure 2.15. The different fields in the frame structure have the following meanings:

- Address Field: It is used on multidrop lines, where the terminal has to be identified. For point to point lines it is sometimes used to distinguish between commands and responses.
- Control Field: Control Field is used for informing sequence numbers, acknowledgments etc.
- Data Field: Data field contains data that has to be communicated. Data field can be of any length but is generally kept under some limit controlled by various factors such as bit error rate, chechsum capabilities, data transmission rate in bps (bit per second) etc.
- Checksum Field: Checksum field is similar to the cyclic redundancy code with CRC-CCITT as generating polynomial.
- Each Frame begins and ends with a special bit pattern (called flag) 01111110 to indicate the beginning and end of the frame. To remove this sequence from the data field (which can contain any arbitrary pattern of bits) the transmitting hardware inserts a 0 in the data stream every time it encounters five consecutive 1's. This is called bit stuffing. The receiver on the contrary destuffs the 0 bit after it encounters five consecutive 1's.

There are three kinds of frames in the protocol: Information, Supervisory and Unnumbered. The basic differences lie in the control field of these frames, hence these control fields are now

8	8	8	$\geqslant 0$	16	8
01111110	Address	Control	Data	Checksum	01111110

Figure 2.15 Typical frame format for a bit-oriented protocol. (From Tannenbaum, A.S., "*Computer Networks*", 2nd. ed., Engelwood Cliffs, N.J., Prentice-Hall, 1988).

discussed. The three kinds of control fields are shown in Figure 2.16.

The control field of an information frame is shown in Figure 2.16 (a). The seq field here indicates the frame sequence number. The Next field is the acknowledgment of the data received. The P/F bit is the poll/final bit. It is set to 1 in the beginning when the computer polls for a terminal to respond. The terminal responds by setting its P/F bit to 1. The P/F bit is also set to 1 when the computer is sending its final frame. The control field of a supervisory frame is shown in Figure 2.16 (b). Here the type field distinguishes between various types of supervisory frames.

Type 0 is an acknowledgment frame used when the receiving terminal does not have to send data.

Type 1 is negative acknowledgment indicating an error has occurred during transmission. The Next field in this case indicates the sequence number of the first frame received incorrectly.

Type 2 is Receive Not Ready (RNR). It acknowledges all the frames received. Next indicates the number of the frame expected. RNR means that the receiver is temporarily out of order and can't accept any more frames for the time being.

Type 3 is the Selective Reject frame. It asks for retransmission of a specific frame indicated by the Next field.

The third type of frame is the unnumbered frame as shown in Figure 2.16 (c). This kind of frame

is used to control information flow between the DTE's. Five bits are available in the type field which provides different kinds of information and requests. Some examples are DISC (disconnect) that allows a machine to announce that it is going down. The reverse is a command SNRM (Set Normal Response Mode) which allows a terminal to announce that it has just come back on the network which is in a Master-Slave communication mode. SABM (Set Asynchronous Balanced Mode) in the type field indicates that both communication parties are equal and resets the line for communication between 'equals' (for e.g. between two computers) as opposed to a master-slave communication style which is communication between a host computer and a terminal.

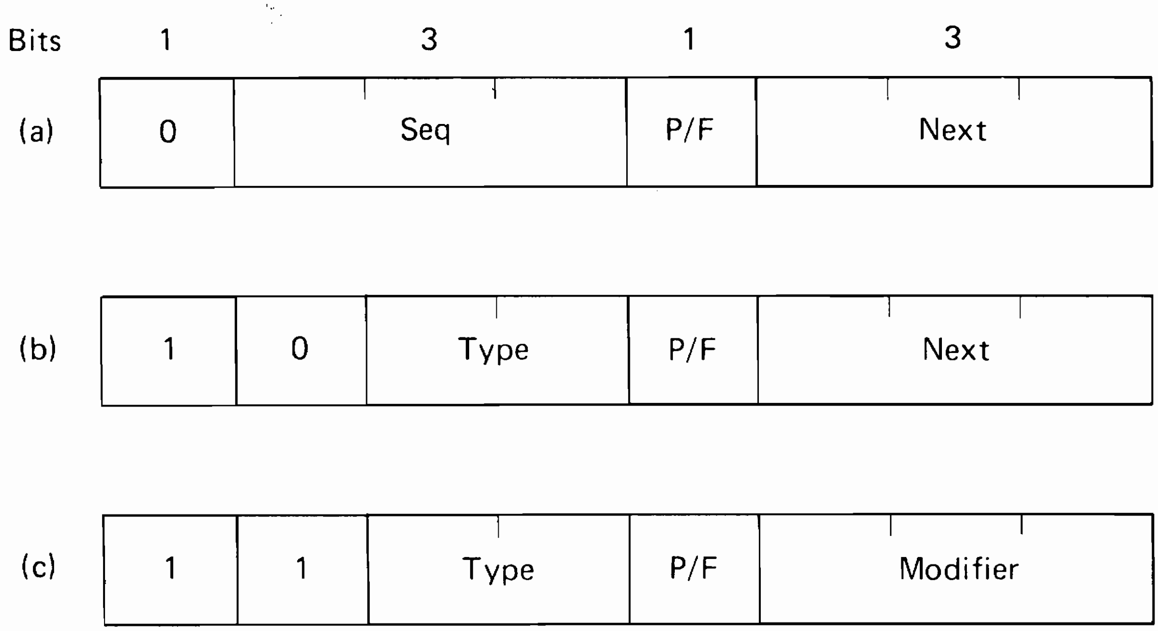


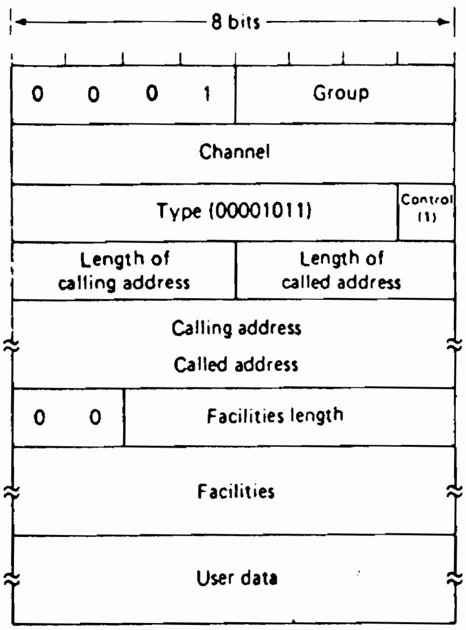
Figure 2.16 Control field of (a) an information frame, (b) a supervisory frame, (c) an unnumbered frame. (From Tannenbaum, A.S., "*Computer Networks*", 2nd. ed., Engelwood Cliffs, N.J., Prentice-Hall, 1988).

2.5.3 The Network Layer in X.25 [17]

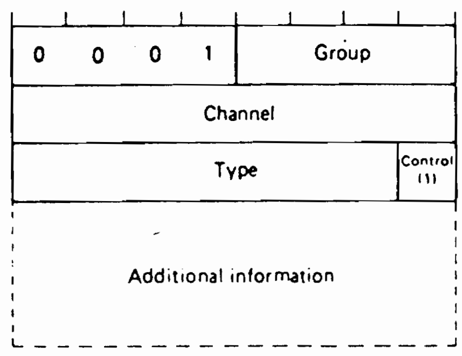
The Network layer in X.25 protocol deals with issues like addressing, flow control of data packets, delivery confirmation, interrupts etc. The format of X.25 data packet is shown in Figure 2.17. The X.25 network layer delivers this packet to the data link control layer of the X.25 protocol which adds a frame header and trailer to this packet structure and then transmits it into the network. The header and trailer added by the X.25 DLC layer have already been discussed before.

The first bit Q in the packet header has the value 1 for control packets at the transport and the higher layers and is 0 for control packets at network layer and for data packets. The next bit shown as D, conveys whether RN has end to end significance or only link significance. It should be noticed here that the path from source node to destination node may pass through a large number of nodes. A link then would be defined as a communication channel between two nodes in the path from the source node to the destination node. D=1 signifies end to end nodes whereas D=0 indicates link significance. The two bits in the modulus field indicates whether the modulus is 8 or 128 for SN and RN. If modulus is 128 the third byte is extended to two bytes to provide seven bits for SN and seven bits for RN. The second byte and last half of the first byte (a total of 12 bits) is the virtual channel number. A virtual channel number is the number assigned to a channel from one node to the next node in the path between the source and destination nodes. This virtual channel number is assigned at every node to the packet hence it keeps on changing at every node in the predefined path of a packet. The third byte of the packet is very similar to the control byte of the standard DLC's. SN and RN refer to the sequence numbers within a session.

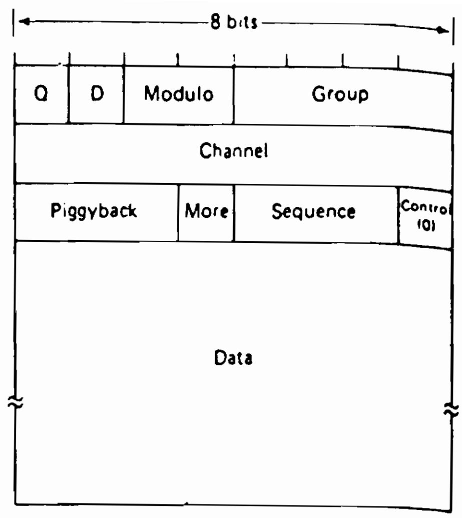
On the network layer many sessions may be taking place at the same time. A session is a communication transaction between end to end nodes. The network layer passes its packets to the DLC layer. The control bit C=0 for data packets and 1 for the control packets. The bit in the M field called the More field is 0 when the packet is the last packet of a session otherwise M is 1. The control packets



(a)



(b)



(c)

Type	Third byte
Data	PPPMSSS0
Call request	00001011
Call accepted	00001111
Clear request	00010011
Clear confirmation	00010111
Interrupt	00100011
Interrupt confirmation	00100111
Receive ready	PPP00001
Receive not ready	PPP00101
Reject	PPP01001
Reset request	00011011
Reset confirmation	00011111
Restart request	11111011
Restart confirmation	11111111
Diagnostic	11110001

(d)

Figure 2.17 X.25 packet formats. (a) call request format. (b) control packet format. (c) data packet format. (d) Type field. (P = Piggy back, S = Sequence and M =More). (From Tannenbaum, A.S., "*Computer Networks*", 2nd. ed., Engelwood Cliffs, N.J., Prentice-Hall, 1988).

are of similar kind as in the DLC layer but they refer to the Network layer. There are control packets like receive-not-ready, receive ready, reject, call request and call accepted control packets.

2.6 File Transfer Protocols [13]

A file transfer protocol is an applications layer protocol that transfers files between users' terminals. Important parameters that need to be well defined in a protocol are:

Block size: This defines the amount of data that will be sent in one continuous block.

Duplex: This defines the communications sequence during a file transfer. The data can flow in both directions in a full duplex protocol and in only one direction in a half duplex protocol.

Handshaking: This concerns with the response of the terminals (at both ends) to the information received,

either correct or erroneous. This also concerns with set up and terminations of the communications session.

Error detection: This defines how the data errors at the receiver will be detected. Data errors need to be detected in both data that needs to be sent and control packets that enable the data transfer.

Error Correction: Error correction procedures correct the errors in the data received with errors. These procedures generally call for retransmissions from the transmitter.

We mention here essential properties of the protocols we have used in our data transfer software.

1. Xmodem: Xmodem protocol is a half duplex protocol. It has a block size of 128 bytes. The protocol attempts to use either CRC or Checksum for error checking. The strategy finally used will depend upon response from the other side of the communication channel. Xmodem does not indicate the file size, name or date to the receiver.

2. Ymodem Batch: Ymodem Batch protocol is a half duplex protocol. It has a block size of 1024 bytes. The error checking scheme used here is CRC. Ymodem Batch first sends a packet numbered 0 which

contains information about the file to be transferred like file name, size, time and date. Ymodem batch can send multiple files one after the other in one communications session.

2.7 Multiaccess Communication [4]

In this section we discuss communication media where the signals received at one node depend upon the signals transmitted by more than one node. Such media, called multiaccess media, underlie the basic functioning of local area networks (LANs), Metropolitan Area Networks (MANs), Satellite Networks and Radio Networks. In these kind of multiaccess channels another sublayer called the medium access control (MAC) sublayer, between the Data Link Control (DLC) and the physical layer, takes care of the allocation of the mutiaccess medium among various modes. Now we briefly discuss some of the mutiaccess media used in the project work.

1. Satellite Channels: In a typical Satellite Network many earth stations can communicate with a geosynchronous satellite. The signal received by the satellite receiver is relayed back to the earth stations at the ground. Signal contention is circumvented by the use of a combination of the following strategies :

a. Use of separate antenna beams: different areas on the ground are served by different antenna beams from the satellite. This way signals from one area on earth do not affect the signals from other areas.

b. Use of FDM and TDM techniques: Frequency Division Mutiplexing or Time Division Multiplexing schemes can be used which enable different earth stations to use the satellite independently.

