TECHNICAL REPORT STANDARD TITLE PAGE

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
TX-92/1245-2		

4. Title and Subtitle	5. Report Date
Communications in Intelligent Vehicle Highway Systems--Part I	November 1991
	6. Performing Organization Code

7. Author(s)	8. Performing Organization Report No.
Costas N. Georghiades, Sittiporn Patarasen, Emina Soljanin, Hadi Jardak, Wesley M. Wills	Research Report 1245-2

9. Performing Organization Name and Address	10. Work Unit No.
Texas Transportation Institute The Texas A&M University System College Station, Texas 77843	11. Contract or Grant No. Study No. 2-1-90/1-1245

12. Sponsoring Agency Name and Address	13. Type of Report and Period Covered
Texas Department of Transportation P.O. Box 5051 Austin, Texas 78701	Interim – January 1990 thru August 1991
	14. Sponsoring Agency Code

15. Supplementary Notes

Research performed with all State Funds
Research Study Title: Texas Advanced Transportation Technology Project

16. Abstract

This report accesses the communications capacity of the traffic management systems being installed in Texas with a view to two-way communications between the traffic management centers and vehicles on the roadway. The study will investigate the general telecommunications requirements for information flow, both from and to vehicles and traffic centers. Various approaches to the communication problem have been studied in terms of data rates and bandwidth requirements, feasibility of implementation and the ability to incorporated future expansions. These approaches include the use of infra-red optical communications in a distributed environment, convential spread sprectrum radio, and the use of a dedicated frequency-band. This report only investigates those communication systems which offer the the most promising methods for communicating between vehicles and a traffic management center. It does not consider, for example, the distribution using a fiber-optic network installed under major highways. The capacity requirements of such a network are not covered in this report. The emphasis is on digital systems that offer unique advantages over analog systems, including the ability to integrate with existing data networks such as the telephone network.

17. Key Words	18. Distribution Statement
Telecommunications, Spread Spectrum, Vehicle Communications	No restrictions. This document is available to the public through the National Technical Information Service, 5285 Port Royal Rd., Springfield, VA 22161

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	97	

Form DOT F 1700.7 (8-69)

COMMUNICATIONS IN INTELLIGENT VEHICLE HIGHWAY SYSTEMS– PART I

by

Costas N. Georghiades
Assistant Research Specialist

and
the following Research Assistants

Sittiporn Patarasen

Emina Soljanin

Hadi Jardak

Wesley M. Wills

Texas Transportation Institute

Research Report 1245-2

Research Study Number 2-1-90/1-1245
Texas Advanced Transportation Technology Project

Sponsored by

Texas Department of Transportation

Texas Transportation Institute
The Texas A&M University System
College Station, Texas 77843

November 1991

METRIC (SI*) CONVERSION FACTORS

APPROXIMATE CONVERSIONS TO SI UNITS

LENGTH

Symbol	When You Know	Multiply By	To Find	Symbol
in	inches	2.54	centimetres	cm
ft	feet	0.3048	metres	m
yd	yards	0.914	metres	m
mi	miles	1.61	kilometres	km

AREA

Symbol	When You Know	Multiply By	To Find	Symbol
in^2	square inches	645.2	centimetres squared	cm^2
ft^2	square feet	0.0929	metres squared	m^2
yd^2	square yards	0.836	metres squared	m^2
mi^2	square miles	2.59	kilometres squared	km^2
ac	acres	0.395	hectares	ha

MASS (weight)

Symbol	When You Know	Multiply By	To Find	Symbol
oz	ounces	28.35	grams	g
lb	pounds	0.454	kilograms	kg
T	short tons (2000 lb)	0.907	megagrams	Mg

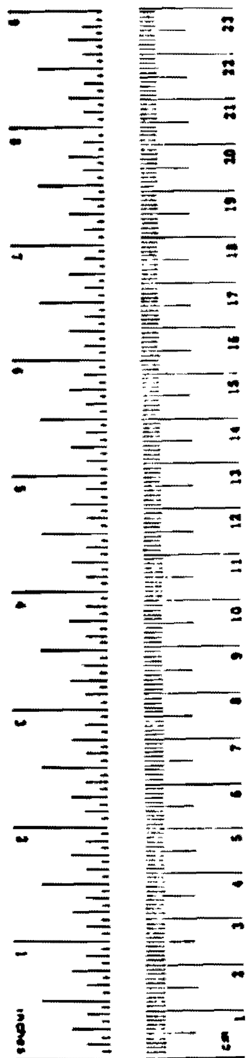
VOLUME

Symbol	When You Know	Multiply By	To Find	Symbol
fl oz	fluid ounces	29.57	millilitres	mL
gal	gallons	3.785	litres	L
ft^3	cubic feet	0.0328	metres cubed	m^3
yd^3	cubic yards	0.0765	metres cubed	m^3

NOTE: Volumes greater than 1000 L shall be shown in m^3.