A combination of the above mentioned techniques is generally used so that the earth stations in a specified area can independently share the satellite transponder.

All these techniques inherently increase delay and do not utilize the medium most efficiently. Delay can be reduced and utilization can be increased if the medium is used on a demand basis. A good discussion of multiaccess communication is given in [4, 5].

2. Multidrop Telephone Lines: one example analyzed in this project work is a communication channel using multidrop telephone lines. Such a network connects a primary node with a number of secondary nodes through a pair of wires. When the primary node transmits a message it is received by all the secondary nodes. A return pair of wires carries the sum of transmitted signals from all the secondary nodes back to the primary node. The mode of operation considered in this project work is the polling method. In this method the primary node requests information from the secondary node in some predefined order. Each secondary mode answers to its poll by either sending data back to the primary node or indicating that it has no data to send. This strategy avoids collision of data from different secondary nodes because at one instant of time only one secondary node is answering the poll from the primary node. The strategy avoids interference successfully but adds a certain amount of inefficiency into the system because some time is lost in polling from the primary node and when the secondary node merely sends a message that it has no data to send.

CHAPTER 3

Delays in Data Networks

This chapter deals with the delay in data networks. Delay is a very important performance measure of a data network. In this project work the satellite network and the multidrop telephone network have been analyzed on the basis of delay in the packet delivery for each network. Concepts developed here will be utilized in Chapters 6 and 7 to analyze the distance education data network.

The total delay in the reception of a data packet can be divided into the following four components [4].

1. Processing delay: This is the time taken by the node to assign the packet to one of the outgoing transmission lines.
2. Queuing delay: This delay is the time a packet stays in a queue at one of the outgoing transmission lines.
3. Transmission delay: This is the time between the first and last bits of the packet being transmitted.
4. Propagation delay: This is the time taken by the packet to propagate on the communication channel between the transmitter and the receiver.

In this section we consider only the queuing and transmission delays in data packets. The reason for this is that we are interested in how traffic on the subnet affects the delay in data packets.

Processing delay and propagation delay are independent of the traffic on the network. Here we assume that computation power is not a limiting resource.

3.1 Model for Queuing Delays [4]

In this section we develop a model to analyze the delay and then calculate delay patterns for various kinds of systems. We assume here that packets arrive at a node for servicing and wait in the queue at the node. We will calculate the service time and the time a packet has to wait in the queue to get serviced. The service time is nothing but the transmission delay as defined earlier. To become consistent with the classical queuing theory we make no distinction between a packet and a customer, the process at the node to be the servicing of the customer by the node. We define here the following important quantities for a node in the network:

Quantity	Definition
L	Packet length in bits
C	Link transmission capacity in bits/second
$N(t)$	Number of customers in the system at time t
$\alpha(t)$	Number of customers who arrived in the time interval [0,t]
T_i	Time spent in the system by the ith arriving customer

The time average of $N(t)$ from time 0 to time t can then be defined as

$$N_t = \frac{1}{t}\int_0^t N(\tau)d\tau \qquad 3.1$$

In many networks of interest we find that N_t tends to a steady state time average N as t increases, mathematically

$$N = \lim_{t\to\infty} N_t \qquad 3.2$$

The rate of arrival of packets (customers) is defined as

$$\lambda_t = \frac{\alpha(t)}{t} \qquad 3.3$$

which also reaches a steady state value, mathematically

$$\lambda = \lim_{t\to\infty} \lambda_t \qquad 3.4$$

the customer delay is defined as

$$T_t = \frac{\sum_{i=0}^{\alpha(t)} T_i}{\alpha(t)} \qquad 3.5$$

which reaches a steady state value of

$$T = \lim_{t \to \infty} T_t \qquad 3.6$$

Now we state mathematically the basic theorem in Queuing theory known as Little's Theorem. It states that the basic quantities N, λ , T in a queuing system are related as follows

$$N = \lambda T \qquad 3.7$$

A network is modeled using the following characteristics:

1. The arrival of the customers. The customers could arrive at a constant rate or their arrival could be represented by a probability distribution function.
2. The nature of the probability distribution of the service time. This depends upon the nature of the server and the length of the packets in bits. It should be noticed here that if L is the length of the packet in bits and C is the capacity of node server then L/C is the packet transmission time.
3. The number of servers at the node.

3.1.1 Queuing System

This section briefly discusses the performance of various queuing systems on the basis of the above three characteristics.

3.1.2 The M/M/1 Queuing system

The M/M/1 queuing system consists of a single server and a single queue. A single server means that there is only one transmission line extending from the node. Data packets are assumed here to arrive according to a Poisson process with rate λ per second.

A Poisson process has the following properties. The number of arrivals that occur in disjoint time intervals are independent. If $A(t)$ represents the total number of arrivals from 0 to time t, then

$$P\{A(t+\tau)-A(t)=n\}=e^{-\lambda t}\frac{(\lambda t)^n}{n!} \qquad 3.8$$

In other words the probability that the number of packets that arrive in the time interval between $t+\tau$ and t is n, is given by Equation (3.8).

The service time distribution is exponential with mean $\frac{1}{\mu}$ seconds. If S_n is the service time of the nth customer, then

$$P(S_n \le S)=1-e^{-\mu S} \qquad S \ge 0 \qquad 3.9$$

Service times S_n are mutually independent and independent of the interarrival times of the packets at the node.

In the queuing system nomenclature (A/S/n) stands for the following: A indicates the nature of the arrival process. Here, M stands for a memoryless or Markovian process which means that the process has a Poisson distribution. G stands for general distribution of interarrival times and D stands for deterministic interarrival times. The second letter S stands for probability distribution of the service times (M,G or D) and the last number (n) indicates the number of servers.

We define now the parameter ρ which is the ratio of λ and μ

$$\rho=\left(\frac{\lambda}{\mu}\right) \qquad 3.10$$

ρ is called the utilization of the network.

The following relations between the quantities defined before can be derived [4]

$$N = \frac{\rho}{1-\rho} = \frac{\lambda}{\mu - \lambda} \quad 3.11$$

where N is the expected number of customers in the system.

Little's formula gives the average delay per customer

$$T = \frac{N}{\lambda} = \frac{1}{\mu - \lambda} \quad 3.12$$

Average waiting time (time in queue only) is given by

$$W = \frac{\rho}{\mu - \lambda} \quad 3.13$$

Little's theorem also gives the average number of customers in queue

$$N_Q = \lambda W = \frac{\rho^2}{1-\rho} \quad 3.14$$

The average number of customers in the service is given by

$$N_s = N - N_Q = \frac{\rho}{1-\rho} - \frac{\rho^2}{1-\rho} \qquad 3.15$$

$$= \frac{\rho(1-\rho)}{(1-\rho)} = \rho$$

which justifies the name for ρ to be the utilization of the communication link.

Another very common queue seen in the network is M/G/1 queue. It is a more general case of queuing in the network and so proves to be more helpful. An M/G/1 queue like the M/M/1 queue has a Poisson arrival process but its service times have a general distribution. The following assumptions are still valid: service times are

- identically distributed
- mutually independent
- independent of interarrival times

Let X_i be the service time of the ith customer. Then we define the average service time

$$\overline{X} = E\{X\} = \frac{1}{\mu} \qquad 3.16$$

the second moment of service times is defined as

$$\overline{X^2} = E\{X^2\} \qquad 3.17$$

The Pollaczek-Khinchin formula gives the waiting time in the queue for a packet for the M/G/1 queue case

$$W = \frac{\lambda \overline{X^2}}{2(1-\rho)} \quad 3.18$$

where ρ is the utilization

$$\rho = \frac{\lambda}{\mu} = \lambda \overline{X} \quad 3.19$$

The expected total time in the system, in queue and in service, is given by

$$T = \overline{X} + W = \overline{X} + \frac{\lambda \overline{X^2}}{2(1-\rho)} \quad 3.20$$

Pollaczek- Khinchin formula and Little's theorem give the number of customers in the queue, N_Q

$$N_Q = \lambda W = \left(\frac{\lambda^2 \overline{X^2}}{2(1-\rho)} \right) \quad 3.21$$

The total number of customers in the system, N which is the sum of N_Q and the expected number of packets being served can be calculated as follows

$$N = \lambda T = \lambda(\overline{X} + W) = \lambda \overline{X} + \frac{\lambda^2 \overline{X^2}}{2(1-\rho)} \quad 3.22$$

$$= \rho + \frac{\lambda^2 \overline{X^2}}{2(1-\rho)}$$

The result for one special case can be derived using the results for the case where the service times are deterministic (M/D/1 queue). Here we have the service times for the packets X to be all equal to $\frac{1}{\mu}$.

$$X_1 = X_2 = X_3 = \ldots\ldots\ldots = X_i = \frac{1}{\mu} \qquad 3.23$$

so we have

$$\overline{X} = \frac{1}{\mu} \qquad 3.24$$

$$\overline{X^2} = \frac{1}{\mu^2} \qquad 3.25$$

waiting time in queue can now be calculated using the P-K formula

$$W = \frac{\lambda\left(\frac{1}{\mu^2}\right)}{2(1-\rho)} = \frac{\rho}{2\mu(1-\rho)} \qquad 3.26$$

A model for queuing delay has been developed in this chapter. Results for M/M/1, M/G/1 and M/D/1 queues have been outlined. These results will be used in chapter 7 to analyze the delays in the terrestrial data network for distance education.

CHAPTER 4

Polling in Data Networks

The multidrop telephone network analyzed in this project uses a technique called polling which manages the flow of data on the leased telephone lines from multiple sites. Figure 4.1 shows the way all the distant class sites have been connected to the main campus in Blacksburg. In these kinds of network provision must be made to avoid collision of data from all the distant sites.

One way discussed in this chapter is by controlling access, either through a central controller or by passing control from one user to another in a decentralized fashion. Polling is the term that collectively describes this type of multiaccess strategy. A central controller is used to avoid the collision of data by polling all the other nodes in the network one by one in some predetermined or adaptive order. This kind of polling technique is called roll-call polling. In hub polling the control of the network is passed from one node to another. This chapter discusses roll-call polling which has been utilized in the leased telephone line network. Time-delay throughput performance evaluation has been analyzed.

This section deals with one of the HDLC (High Level Data Link Control) protocol modes specifically developed for this kind of polling. HDLC has three modes of operation. One of these, the Normal Response Mode (NRM), is used specifically in a multipoint configuration. In this mode one node computer controls the functioning of the network including all the computers connected to it by

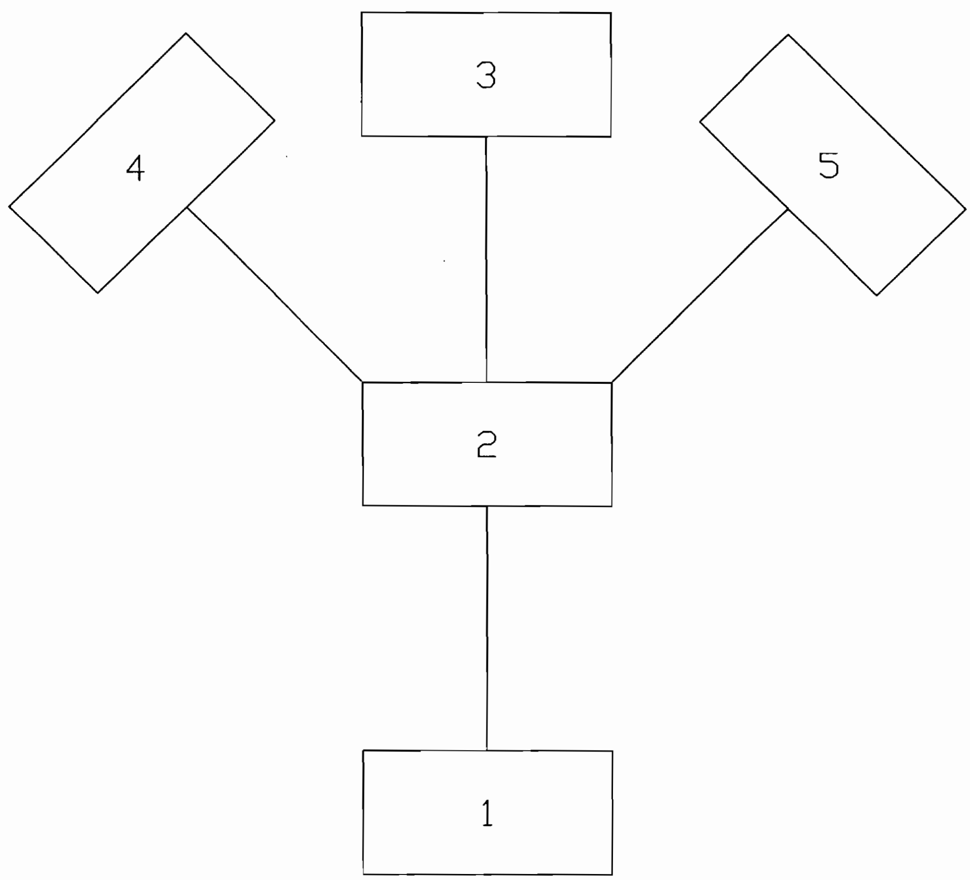


Figure 4.1 Network Topology of leased telephone lines network for Virginia Tech.

1. Main Campus in Blacksburg, Va. 2. Telephone bridge 3,4,5 Distant class sites (only three shown).

issuing commands to them and receiving their responses. The HDLC is similar to the Data Link Control Protocol discussed in Section 2.5, as already has been mentioned there. The HDLC frame format is shown in Figure 4.2. The F field is the 8 bit flag sequence 01111110 at the beginning and end of a frame. The A field indicates the address of the secondary station interrogated by the central supervisor (Primary station). Frame check sequence (FCS) is 16 bits long and is used to detect errors in the frame received. HDLC has three types of frames:

- I or information frames, that carry data in the I field.
- S or Supervisory Frames, that contain no I field.
- U or Unnumbered frames.

Figure 4.3 shows the control field format for HDLC protocol. N(S) and N(R) have the same significance as discussed in the Section 2.4.

NRM includes the use of I frames and ready-to-receive (RR) and not ready-to-receive (RNR) S frames. Other frames used in the NRM are (SREJ) Selective Reject S frame and nonsequenced Unnumbered Information (UI) frames. A simple way in which the NRM functions is now discussed. The P/F bit is used to initiate the polling function. The P bit is set in a frame from the primary station with address of the secondary station to be polled. When the secondary station receives the frame it sends as many I frames as it needs to send according to the retransmission strategies discussed in Section 2.3. The last frame from the secondary station has its F bit set to 1, indicating that it has no more frames to send. If the station polled has no data to send, it sends an RR frame with F=1.

This section analyzes the roll call polling discussed in the previous section. Here we calculate the access delay which is the time a data packet waits in the secondary station queue. The inbound time is the access time plus the frame transmission time.

Time required by one complete cycle of data transfers has following components:

1. Time taken by the transfer of permission from primary station to secondary station.
2. Time required to actually transfer the data from all the secondary stations to primary station.

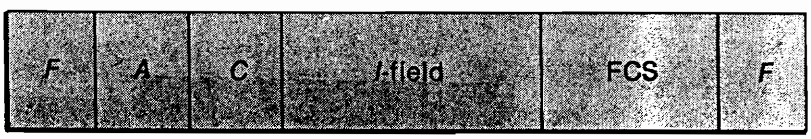


Figure 4.2 HDLC frame format. (From Schwartz, M. "*Telecommunications Networks: Protocols, Modeling and Analysis*", Addison Wesley Publishing Company, 1987).

Bit number	1	2	3	4	5	6	7	8
I-frame	0	$N(S)$			P/F	$N(R)$		
S-frame	1	0	S	S	P/F	$N(R)$		
U-frame	1	1	M	M	P/F	M	M	M

Figure 4.3 Control field format of HDLC. (From Schwartz, M. "*Telecommunications Networks: Protocols, Modeling and Analysis*", Addison Wesley Publishing Company, 1987).

Component number 1 above consists of the time required to transmit the polling message, for the polling message to propagate on the line to the secondary station, the time the secondary station takes to recognize its address and then to take action on the request (start transmitting), and the propagation time required by the message to go back to the primary station. This time for polling from one station to another sequentially for one complete cycle is called the total walk time of the system.

Figure 4.4 shows that the total cycle time t_c consists of alternate walk and transmission times. The ith station is assumed to have walk time equal to w_i and transmission time equal to t_i . The scan or cycle time t_c is then given by

$$t_c = \sum_{i=1}^{N} w_i + \sum_{i=1}^{N} t_i \qquad 4.1$$

Averaging the times will yield

$$\overline{t_c} = \sum_{i=1}^{N} \overline{w_i} + \sum_{i=1}^{N} \overline{t_i} \qquad 4.2$$

$$L + \sum_{i=1}^{N} \overline{t_i} \qquad 4.3$$

L, then is the total walk time of the complete polling system.

It can be shown [18] that

$$\overline{t_c} = \frac{L}{(1-\rho)} \qquad 4.4$$

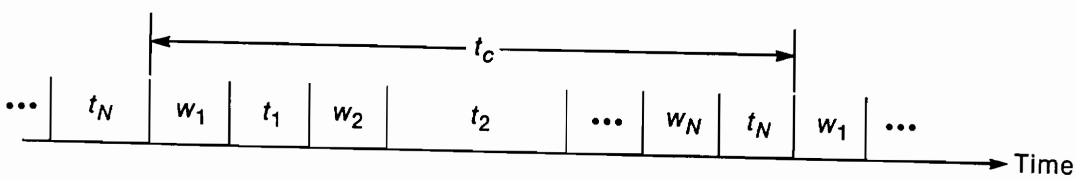


Figure 4.4 Cycle time of a polling system. (From Schwartz, M. "*Telecommunications Networks: Protocols, Modeling and Analysis*", Addison Wesley Publishing Company, 1987).

where

$$\rho = \sum_{i=1}^{N} \rho_i \qquad 4.5$$

$$\rho_i = \lambda_i \overline{m_i} \qquad 4.6$$

λ_i = average arrival rate of the packet at station i (packets/second)

$\overline{m_i}$ = average frame length in units of time $\left(\frac{\bar{l}}{C}\right)$

$\bar{l}$ = the average length of a packet

C= channel capacity in bits/sec

Now we give the result of the analysis of polling systems for a special case when we assume that each station has the same packet-arrival rate λ, the same frame-length statistics and the same average walk-time [18, 19]. The average access delay can be given by

$$E(D) = \frac{\bar{t}_c}{2}\left(1 - \frac{\rho}{N}\right) + \frac{N\lambda\overline{m^2}}{2(1-\rho)} \qquad 4.7$$

$$= \frac{L}{2}\frac{\left(1 - \frac{\rho}{N}\right)}{(1-\rho)} + \frac{N\lambda\overline{m^2}}{2(1-\rho)}$$

where $\overline{m^2}$ is the second moment of the frame length in $(\sec)^2$ and

$$\rho = N\lambda\overline{m} \qquad 4.8$$

Now assume that the polling message transmission time be a fixed value t_p sec. The synchronization time required by each secondary station is t_s sec, and the total propagation delay for the entire N-station system is τ' sec. Then the total walk time for roll call polling is given by

$$L = Nt_p + Nt_s + \tau' \qquad 4.9$$

τ' depends upon the propagation speed of energy in the medium and on the topology of the network. The result stated in this chapter will be modified slightly and will be used to analyze the leased telephone line network in Chapter 7.

CHAPTER 5

Satellite Links

This chapter discusses the basic principles of satellite link design and bit error rate calculation. Section 5.1 develops the theory for satellite link design and Section 5.2 develops the theory for the calculation of bit error rate on satellite links. The theory developed here will be used in Chapter 7 to calculate various parameters for the satellite data network studied in this project .

5.1 Satellite Link Design

In this section techniques are introduced which are used to calculate antenna parameters, transmission power, bandwidth and noise for a satellite link. A good discussion of satellite link design is given in [20, 21]. Figure 5.1 depicts a satellite link.

Consider an antenna that radiates equal power in all directions (an isotropic antenna). It can be shown that flux density at a distance R from the antenna is given by [20]

$$F = \frac{P_t}{4\pi R^2} \ W / m^2 \qquad 5.1$$

where P_t W is the output power of the antenna. If the gain of a transmitting antenna is G_t [20] then

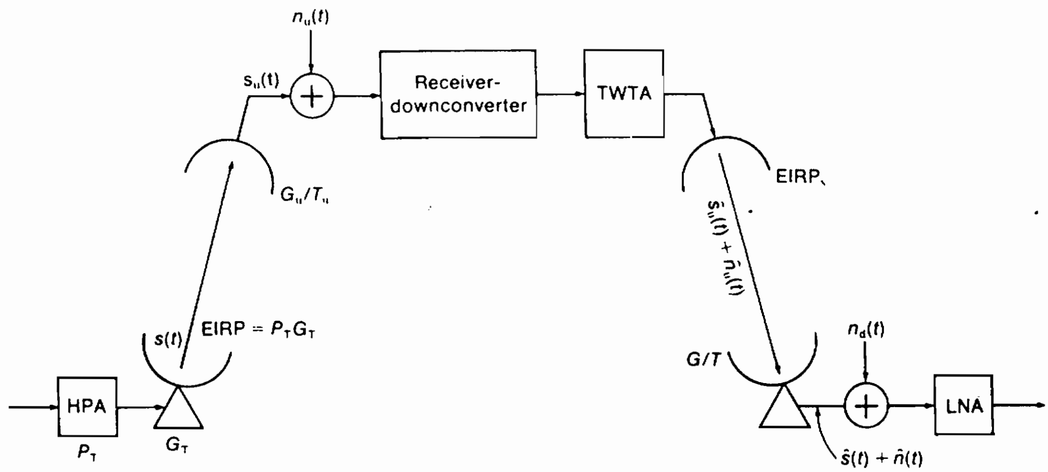


Figure 5.1 Various components in a satellite link design equation. (From Ha., T.T.,"Digital Satellite Communications", 2nd. ed., John Wiley and Sons, New York, 1989).

the flux density at R meters from the boresight of the antenna is

$$F = \frac{P_t}{4\pi R^2} G_t \ W/m^2 \qquad 5.2$$

The power received, P_r, by a receiving antenna is then given by

$$P_r = \frac{P_t G_t}{4\pi R^2} A_e \qquad 5.3$$

where A_e is the effective receiving area given by

$$A_e = \eta A_r \qquad 5.4$$

A_r is the physical aperture area of the receiving antenna and η is called the aperture efficiency of the antenna. The gain of an antenna is given by [20]

$$G = \frac{4\pi A_e}{\lambda^2} \qquad 5.5$$

From Equations (5.3) and (5.5) we get the Friis transmission equation

$$P_r = P_t G_t G_r \left[\frac{\lambda}{4\pi R}\right]^2 \qquad 5.6$$

Here G_r is the gain of a receiving antenna. Path loss is part of the R.H.S. of Equation (5.6). It is given by

$$L_p = \left(\frac{4\pi R}{\lambda}\right)^2 \qquad 5.7$$

The decibel form of the Friis transmission equation is

$$P_r = (EIRP + G_r - L_p) \text{ dBW} \qquad 5.8$$

where

$$EIRP = 10\log_{10}(P_t G_t) \text{ dBW}$$

$$G_r = 10\log_{10}\left(\frac{4\pi A_e}{\lambda^2}\right) dB$$

$$L_p = \text{path loss} = 20\log_{10}(\frac{4\pi R}{\lambda}) \text{ dB}$$

In addition to the losses due to the path we also have losses due to attenuation by rain, losses in receiving and transmitting antennas and loss due to antenna mispointing. Which brings us to the following equation

$$P_r = EIRP + G_r - L_p - L_a - L_{ta} - L_{ra} \text{ dBW} \qquad 5.9$$

where the symbols have following significance:

- L_a = losses due to attenuation in atmosphere
- L_{ta}=losses added by the transmitting antenna
- L_{ra}=losses added by the receiving antenna

The noise power at the demodulator input of the receiver is given by [20]

$$P_n = kT_sBG \tag{5.10}$$

where

- k=Boltzmann's constant=1.38×10^{-23} J/K = -228.6 dBW/K/Hz
- T_s =system noise temperature in Kelvins
- B = Bandwidth of the receiver in hertz
- G = overall RF and IF gain of the receiver

If the signal power is P_r at the input of the RF receiver, the signal power at the demodulator input is P_rG. Hence the carrier to noise ratio is given by

$$\frac{C}{N} = \frac{P_r}{kT_sB} \tag{5.11}$$

Using the Equations (5.6) and (5.11) we get

$$\frac{C}{N} = \frac{P_tG_t}{kB}\left(\frac{\lambda}{4\pi R}\right)^2\left(\frac{G_r}{T_s}\right) \tag{5.12}$$

$\frac{G_r}{T_s}$ is sometime called figure of merit and is shortened often to G/T ratio. Equation (5.12) governs the system design of a satellite link.

5.2 Bit Error Rates on Satellite Links

Satellite links usually employ phase shift keying modulation, which has been discussed in section 2.1 in some detail. Binary Phase Shift Keying (BPSK) and Quadrature Phase Shift Keying (QPSK) are the most popular phase shift keying modulation with satellite link designers.