TEMPERATURE (exact)

Symbol	When You Know	Multiply By	To Find	Symbol
°F	Fahrenheit temperature	5/9 (after subtracting 32)	Celsius temperature	°C



APPROXIMATE CONVERSIONS TO SI UNITS

LENGTH

Symbol	When You Know	Multiply By	To Find	Symbol
mm	millimetres	0.039	inches	in
m	metres	3.28	feet	ft
m	metres	1.09	yards	yd
km	kilometres	0.621	miles	mi

AREA

Symbol	When You Know	Multiply By	To Find	Symbol
mm^2	millimetres squared	0.0016	square inches	in^2
m^2	metres squared	10.764	square feet	ft^2
km^2	kilometres squared	0.39	square miles	mi^2
ha	hectores (10 000 m^2)	2.53	acres	ac

MASS (weight)

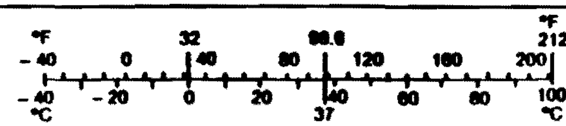
Symbol	When You Know	Multiply By	To Find	Symbol
g	grams	0.0353	ounces	oz
kg	kilograms	2.205	pounds	lb
Mg	megagrams (1 000 kg)	1.103	short tons	T

VOLUME

Symbol	When You Know	Multiply By	To Find	Symbol
mL	millilitres	0.034	fluid ounces	fl oz
L	litres	0.264	gallons	gal
m^3	metres cubed	35.315	cubic feet	ft^3
m^3	metres cubed	1.308	cubic yards	yd^3

TEMPERATURE (exact)

Symbol	When You Know	Multiply By	To Find	Symbol
°C	Celsius temperature	9/5 (then add 32)	Fahrenheit temperature	°F



These factors conform to the requirement of FHWA Order 5190.1A.

* SI is the symbol for the International System of Measurements

IMPLEMENTATION STATEMENT

The Texas Department of Transportation is investigating methods of providing increased mobility for the major urban areas of Texas. One area being examined is to install Traffic Management Centers that would control the signal systems on freeway ramps as well as frontage roads and other major arterial in the vicinity of the freeway. This report investigates strategies of communicating with vehicles that are using the freeway and urban street system. While the findings of this report may not be installed as a part of the initial traffic management centers they should provide a context for planning the centers so as to be compatible with the most recent state-of-the-art communications techniques.

DISCLAIMER

ABSTRACT

This report accesses the communications capacity of the traffic management systems being installed in Texas with a view to two-way communications between the traffic management centers and vehicles on the roadway. The study will investigate the general telecommunications requirements for information flow, both from and to vehicles and traffic centers. Various approaches to the communication problem have been studied in terms of data rates and bandwidth requirements, feasibility of implementation and the ability to incorporated future expansions. These approaches include the use of infra-red optical communications in a distributed environment, conventional spread spectrum radio, and the use of a dedicated frequency-band. This report only investigates those communication systems which offer promising alternatives for communicating between vehicles and a traffic management center. It does not consider the distribution using a fiber-optic network installed under major highways. The capacity requirements of such a network are not covered in this report. The emphasis is on digital systems that offer some unique advantages over analog systems, including the ability to integrate with existing data networks such as the telephone network.

TABLE OF CONTENTS

CHAPTER | Page

CHAPTER Page

LIST OF TABLES

LIST OF FIGURES

FIGURE Page

OVERVIEW OF REPORT

This report summarizes research performed between January 15, 1990, and August 31, 1990, for the project "Telecommunications Requirements for Intelligent Vehicle Highway Systems," funded by the Texas Transportation Institute (TTI).

The research was done at Texas A&M University by the following Electrical Engineering Department graduate students, under the direction of C.N. Georghiades acting as the principal investigator: Ning Kong (Ph.D.) (who joined the project during the last three months), Sittiporn Patarasen (Ph.D.), Emina Soljanin (Ph.D.), Hadi Jardak (MS) and Wesley Wills (MS). The main thrust of the research concentrated on four areas briefly described next and in more detail in the main body of the report that follows. Further work on these topics is currently in progress.