In PSK modulation [20] the modulator changes the phase of the carrier corresponding to the incoming data stream. The time it takes the modulator to change from one phase to another phase (transition time) and the time spent at the desired phase constitutes a symbol period. The transmitted waveform during a symbol period is called a symbol. The complete set of symbols for a particular type of PSK is called its alphabet. If the number of symbols in a alphabet is M then N_b bits are mapped into one of the M symbols, where N_b and M are related by

$$N_b = \log_2 M \tag{5.13}$$

Standard practice is to make M a power of 2.

The bit error rate (BER) or bit error probability (P_b) is defined as the probability that a single bit sent over the communication link will be received incorrectly. Symbol error probability (P_e) is the probability that a symbol sent over the communication link will be received incorrectly. For BPSK, the following relations have been derived in [20]

$$P_b = \frac{1}{2} erfc\left(\sqrt{\frac{E_b}{N_0}}\right) \qquad 5.14$$

$$P_e = \frac{1}{2} erfc\left(\sqrt{\frac{E_s}{N_0}}\right) \qquad 5.15$$

where E_s is the energy per symbol

E_b is the energy per bit

N_0 is the noise power spectral density (noise power divided by bandwidth)

The complementary error function is given by

$$erfc(x) = \frac{2}{\sqrt{\pi}} \int_x^{\infty} e^{-u^2} du \qquad 5.16$$

P_b and P_e are equal for BPSK because a bit and a symbol are the same in BPSK.

Similarly the results for QPSK as derived in [20] are

$$P_e = erfc\left(\sqrt{\frac{E_s}{2N_0}}\right) \qquad 5.17$$

$$P_b = \frac{1}{2} erfc\left(\sqrt{\frac{E_b}{N_0}}\right) \qquad 5.18$$

The symbols have the same meaning as for the BPSK modulation. These results will be used to analyze the satellite network in Chapter 7.

CHAPTER 6

Data Transfer Strategies

In this thesis two strategies have been considered for data communication between the main campus in Blacksburg and all the distant class sites. The first method is the use of the current network of leased telephone lines which is presently being used for voice connection between the main campus class room and all the class sites. The other alternative is the use of a 19.2 kbps data carrier available through Telstar 401 satellite transponder which is currently being used for carrying TV signals from the TV class room in Blacksburg to all the distant class sites. The following sections explain the systems in detail.

6.1 The Leased Line Network

Three cases have been analyzed that use the leased line network. They are:

- Use of the leased line network for both voice and data communication
- Use of the leased line network for only voice communication
- Use of the leased line network for only data communication

This chapter explains the system and discusses the data protocol developed to control the information flow in the network. In Chapter 7 the data networks discussed in Chapter 6, have been analyzed. The present leased line network for the off campus program of Virginia Tech is shown in Figure 4.1. It consists of leased telephone lines from each class site connected to the telephone bridge. One more leased telephone line connects the telephone bridge to the TV studio in Blacksburg. The studio is located in 281 Whittemore Hall at Virginia Tech's main campus in Blacksburg, Virginia. These telephone lines remain connected to the telephone bridge throughout the one hour and fifteen minutes duration of a class. Students at off campus sites use these lines to get a voice connection with the course instructor during an on going lecture. The network is utilized when somebody at an off-campus site wants to ask a question of the instructor.

6.2 Satellite Network

The other alternative considered in the project is the use of 19.2 kbps data link available via the satellite transponder. The transponder bandwidth is currently utilized in sending the TV signal from the main class room in Whittemore Hall to all the off campus sites. The 19.2 kbps link is on a subcarrier generated at the V.T. uplink site and transmitted to all distant sites.

6.3 Strengths and Weaknesses in the Leased Line Network

The leased line network has been used for some time now. Some strengths and weaknesses have become evident through long term use. The strengths of the network are:

1) There is no delay between the request for a connection and the initiation of a voice session because the leased lines remain connected as in Figure 4.1 throughout the lecture.

2) The network by and large works well and provides reliable voice communication between the off campus sites and Virginia Tech's main campus.

This network has some weaknesses though, namely:

1) There is a lot of noise in the network because up to 37 telephone lines may be connected to the telephone bridge. Their noise is accumulated at the receiver site. There is also a lot of interference of other telephone lines which becomes evident to all the students taking a TV course.

2) The network is used very inefficiently. All the lines remain connected throughout an on-going class but they are only utilized when someone at the off campus site asks a question. That time is hardly ever more than 15 minutes for a seventy five minute class.

6.4 Efficient use of the present network

The present network could be more efficiently used if, during the idle time of the network, data could be transferred between the off campus sites and Virginia Tech's main campus in Blacksburg and vice-versa. This data could be homework solutions from off campus students or the notes an instructor would like to send to off campus class sites. Presently the notes and homework solutions are sent through mail which takes two to five days to reach the destination, or by courier service to achieve more rapid delivery at a higher price.

The advantages of this set up would be :

1) The present network would be used more efficiently.

2) The time lost in communications through mail will be highly reduced making possible real time data communications between all the distant class sites and the main class room in Blacksburg.

This study is undertaken to:

1) Find out if the present network could be used for data communication in addition to voice communication for which it is being currently utilized.

2) Study alternate ways for data communication between off campus sites and the TV class room in

Blacksburg and to compare them with the improved leased line network.

6.5 Alternative Data Transfer Strategies

This project studies two strategies for data communication between the distant class sites and the main campus in Blacksburg. The next few sections study these two alternatives in detail. Software has been written to simulate both the strategies. In both the alternatives the software package Procomm Plus has been used for the point to point file transfer protocols available in the software. Section 6.7 discusses the capabilities of the Procomm Plus software.

6.5.1 Terrestrial Network

The terrestrial network is shown in Figure 4.1. The software that controls the communication sessions between the main class room and all the distant class sites resides in a 486 class PC kept in the main class room in Blacksburg. We will call this PC a Network Controller for obvious reasons. The network will work satisfactorily if it can sustain a data rate of about 9kbps. Procomm file transfer protocols have been used for data transfer. It is assumed that the data will be available in the form of files. These files will be transferred to the network controller. The network controller is a 486 PC with a modem connected to a telephone line. The 486 PC has a source code in the PCplus directory written to take control of all the communication sessions. The software is in the script language of Procomm Plus and controls all the data and voice communications sessions.

The data communication strategy effectively utilizes the polling technique for data communication which has already been discussed in some detail in Chapter 4. At one point in time only one remote site can successfully communicate with the network controller in the TV class room in

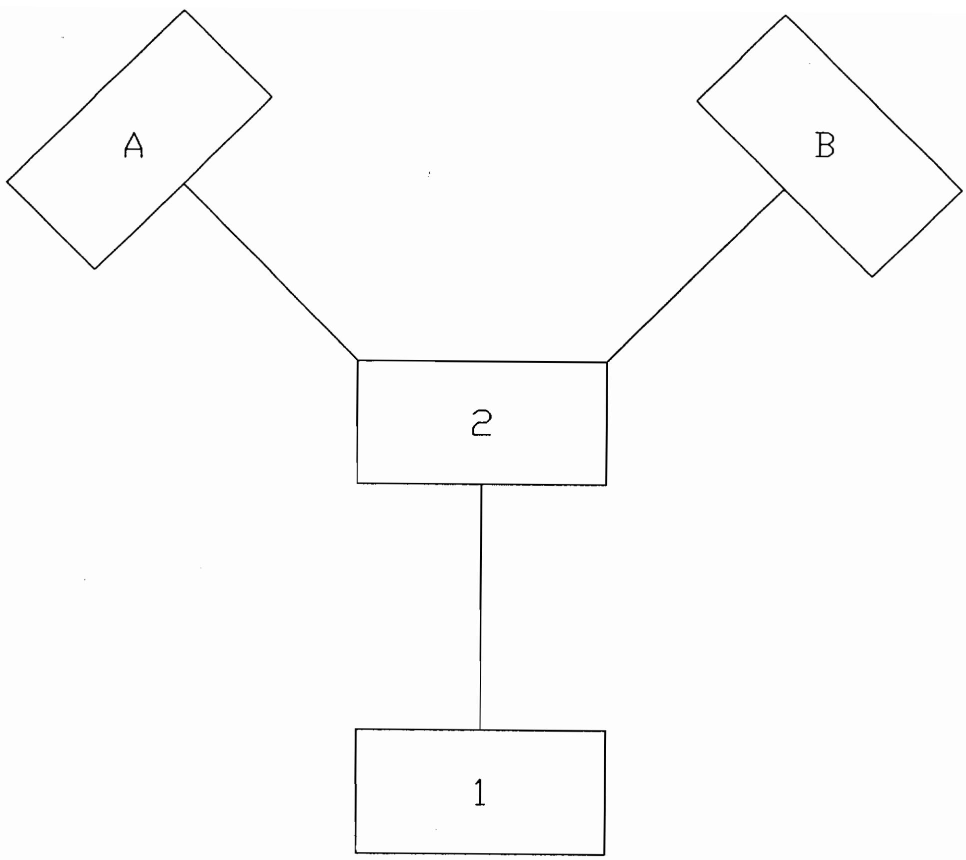


Figure 6.1 Terrestrial Network with two distant class sites A and B. 1. Main campus, Blacksburg, Virginia. 2. telephone bridge.

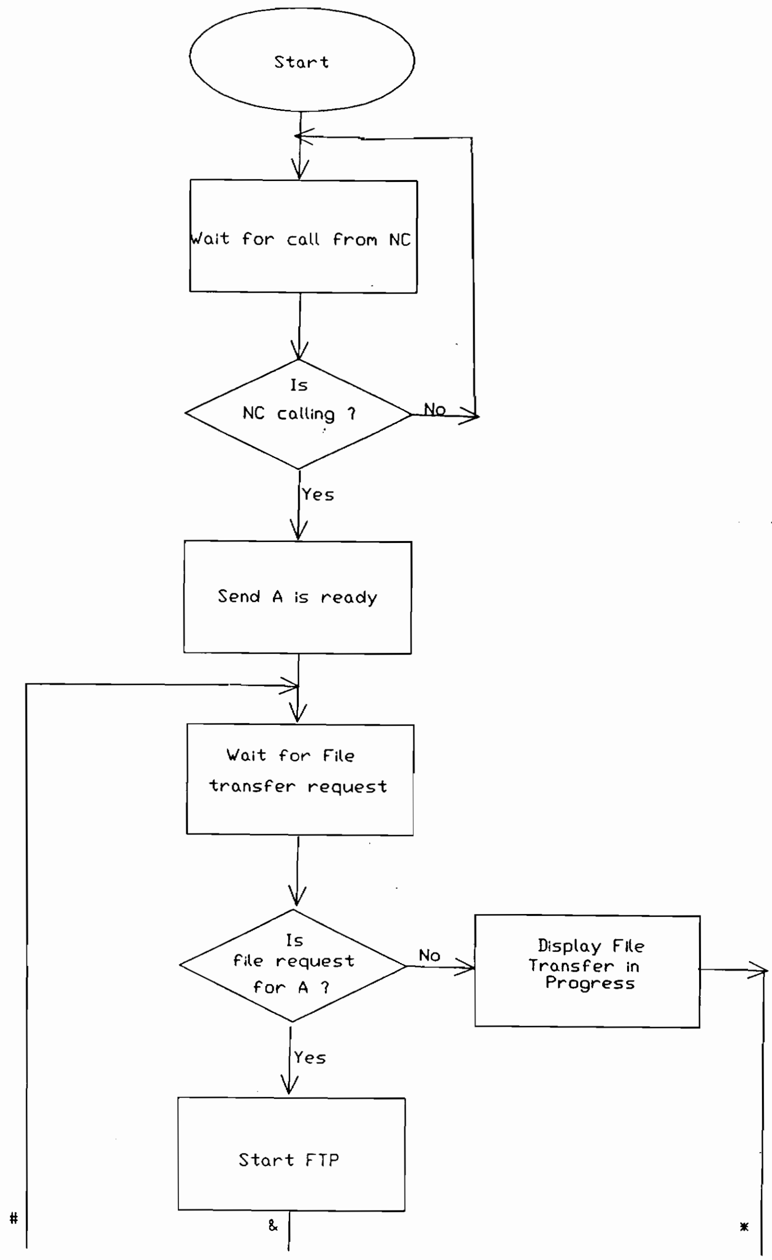


Figure 6.2 Flow chart shows sequence of events taking place at a distant site (A here) continued on next page.

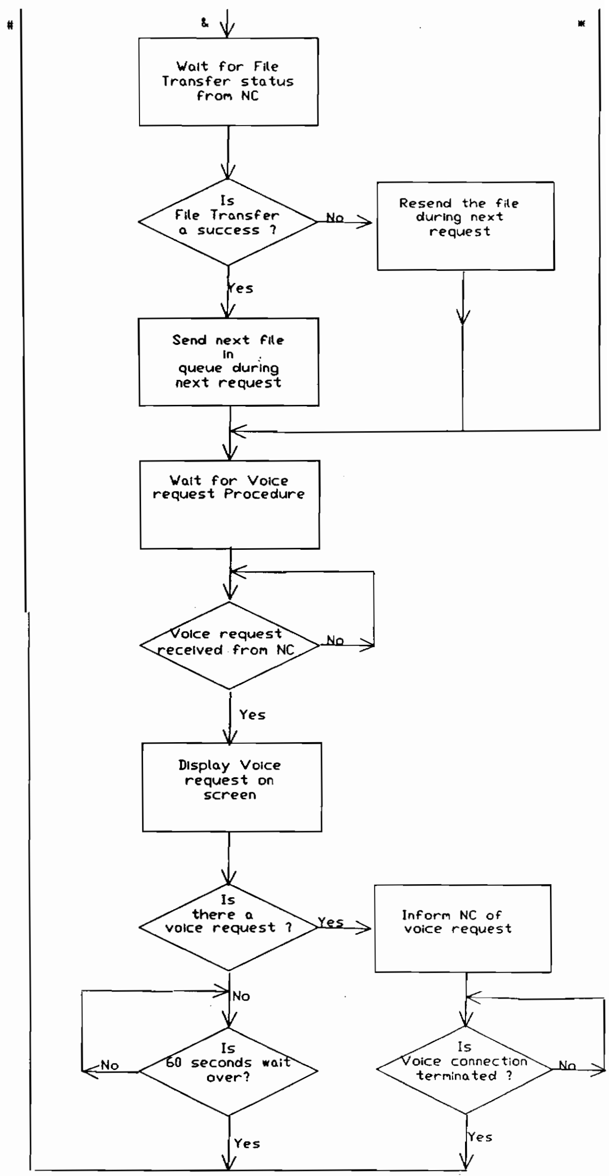


Figure 6.2 continued from previous page.

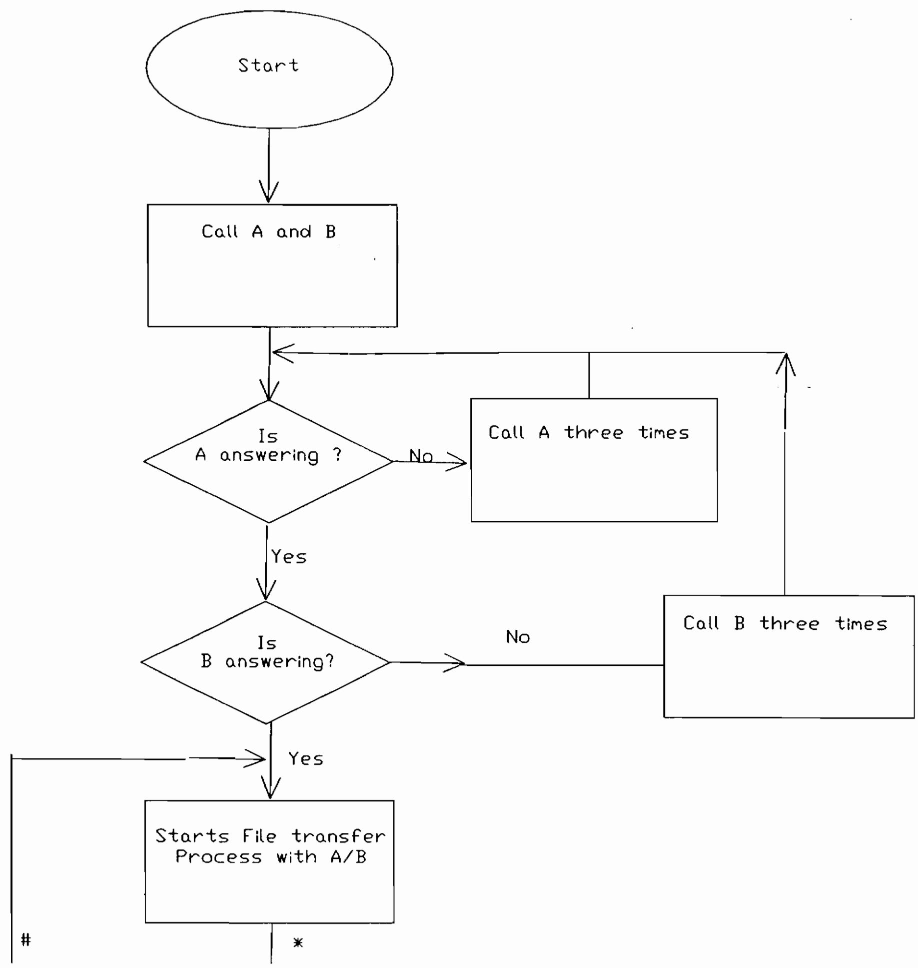


Figure 6.3 Flow chart shows sequence of events taking place in Network Controller (N.C.) continued on the next page.

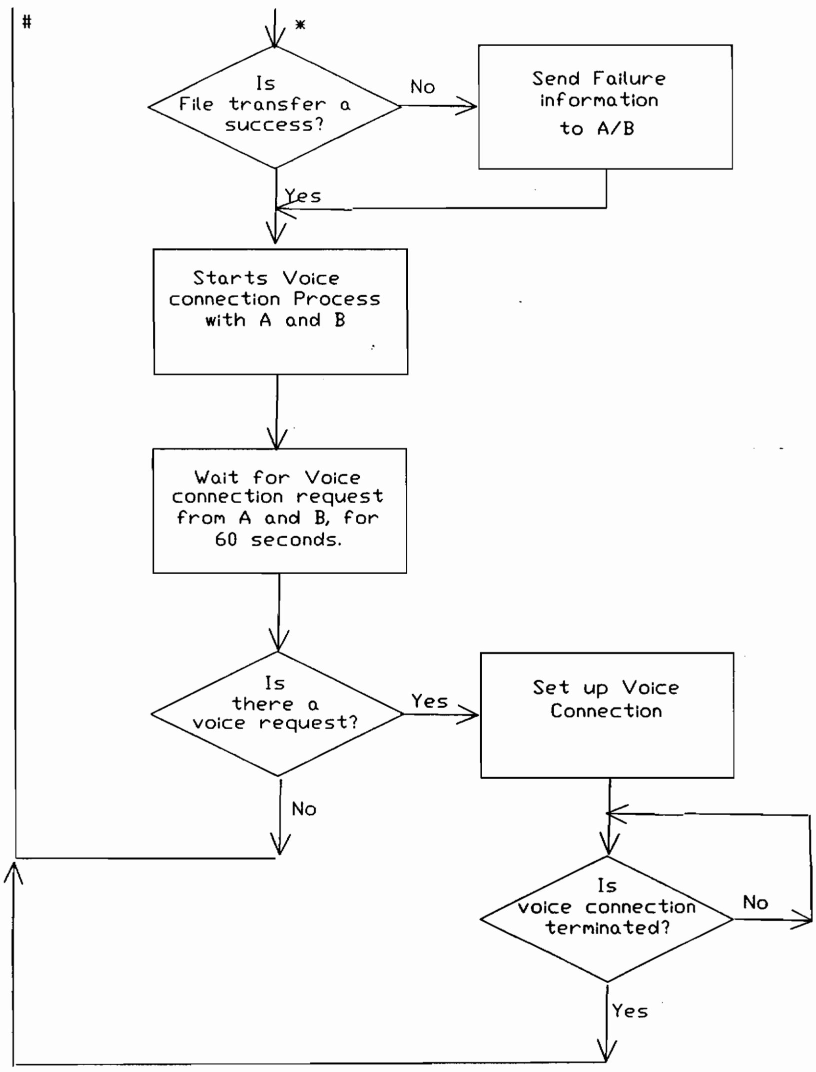


Figure 6.3 continued from previous page.

Whittemore Hall. The strategy can be explained using a simple example from Figures 6.1, 6.2 and 6.3. In explaining the strategy here we assume, for the sake of simplicity, that there are only two distant class sites which are connected to the network controller as shown in the network topology of Figure 6.1. We name these sites as A and B. There may actually be as many as 30 sites connected into the network. The path is set up between site A and the network controller (N.C.) when the network controller at Blacksburg sends a packet specific to site A, indicating that the NC is ready to receive a file from site A. Of course, due to the network topology the packet is also received by all the other computers connected to the telephone bridge, two in this special case. This packet indicates that a path is set up between the network controller and the site computer at A. As the computer at site A receives this packet it initiates the file transfer procedure using the Procomm Plus software. Every time the packet is received the computer at A sends one file from a specific directory to the NC in Blacksburg. Various file transfer protocols that have been used in the project have already been explained in Chapter 2.

After the file transfer is over, the Network controller sends another specific packet to every computer connected to the telephone bridge carrying the information about the status of the file transfer. If the file transfer is a failure the computer at A will have to send this file again at a later time. After one file transfer attempt is over, irrespective of success or failure, the network controller sends a packet to indicate that the network is idle and can be used for voice communication. The network controller waits for one minute for any voice request. If in that duration there is a voice connection request then the connection is set up immediately after the request and the network controller waits for the termination of the voice connection. If during the one minute delay there is no request for the voice connection, the network controller in Blacksburg sends another packet to initiate a file transfer session, to the site computer at A. The site computer sends a particular file depending on the status of the previous file transfer (success or failure). If there are no more files to be sent, the computer sends back a packet indicating that the file transfer is complete. In that case the Network controller initiates the voice request session again. After this session is over it initiates a file transfer with another computer at site B as shown in the Figure 6.1. In this way the whole cycle of file transfer and voice connection repeats itself in this strategy until all the files have been transferred from the distant sites. After all the files have been transferred the network remains open for voice connection all the time. One point to be observed at this juncture is that if there is a request for voice connection during a file transfer it remains pending until the file transfer is over.

It is assumed that the files transferred in these sessions are not larger than 10 Kbyte, in which case it will not take more than 15 seconds for the complete file transfer. If there is a voice connection request just after the initiation of the file transfer the request remains pending for a maximum of 15 seconds. There is no way to reduce this delay once the file transfer has been initiated because the current network topology does not allow more than one point to point communication link. This statement deserves some explanation. Assume that a data transfer session has just begun between the network controller and the site computer at B. Packets are being transferred to and from the site computer at B through the telephone bridge and passing through line 1 to the network controller at the end of the line. Assume now that at this point somebody at site A class room requests a voice connection. If a packet is sent to the network controller at this point from the site computer at A, indicating a voice request from A, there occurs a collision at the telephone bridge between the packets sent by the site computer at B and the packet sent by the site computer at A. This will corrupt all the data that is being sent and will put the network in deadlock. The file transfer protocol in Procomm Plus will wait for an error free packet for some pre-specified time (varies between 5 to 30 seconds) and will then abort the file transfer. That is why it is important that only one computer communicates with the Network controller at a time.

In the network that is currently being used for voice communication only, there is no delay between the request for connection and the beginning of a session. But then the network is being used inefficiently. In the proposed network there is a delay of 15 seconds but the network is being used almost all the time (depending upon the availability of files to transfer). Nevertheless it is a weak point of the proposed strategy.

6.5.2 Data Transfer Through Satellite Links

An alternative to the above strategy considered in this study is the transfer of data using the satellite. Currently the transponder bandwidth available on the satellite is used only for sending TV signals. There is a subcarrier available on the uplink which can sustain a 19.2 kbps data rate. This is a superior channel to the terrestrial network available to us which can only sustain a data rate of 9.6 kbps. One disadvantage of this network is that it can only be used for sending data from Blacksburg to all the distant class sites, but it cannot be used for sending data from the distant sites to the main campus in Blacksburg. The information that needs to be sent from main campus to all the distant class sites is in the form of class notes and exam and homework questions. The two communication strategies are complimentary to each other but can be compared in terms of data rate, power required, failure rate, maintenance expenditure, equipment overhead etc.

The data to be sent is divided into packets of 720 bytes (the reason for this size will become evident in a short while) and transmitted at the rate of 19.2 kbps to all the distant class sites. All the sites receive the data at the same time. This type of data communication, called broadcast data communication was discussed in Chapter 2. Each packet is sent three times to the distant sites because an error can occur in the channel which may render the data unreadable. The probability of this happening is calculated in Section 6.6. All the required information is embedded in the packet being transmitted which is used at the receivers to check for any errors and to reconstitute all the packets to form the original document. If an error is detected in a packet, that packet is discarded and a copy of that packet is used in its place. When all the packets have been successfully reconstituted into a file, the receiver is ready to receive more data from the network controller in Blacksburg. Procomm Plus scripting capability was used to develop software which controls the communication sessions taking place in the Data Transfer. The software present in the Pcplus directory controls all the processes in the network controller. It reads the data into 720 byte files and embeds enough information in the packet, as shown in Figure 6.4, for the receiver to reconstitute the data back to its original form. The software then transmits the packets three times using the satellite channel.

6.6 Error Correction for the Satellite links

This section discusses the error correction used for the satellite link. In the case of digital satellite communication we do not have a return path from the distant sites in the form of a satellite channel. That is why we have to embed enough information in the transmitted packet so that the receivers at all the sites are able to at least detect and if possible correct errors that occurred in the channel during transmission. Two schemes have been analyzed in this section. They are explained in the following sections.

6.6.1 Repetitive Transmission of Data

The error detection has been made possible here by sending each packet three times. The receiver compares the three packets and discards the packet that does not exactly match with the other two packets. It is assumed that out of the three packets received by the receiver only one, at most, will have any errors. To keep the probability of error in the packets low, the packet size is kept low (720 bytes in this case). When the receivers detect an error they discard the whole packet received and use another copy of the packet for reconstitution into the original document. This is why one packet is transmitted three times and the packet length is made not too large so that if a packet has to be

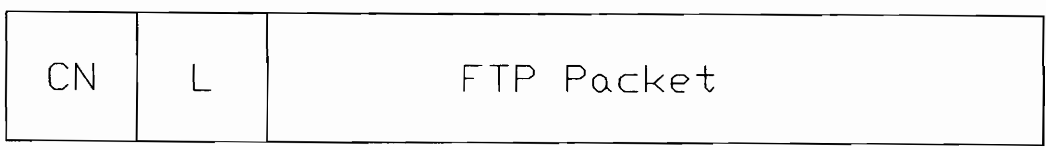


Figure 6.4 A data packet for satellite network

CN is copy number of the packet (1,2 or 3)

L is 1 if it is the last packet of the file being transferred, otherwise it is zero

FTP packet is the file transfer protocol packet.

discarded it does not account for a big loss of transmission time. The packet length is critical for the effective data rate of the system. The probability of an error occurring in one packet of 720 bytes or 5760 bits, given the BER on the link is P_b, is given by

$$P_{bp} = \{1-(1-P_b)^{5760}\} \qquad 6.1$$

It is assumed that out of the three packets received at the receiver only one, at most, will have any errors. The probability that our assumption fails i.e. 2 out of three or all the three packets received have errors, is given by

$$P_f = 3\times\{(P_{bp})^2\times(1-P_{bp})\}+(P_{bp})^3 \qquad 6.2$$

this is the probability that a bad packet is accepted as a good one. As an example, if $P_b = 10^{-6}$ we have P_{bp}=5.743×10^{-3} and $P_f = 9.85\times10^{-5}$. This shows that one packet in approximately 10000 will be accepted in error.

6.6.2 Use of Convolutional Coding

Convolutional coding can be used as a FEC (forward error correction) scheme for the satellite network. This scheme is popular because it can be easily implemented and can provide large coding gains. A good discussion of FEC scheme for satellite links is given in [20, 21].

The performance of convolutional codes with Viterbi decoding can be estimated by using the generating function $T(D,N,L)$. In the augmented state diagram, each branch is associated with a branch gain D^i where i is the weight of the n output bits associated with that branch, a parameter N^j, where j is the weight of the k input bits on a branch, and an indeterminate L associated with the length of the input sequence .

For a binary input AWGN (additive white gaussian noise) channel with Q finite output quantizations, the probability of bit error with soft decision Viterbi decoding is upper bounded by

$$P_b < \frac{1}{k}\frac{\partial T(D,N,L)}{\partial N}\Big|_{D=D_0,N=1,L=1} \qquad 6.3$$

For a PSK signal and a binary input AWGN channel with no output quantization, D_0 is replaced by $\exp(-RE_b / N_0)$ where

R is the rate of the convolutional code and

E_b is the energy per bit and

N_0 is the noise power spectral density (noise power divided by bandwidth)

Table 6.1 lists coding gain (dB) for PSK with convolutional codes of rate R=1/2 with various constraint lengths using soft decision Viterbi decoding [21].

The FEC scheme discussed in this section is superior to the scheme discussed in the Section 6.6.1 because only two times the data needs to be sent in this scheme as compared to three times in the scheme discussed in Section 6.6.1. The coding gain in the convolutional coding makes the BER low at the receiver, as discussed in this section. This error correction scheme should be used for satellite data network analyzed in this project.

6.7 Procomm Plus and scripting capability

Procomm Plus software has been used in this study for establishing point to point data communication. Most of the source code is written in Procomm Plus software's ASPECT script language. Procomm plus software was written by Datastorm Technologies, Inc. It requires an IBM PC,

Table 6.1

Coding gain (dB) for PSK with convolutional codes of rate R=1/2 with various constraint lengths (K)

E_b / N_0 Uncoded (dB)	P_b	K=5	K=6	K=7
6.8	10^{-3}	3.3	3.5	3.8
9.6	10^{-5}	4.3	4.6	5.1
11.3	10^{-7}	4.9	5.3	5.8
-	Upper bound	5.4	6.0	7.0

XT, AT or PS/2 computer (or a computer that is compatible with any of these). It requires at least 192 K of available RAM. It works on PC-DOS or MS-DOS, version 2.0 or later. It can work with a very wide variety of Modems and requires a serial port if an external modem is being used.

Procomm Plus supports 13 standard protocols, and allows three more external protocols to be used with the software. Procomm software's ASPECT script language is a powerful programming language which supports conditionals like IF/THEN/ELSE/ etc. The ASPECT script language also supports advanced screen handling, file I/O, some elementary mathematical functions, looping, memory variables and subroutines. Procomm Plus emulates 16 popular video display terminals, which can be used to connect a PC to a mainframe and run full mainscreen applications. Procomm Plus operates at the following baud rates: 300, 1200, 2400, 3800, 9600, 19,200, 38,400, 57,600, 115,200. It can operate with space, even, odd, mark or no parity; 7 or 8 data bits; and 1 or 2 stop bits. It supports up to 8 user definable serial ports. The file transfer protocols it can support are as follows: Xmodem, Kermit, ASCII, Ymodem, Ymodem Batch, Modem7, Telink, Wxmodem, Sealink, Compuserve B, Ymodem-G, Xmodem-G batch, Imodem, and 3 external user-definable protocols. These file transfer protocols are explained in some detail in Section 2.6.

CHAPTER 7

Analysis of Networks

This chapter includes the link design and bit error rate (BER) calculations for satellite network, delays and data transfer capability for the leased telephone line network. Section 7.1 deals with the satellite network and Section 7.2 deals with the leased telephone line network.

7.1 Satellite Data Network

The TV signal from the main class room is transmitted to all the distant class sites via a C-band satellite with parameters similar to those of Galaxy V. These parameters are used for the analysis in this section. The basic equations of link design can be found in the Sections 5.1 and 5.2.

The TV signal has an RF bandwidth of 27 MHz. The modulation used is frequency modulation (FM). The uplink power of the Virginia Tech earth station is taken to be 100 W. This gives the power spectral density (PSD) in the signal, assuming uniform spectral density across the 27 MHz bandwidth, to be equal to

$$\frac{100}{27 \times 10^{6}} W / Hz = 3.7 \times 10^{-6} \text{ W/Hz} \qquad 7.1$$

If we use a 76.8 kHz RF bandwidth for a 19.2 kbps data channel (assuming BPSK modulation and R=1/2 convolutional coding with roll-off factor for ideal Nyquist filters equal to 1) and the PSD of the signal is the same as calculated in Equation (7.1), the power of the signal in the data channel is given by:

$$P = \frac{76.8 \times 10^3 \times 100}{27 \times 10^6} = 0.2844W \qquad 7.2$$

With the PSD's calculated in Equations (7.1) and (7.2), the following link budget analysis has been done for the data and TV link. We have assumed a 9 meter dish for the transmitting earth station and a 2.4 meter dish for the receiving earth station. Four 19.2 kbps data links are available on the uplink. The power that remains for the TV link after feeding power to four data links is 98.862 W.

7.1.1 Television Link

This section includes the link design for the TV link. The values wherever needed have been taken from [22] which gives parameters for the Galaxy V satellite. Most domestic C-band GEO satellites serving the U.S. have similar characteristics.

Slant Range (assumed)	40000 km
Uplink Frequency	6 GHz
Downlink Frequency	4 GHz
Clear air atmospheric loss	0.5 dB
Pointing error in receive antenna	1 dB
Polarization loss in receive antenna	0.5 dB

Uplink free space path loss	200 dB
Downlink free space path loss	196.5 dB
Boltzmann's constant	-228.6 dBW/K/Hz
$EIRP_{(sat)}$	38.5 dBW
(allowing 1dB backoff at the transponder output)	
$\left(\frac{G}{T}\right)_{sat}$	3.4 dB/K
Gain of the transmitting	
Earth station antenna	53.3 dB
Gain of the receiving earth station antenna	38.3 dB
System noise temperature of the receiving	
Earth station	100 K
$(EIRP)_{es(t)}$	73.3 dB
$\left(\frac{G}{T}\right)_{es(r)}$	18.3 dB/K

$es(t)$ and $es(r)$ stand for transmitting and receiving earth stations respectively.

Throughout the calculations in this chapter the aperture efficiency used for the antennas is taken to be 0.67.

Using the above values the following carrier to noise ratios have been calculated:

$\left(\frac{C}{N}\right)_u = 30.5$ dB and $\left(\frac{C}{N}\right)_d = 12.6$ dB.

$\left(\frac{C}{N}\right)_u$ and $\left(\frac{C}{N}\right)_d$ stand for $\left(\frac{C}{N}\right)$ ratios on the uplink and downlink respectively.

$$\left(\frac{C}{N}\right)_{link} = \frac{1}{\dfrac{1}{\left(\frac{C}{N}\right)_u} + \dfrac{1}{\left(\frac{C}{N}\right)_d}} \qquad 7.3$$

$\left(\frac{C}{N}\right)_{link}$ for the video link is equal to 12.5 dB. This is a good value for $\left(\frac{C}{N}\right)_{link}$ which for practical values should result in a $\left(\frac{S}{N}\right)$ for the video to be equal to 43 dB (assuming a de-emphasis improvement of 8 dB and a subjective improvement of 7 dB).

7.1.2 Data Link budget analysis

The power spectral density is assumed to be the same as in the TV signal. In this section the data link budget is analyzed and the degradation of the video link C/N is also calculated when power sharing takes place between the data and TV links. The following values have been used in the calculations for the data link:

$(EIRP)_{es(t)}$	47.8 dBW
Noise Bandwidth	19.2 KHz
$(EIRP)_{satellite}$	13 dBW

all the rest of the values are same as in Section 7.1.1

$\left(\frac{C}{N}\right)_{link}$ in this case is 12.5 dB. The $\left(\frac{C}{N}\right)_{link}$ will be further degraded by the interference from adjacent satellite links. The worst case in North America is a 2° separation between the satellites.Figure 7.1 illustrates the case. The equations below give relevant information about calculating the carrier-to-Interference ratio [21]

$$\left(\frac{C}{I}\right)_u \approx \left(\frac{EIRP}{EIRP'}\right)\left(\frac{G_u}{G_u'}\right) \qquad 7.4$$

here $EIRP'$ and G_u' are EIRP of interference signal in direction of interfered satellite and antenna gain

of interfered satellite in direction of interfering earth station respectively. The downlink equation has a similar form as Equation (7.4) above. For the above case the values calculated are:

$$\left(\frac{C}{I}\right)_u = 32.7dB$$
$$\left(\frac{C}{I}\right)_d = 29.2dB$$

The $\left(\frac{C}{N}\right)_{link}$ for the case where interference is also considered is given by [20]

$$\left(\frac{C}{N}\right)_{link} = \left(\frac{1}{\frac{1}{\left(\frac{C}{N}\right)_u}+\frac{1}{\left(\frac{C}{N}\right)_d}+\frac{1}{\left(\frac{C}{I}\right)_u}+\frac{1}{\left(\frac{C}{I}\right)_d}}\right) \qquad 7.5$$

Using the above equation and the values calculated above, the value of $\left(\frac{C}{N}\right)_{link}$ is equal to 12.3 dB for

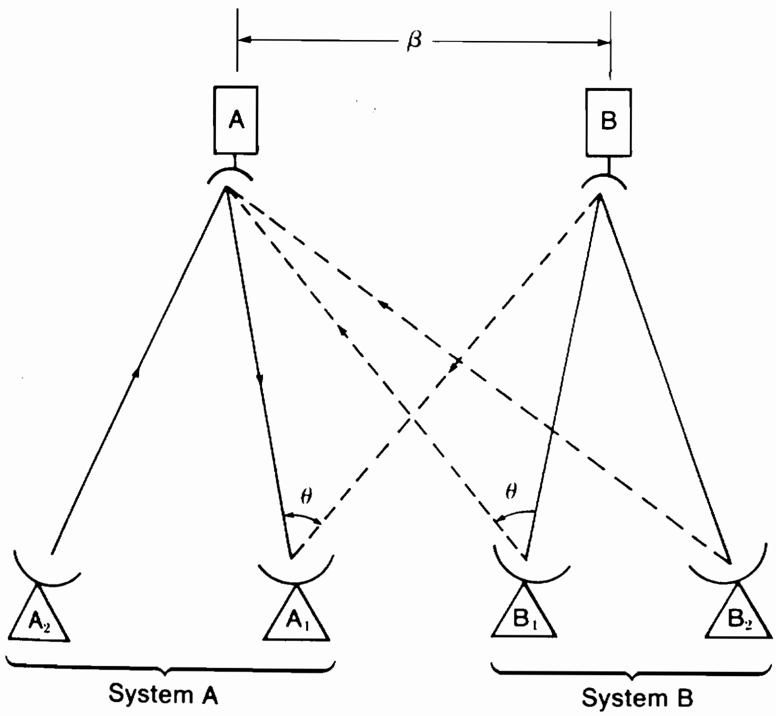


Figure 7.1 An illustration of interference from adjacent satellite. (From Ha., T.T.,"Digital Satellite Communications", 2nd. ed., John Wiley and Sons, New York, 1989).

the data channel.

7.1.3 Reduction in C/N Ratio for TV Link

In this section reduction in the $\left(\frac{C}{N}\right)_{link}$ for TV signal is calculated. The power in the TV signal is reduced because power is shared among the data and TV signals. The PSD of all the signals is assumed equal. Four data channels are assumed each having RF bandwidth of 76.8 kHz.

The power in the TV signal is calculated as follows:

$$\mathrm{P} = 100 \times \left(\frac{27 \times 10^6 - 4 \times 76.8 \times 10^3}{27 \times 10^6} \right) = 98.8\mathrm{W} = 19.9\mathrm{dBW}$$

One more parameter that is changed due to power sharing is $(EIRP)_{satellite}$ allocated to the TV signal. The new value for the TV signal is given by:

$$7079.45 \times \left(\frac{27 \times 10^6 - 4 \times 76.8 \times 10^3}{27 \times 10^6} \right) = 11092.524 = 38.4\mathrm{dBW}$$

All the rest of the parameters remain the same. Hence the new values for the $\left(\frac{C}{N}\right)$ ratios are

$\left(\frac{C}{N}\right)_u$ = 30.4 dB , $\left(\frac{C}{N}\right)_d$ = 12.5 dB and $\left(\frac{C}{N}\right)_{link}$ = 12.3 dB.

The $\frac{C}{N}$ for the data link at the receiver is 12.3 dB which gives a BER, using Equation (5.15), equal to

6E-9. This figure of BER is very small for transmitting text files which indicates a fairly good data link.

The TV signal because of its very high bandwidth compared with the bandwidth for the data signals is not appreciably affected by power sharing with the data channels. There is only a difference of 0.2 dB in the value of $\left(\frac{C}{N}\right)_{link}$ for the TV signal. If the BER needs to be improved to 1E-6 (for the case of QPSK modulation) the power of the data channels can be increased slightly to compensate. The S/N for TV will not fall much due to more power sharing. One dB more in C/N for data link will produce desired results.

The results indicate that the 19.2 kbps data channel is a very good medium for transmitting text files to the distant sites. There is no appreciable reduction in the TV signal $\left(\frac{C}{N}\right)_{link}$ ratio due to power sharing with data channels. We also calculate here the probability of packet error given BER equal to 1E-6. Using Equation (6.1) we get probability of packet error to be equal to 5.743 E(-3). Which indicates that one in 200 packets (720 bytes each) will be received with error and one in 10000 will be accepted as a good packet, with no rain fade margin in the link. The use of convolutional coding as explained in Section 6.6.2 results in much better results, a BER of 10^{-7} can be achieved with a $\left(\frac{C}{N}\right)_{link}$ equal to 6.4 dB (as shown in Table 6.1).

7.2 Leased Telephone Line Network

The leased telephone line network for the Virginia Tech distance education program has been outlined in Section 4.1. Figure 4.1 depicts the network architecture. This section calculates the various delays which occur in the present and proposed network. The theory has been outlined in Sections 3.1

and 3.2 and Chapter 4.

7.2.1 Average Time Between Two Voice Connections

The first calculation deals with the average delay between the two concurrent voice connections in the proposed protocol. This involves the calculation of average walk time as defined in the Chapter 4 and average time required to transmit a typical file over the network. The time required by the primary station to transmit the polling message, for the polling message to propagate on the line to the secondary station, for the secondary station to recognize its address and then to take action on the request and the time required by the message to go back to the primary station is the total time taken by the transfer of permission from primary station to secondary station. This time for polling from one station to another sequentially for one complete cycle is called the total walk time of the system. Figure 7.2 (a) shows the sequence of tasks being done by the network. As can be seen from the Figure 7.2(a) the average time between t_1 and t_2 is given by

$$\sum_{i=1}^{N} w_i + \overline{f} \qquad 7.6$$

where w_i is the walk time for the ith station in a network of N stations and $\overline{f}$ is the file transfer time for an average size file. The total walk time is given by

$$L = Nt_p + Nt_s + \tau' \qquad 7.7$$

where t_p is the time to transmit the polling message, t_s the time taken by the station to synchronize and τ' is the propagation delay which depends on the architecture of the network. We assume here that the

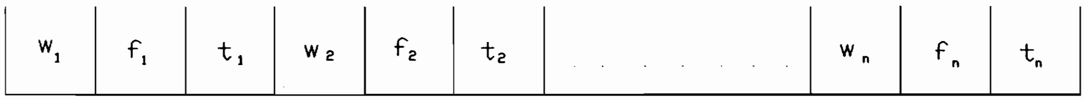


Figure 7.2(a)

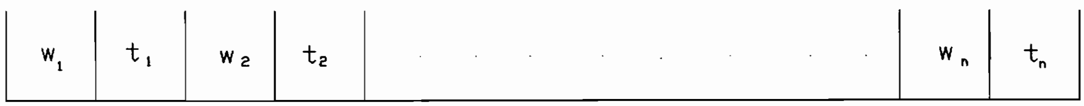


Figure 7.2(b)

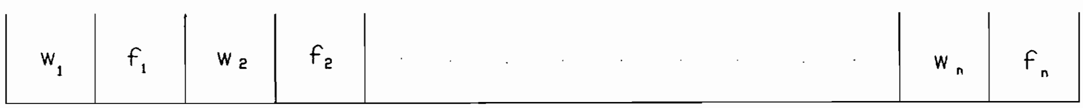


Figure 7.2(c)

Depiction of processes in the terrestrial network. The diagrams show sequence of events taking place in the terrestrial network.

w_i = walk time.

f_i = file transfer time.

t_i = time duration of a voice connection.

time taken to synchronize by each station on the network is the same. To calculate the propagation delay we assume that there are 30 distant class sites, 15 of which are located around the Tidewater area in Virginia and the rest are located in Northern Virginia area. The speed of propagation of energy in the medium of the network is assumed to be 50 miles/msec [18]. Hence τ' in this case is equal to 194.4 msec. The synchronization time t_s is assumed to be 10 msec hence Nt_s is equal to 300 msec. The polling message in the proposed protocol is 128 bits long as explained in Section 6.3. We assume here the line bit rate is 2400 bits per second. Hence Nt_p is equal to 1600 msec. Hence total walk time L is given by

$$L = 194.4 + 300 + 1600 = 2114.4 m\sec$$

average walk time for one station is then $\frac{L}{30}$.

$$\overline{w} = \frac{L}{30} = \frac{2114}{30} = 70.4 m\sec \approx 70 m\sec$$

To calculate $\overline{f}$ we assume that the average frame length statistics are the same at each station and the frame lengths are exponentially distributed with a mean of 10kbytes (which is equivalent to about 4 type- written pages). On a 2400 bps line it will take 80000/2400 =33.333 seconds to transmit such a file. Hence the average time elapsed between the two voice connections is equal to 33.3333 sec + 70 msec = 33.403 seconds.

7.2.2 Polling in the Voice Network

If the terrestrial network is used solely for voice connection and the voice connections are given to the different stations through polling, the expected delay will be appreciably higher than the above value calculated in 7.2.1. This section calculates the average delay for the above case. Figure 7.2(b) shows the sequence of events taking place in this case. The expected delay is given by Equation (4.6) mentioned here again for convenience:

$$E(D) = \frac{\overline{t_c}}{2}\left(1 - \frac{\rho}{N}\right) + \frac{N\lambda\overline{m^2}}{2(1-\rho)} \qquad 7.8$$

We assume here that the durations for voice connections are exponentially distributed with mean 180 seconds. Other values in Equation (7.8) are N=30 $\overline{t_c}$=2.114 sec as calculated in Section 7.2.1. We have $\rho = N\lambda\overline{m}$, where λ is the request rate for voice connection, and $\overline{m}$ is the average duration of one voice connection. We assume here that the average request for voice connection is the same for each distant class site which is equal to 0.5 in one hour. Hence we have the following values for the parameters in Equation (7.8):

$$\lambda = \frac{0.5}{3600} / \sec^2$$

$$\rho = 0.5$$

The expected delay E(D) is then given by Equation (7.8). E(D) for this typical case is equal to 271.039 seconds or approximately 4 minutes and 31 seconds. The expected delay for this case is very high and will not be acceptable to the students at the distant class sites. But this value is helpful is comparing this scheme with others examined in this project.

In the present leased telephone line network which allows voice connection without any polling strategy the voice connection is available for 100% of the time the telephone lines remain connected. For the polling case discussed in this section the voice connection is not available when one of the stations is being polled. One voice session takes $\overline{w} + \overline{t}$ sec to complete. The value of $\overline{w} + \overline{t}$ is equal to $\frac{2.114}{30} + 180\,\text{sec}$ = 180.071 sec. This indicates that the network is being utilized for voice connection 99.96% of the time. The efficiency of the polling scheme is very good but the expected delay is very high.

7.2.3 Use of Terrestrial Network for Data Transfer only

This section examines the use of the terrestrial network for the data transfer from the distant class sites to Blacksburg, when there is no provision for the voice connection through the same network. Figure 7.2(c) shows the sequence of events in this case. The calculations for the $\overline{w} + \overline{f}$ is the same for the case in Section 7.2.1. The average file 10 kbytes long will take 33.403 seconds to transmit. For a worst case of 75 off-campus students it will take 2505.225 sec to transfer all the data to Blacksburg which is approximately equal to 41 minutes and 45 seconds. Here it is assumed the file size remains the same as in Section 7.2.1. For a one hour period of data transfer this network can transfer about 5.5 pages of text from each student for the worst case of 75 off-campus students.

7.2.4 Prioritized Data and Voice Connection

There is one more data transfer scheme that needs to be discussed here, which improves the delay in getting the voice connection, in the distant class sites. In this scheme we can give priority to voice connection all the time. When there is a voice connection request the data transfer stops immediately and the network becomes available for voice connection. When the conversation is over the instructor can revert the network back to data transfer mode. Voice can be preceded by high level burst of tone (e.g. at 3 KHz) that kills all the data bits on the line. When the network controller recognizes the tone, it blocks the data transfer and makes the network available for the voice connection. Delay here will be of the order of 0.5 second: the time it will take the tone to travel to the network controller and the time it will take network controller to recognize the tone and take action on it. This network is suitable for voice connection but becomes deficient for data transfer. After every voice request the data transfer has to be stopped and for every transfer from data to voice some packets will have to be retransmitted. Reversion from voice to data transfer will require some synchronization between network controller and the secondary station that is sending data to Blacksburg. The next section proposes a superior solution to the above problems.

7.3 A VSAT Network for Distance Education

The data transfer strategies that have been proposed in Chapter 6 and analyzed in Chapter 7 are deficient in one or other respect. The terrestrial network doesn't work well when both data and voice are transmitted on the same network. Data communication blocks the voice connection and introduces unacceptable and unpleasant delays in voice connection. The terrestrial network is not a very appropriate network for transferring data from distant sites to Blacksburg because of its low capacity. Even then it can be used satisfactorily for transferring data to Blacksburg from the distant class sites if time available for data transfer is not a consideration. The data rates achievable over the network with fairly low bit error rates are not higher than 9600 bps. The occurrence of burst errors will be very high because the interference and noise level in the network is high.

This section attempts to propose an alternative network primarily to be used for data transfer from the distant class sites to Blacksburg. A VSAT network has been proposed with the hub location in Blacksburg and one VSAT terminal for each distant class site. VSATs can be defined as a class of very small aperture, intelligent satellite earth stations suitable for easy on-premise installation, usually operating in conjunction with a large hub earth station and capable of supporting a wide range of two-way, integrated telecommunication and information services [23].

The proposed VSAT network could also be used to conduct interactive video communication between the instructor and students in distant classes during instructor's office hours. This will entail the use of the network beyond class time and thus the VSAT network could be more efficiently used. The parameters in the VSAT network are proposed keeping in consideration that the network will one day be used for video interactive communication in addition to data transfer from distant class sites to Blacksburg. The basic requirements for such a network are:

1. A high data rate which can sustain real time video communication

2. Communication capability between all the class sites and main campus in Blacksburg. There is no need for communication between two class sites directly

3. Data transfer capability from distant class sites to Blacksburg

4. Different communication needs will need different protocols hence a need for the capability at VSAT to use different communication protocols adaptively.

In the United States VSAT networks usually use 14/12-GHz band operation because of the absence of terrestrial and adjacent satellite interference in this band. The employment of the 6/4 -GHz band requires the use of spread spectrum techniques to reduce the signal's power spectral density. This technique does not utilize the satellite capacity very efficiently. It is proposed that the 14/12-GHz band be used for the network. Rain attenuation affects the performance in this band considerably, hence a link margin should be provided for rain attenuation.

Network configuration is another important design feature in VSAT networks. As already outlined there is no need for a communication capability between two distant class sites. Hence a VSAT star network as shown in Figure 7.3 should be used. The Hub station will be located in Blacksburg and one VSAT will be located in each distant class site. Low cost of the VSAT is possible due to small dish antennas and low power transmitters in the terminals. Quaternary phase shift keying (QPSK) should be used as the modulation technique. The communication entails stream traffic as opposed to bursty traffic because the network will be used for data transfer in bulk form, from distant class sites and for two way video communication. For this type of traffic the multiple access schemes that are most popular are time division multiplexing (TDM) and single channel per carrier (SCPC). The TDM technique includes communication between different pairs separated by time and SCPC is the technique where each VSAT communicates with the hub station, inbound or outbound (inbound is VSAT to hub and outbound is hub to VSAT communication), by using one preassigned channel with one carrier. A good discussion of multiplexing and multiaccess schemes for VSAT networks (and for satellite networks in general) is given in [21, 23, 24]. Two carriers may be necessary for outbound and/or inbound SCPC (depends on the error control scheme we use:ARQ or FEC) because at one time two communication sessions may be going on: data transfer from distant class sites to Blacksburg and interactive video communication between course instructor and student at the distant class site. Carrier configuration is an important component of the VSAT network design. The proposed network will primarily require about 50 VSAT stations, mostly located in the state of Virginia. The requirement for the network suggests that we should use SCPC for the inbound and either SCPC or TDM-FDMA for the outbound. These receivers will be cheaper than receivers for the DAMA (demand assigned multiple access) networks. DAMA networks use the satellite capacity more efficiently but make the network components more expensive. Two carriers may be used per VSAT one each for data and video, for the inbound.

VSATs with antenna diameter of 1.8 meters and transmitted power of as low as 0.35 W and a hub with an antenna diameter of 7.0 meters will suffice for our purposes. The analysis in [21, 23]

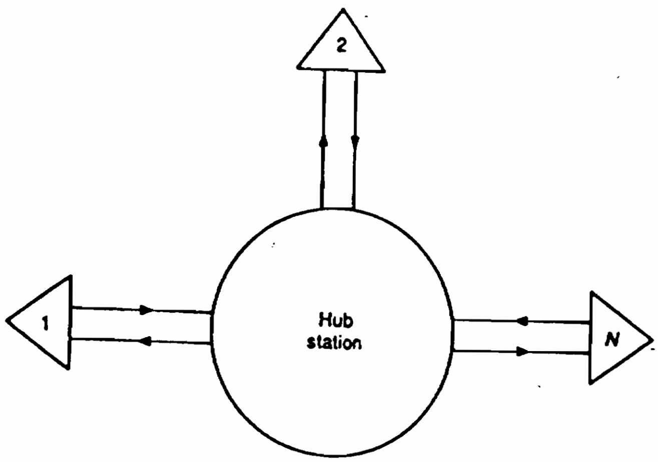


Figure 7.3 A VSAT network in star configuration. (From Ha., T.T.,"*Digital Satellite Communications*", 2nd. ed., John Wiley and Sons, New York, 1989).

suggests that with the above values and practical values for satellite parameters as shown in Table 8.1, the data rate that can be sustained by inbound and outbound links is 5.75 Mbps each, for a BER of 1E-6. This indicates the network capacity of 11.5 Mbps or 204 SCPC 56 kbps carriers. Table 8.1 lists the values for satellite and earth station parameters for a viable VSAT network for distance education [21].

One more aspect of the network design that needs to be discussed is the error control in the network. Forward error correction (FEC) or automatic repeat request (ARQ) techniques could be used. ARQ requires a return path and adds unpleasant delays in the VSAT network when it is being used for query/response type of communication (as in instructor's office hours over the network). An FEC codec on the other hand constitutes a large part of the cost of a VSAT receiver. FEC could be used for the video interactive communication and ARQ scheme could be used for the data transmission.

Finally we state the expected cost per VSAT per month for this type of network . Cost per month per VSAT location is given by [25]

$$(X+H\times M/N)\times(\text{Capitalization factor})+(S_1+S_2\times K)/N \quad \text{where} \qquad 7.9$$

- X: installed cost per VSAT
- N: number of VSATs
- H: installed cost per hub
- M: number of hubs necessary to support N VSATs
- Capitalization factor based on five year amortization factor at 10% rate of return=0.212
- S_1: transponder cost per month for inbound carrier (one carrier of adequate bandwidth)
- S_2: transponder cost per month for outbound carrier
- K: Number of outbound carriers necessary to support N VSATs

Installed cost per VSAT can be taken as $10000 for a 1.8 meter VSAT. Hub installed cost is about $ 50000 for the proposed VSAT network. These characteristic values will result in cost per month per VSAT for the proposed network to be about $ 750. This cost analysis assumes an inbound carrier of 224 kbps and an outbound carrier size of 1.544 Mbps.

Table 8.1

Satellite and earth station parameters for a viable VSAT network for distance education [21]

Satellite	
saturation flux density (dBW/m^2)	-85.0
saturation EIRP (dBW / m^2)	43.0
output backoff (dB)	5.0
G/T(dB/K)	2.5
Hub station	
antenna diameter (m)	7.0
gain at 14 GHz (dB)	58.0
gain at 12 GHz (dB)	56.7
transmitter power (dBW)	7.6
G/T (dB/K)	30.0
Remote VSAT	
antenna diameter(m)	1.8
gain at 14 GHz (dB)	46.5
gain at 12 GHz (dB)	45.0
total transmitter power for all VSATs (dBW)	15.5
(star network)	
G/T (dB/K)	21.0

CHAPTER 8

CONCLUSIONS

The existing network for distance education for Virginia Tech needs improvements because:

- it has no capability to transfer data from distant class sites to Blacksburg
- it has no capability to transfer data from Blacksburg to distant class sites
- the terrestrial leased line network is not efficiently used
- the subcarrier available on the uplink of the satellite network, which can sustain a 19.2 kbps data rate is not used at all.

This project has proposed a protocol that will utilize the leased telephone line network more efficiently. The protocol utilizes a roll call polling method to control the flow of information on the network. The polling method, as expected, adds unacceptable and unpleasant delays in the leased line network. When both data and voice utilize the same leased line network the average time between two voice connections is close to 30 seconds. This delay is unacceptable for a distance education network. The leased line network is also not an appropriate network to utilize for data transfer between distant class sites and Blacksburg because of its low capacity. The data rates achievable over the network with fairly low bit error rates are not higher than 9600 bps. When terrestrial network is used only for data transfer, approximately six pages of text from each student (assuming 75 off-campus students) can be transferred to Blacksburg from the distant sites in a one hour period. The data throughput for the network is not very high but it could be used if the network of leased lines could be leased for a longer time. If the network is available all the time it could be used for voice connection during a class and for data transfer the rest of the time.

The subcarrier available on the uplink of the satellite which is used for transmitting the TV signal to distant class sites, can sustain four 19.2 kbps data streams. This capability should be utilized to transfer data from Blacksburg to all the distant class sites. This data can include exam papers, class notes and graded homeworks and exams. A protocol has been proposed in this thesis that controls the data transfer through the satellite uplink. A link design for this case has also been presented. It has been shown that the power sharing between the subcarriers and TV signal does not reduce the S/N for the TV signal at the receivers appreciably. This is a very good medium to transfer data from Blacksburg to all the distant class sites.

A VSAT network has been proposed to fulfill all the needs for the distance education network for Virginia Tech. This network could easily be used as a distance learning network for the state of Virginia. The network consists of one VSAT station at every distant class site and a hub station in Blacksburg. The capabilities include two way data carriers between all the distant class sites and Blacksburg. Video communication between each of the distant class sites and Blacksburg has also been discussed. Video communication allows instructors to conduct office hours with distant site students.

Areas of future work include:

- A collision detection scheme of data transfer and voice communication using the terrestrial telephone lines should be studied. Viability of such a scheme should be examined
- Cost analysis of all the proposed data networks should be done to conclude which scheme will be most economic.

REFERENCES

1. Portway, P.S., "How corporate America trains by telecommunications", *Communications News*, Vol 30, Iss 2 February 1993.
2. Elisabeth Charron and Kim Obbink, "Long-distance learning", *The Science Teacher*, vol 60, n3, March 1993.
3. Ohler Jason, "Why Distance Education ?", *The Annals of the American Academy of Political and Social Science*, v514, March 1991.
4. Bertsekas D. and Gallagher R., "*Data Networks*", Prentice Hall Inc., Englewood Cliffs, NewJersey, 1990.
5. Couch II L.W., "*Digital and Analog Communication Systems*", Macmillan Publishing Company New York ,1989.
6. Lathi B.P., "*Modern Digital and Analog Communication Systems*", Holt, Rinehart and Winston, New York, 1983 (2nd ed., 1989).
7. Taub, H., and D. L. Schilling, "*Principles of Communication Systems*", 2nd ed., McGraw-Hill Book Company, New York, 1986.
8. Nyquist, H., "Certain Topics in Telegraph Transmission Theory," *Transactions of the AIEE*, vol. 47, February 1928, pp. 617-644.
9. Jerri, A.J., "The Shannon's Sampling Theorem-Its Various Extensions and Applications: A Tutorial Review," *Proceedings of the IEEE*, vol. 65, no.11, November 1977, pp. 1565-1596.
10. Ziemer, R.E., and Peterson, R.L., "*Introduction to Digital Communications*", Macmillan Publishing

Company, New York, 1992.

11. Forney, Jr., David and M. Vedat Eyuboglu, "Combined Equalization and Coding Using Precoding",*IEEE Communications Magazine*, Vol 29, No. 12, pp 25-34, December 1991.
12. Shannon, C.E., "Communication in the Presence of Noise," *Proceedings of the IRE*, vol 37, January 1949, pp. 10-21.
13. Tanenbaum, A.S., "*Computer Networks*", 2nd. ed., Englewood Cliffs, N.J., Prentice-Hall 1988.
14. Bhargava, V.K., "Forward Error Correction Schemes for Digital Communications,"*IEEE Communications Magazine*, vol 21, January 1983, pp. 11-19.
15. Lin, D.J. Costello, Jr., M.J. Miller, "Automatic-Repeat-Request Error-Control Schemes, "*IEEE Communications Magazine*, 20, 5-7, December 1984.
16. Sherman, K., "*Data Communications :A user's Guide*", 2nd ed., Reston Publishing Company, Inc. Reston, Virginia.
17. Stallings, W., "*Data and Computer Communications*", Macmillan Publishing Company, 1985.
18. Schwartz, M. "*Telecommunications Networks: Protocols, Modeling and Analysis*",. Addison Wesley Publishing Company, 1987.
19. Konheim, A.G. and B. Meister, "Waiting Lines and Times in a System with Polling," *J. ACM*, vol.21, no.3, July 1974, 470-490.
20. Pratt, T. and C.W. Bostian, "*Satellite Communications*", John Wiley and Sons, New York,1986.
21. Ha, T.T., "*Digital Satellite Communications*", 2nd ed., John Wiley and Sons, New York, 1989.
22. Long, M., "*The 1993/1994 World Satellite Annual*", MLE Inc.
23. *IEEE Communications Magazine*, "Special Series on VSAT Communications Networks:Technology and Application", May, July, September and November 1988 and May 1989.
24. Puente, J.G., William G. Schmidt, and Andrew M. Werth, "Multiple-Access Techniques for Commercial Satellites", *Proceedings of the IEEE*, 59,218-229, February 1971.
25. Sharma, R., "VSAT Network Economics: A Comparative Analysis", *IEEE Communications*

Magazine, vol 27, pp.31-35, February 1989.

VITA

Shikhar Kishore Srivastava was born on June 20, 1968 in New Delhi, India. He attended Kendriya Vidyalaya Tagore Garden in New Delhi, India from April 1977 to March 1986. He received B.E. in Electrical Engineering from Delhi College of Engineering, Delhi University, India, in July of 1990. He attended Virginia Polytechnic Institute and State University between August 1991 and May 1994. He had an opportunity to work as an Intern in Entergy Corporation in New Orleans, Louisiana in the summer and fall of 1992. He received M.S. in Electrical Engineering in May of 1994. He has accepted employment with TSI in Arlington, Virginia.

Shikhar Srivastava