- **Spread-spectrum road-automobile communication**: This part of the project was researched by Emina Soljanin and Wesley Wills and consists of a study of the communication problem between roadside beacons and moving vehicles, and vice versa, utilizing spread-spectrum techniques. The power, bandwidth and complexity requirements in order to achieve a desired error-probability performance and communicate a given amount of information between a beacon and moving vehicles was studied, under some worst-case traffic scenario.

- **Infrared road-automobile communication:** This second part of the report describes work done by Sittiporn Patarasen on an infrared beacon-to-automobile communication system. Here, the various noise sources that affect communication were studied and quantified, and the performance of the system in terms of power requirements, error-probability and throughput was determined.

- **Automobile positioning**: This is a tutorial part done by S. Patarasen that describes the various approaches currently available for obtaining the position of moving vehicles, their advantages and disadvantages, and some recommendations.

- **Optical fiber communications**: This final part of the report, done by Hadi Jardak, studies the expected performance of a fiber-optic network to be installed under highways which will distribute information to and from roadside beacons. The advantages and disadvantages of fiber-optic systems over coaxial cable systems are studied, as well as both direct-detection and heterodyne optical systems. For direct-detection systems, a Fortran program is included that determines the throughput of such systems, given various practical parameters.

CHAPTER I

SPREAD-SPECTURM ROAD-AUTOMOBILE COMMUNICATIONS

A. Introduction

Communications between roadside beacons and the vehicles on the highway may be provided by a radio communication system. In a setting where each beacon serves only its limited area, but is able to communicate with all the vehicles within that area, a multiple user radio communication system is required. There are three major techniques used in multiple user radio communication systems: time division multiple access (TDMA), frequency division multiple access (FDMA), and spread spectrum multiple access (SSMA).

In time-division multiple access radio communication systems, time available for communications is divided into small segments called slots. A certain number of these slots are then grouped into frames. A particular user is assigned to one slot and communicates within each frame only during the assigned slot. As long as no two users are assigned the same slot, there is no interference; this is the main advantage of the TDMA systems. Fukui, Noki, and Hashizume discuss a road automobile communication system which uses a TDMA scheme [1] with the frames containing 20 slots. The system consists of a series of beacons separated by between two and five kilometers. Each of the beacons covers an area of approximately 100 meters in diameter. The main disadvantage of this system is the inefficiency of time usage: each user is allotted only 1/20th of the available time. In applications where the available time is limited, this can be a major obstacle. Since the allotted time is reduced, the data rate must be increased which either reduces performance or requires more power.

Frequency-division multiple access communications systems (analog or digital), for transmission of the data, use an auxiliary signal called carrier. The amplitude (AM systems), phase, or frequency (FM systems) of the carrier signal is modulated (modified) according to the information signal to be sent. Thus, even though the information signals from different systems share the same frequency band, the modulated AM or FM signals are separated in frequency. This frequency separation makes it possible to extract a particular signal from the other signals transmitted at the same time. A particular receiver selects a specific signal by using a filter tuned to the appropriate frequency. All other signals are thus eliminated.

Spread spectrum (SS) radio communications systems use, besides the carrier, a second auxiliary signal in the modulation process. This signal is a user-specific digital spreading code. The spreading codes are binary sequences which, when long, seem to be random. The power spectrum of the information signal to be sent is spread over a wide frequency band according to its assigned spreading code. While in AM or FM systems signals are differentiated based on their carrier frequency, in SS systems signals are differentiated based on their spreading code — the carrier frequency of all signals is the same. A particular receiver selects a specific signal by despreading its power spectrum — again according to its assigned spreading code — while leaving the power of other signals spread over the wide frequency band. Although not completely eliminated, all other signals thus produce an effect of low level interference similar to noise. The spreading codes are pseudo-random. This makes spread spectrum transmission private: a casual listener would not be able to intercept messages. Spread spectrum modulation provides low susceptibility to interference due to multipath propagation or due to multiple users on a single channel, and promises more efficient utilization of available frequency spectra.

The block diagram shown in Fig. 1 illustrates the basic elements of a spread spectrum digital communications system. In addition to the channel encoder and decoder and modulator and demodulator, there are two identical pseudo-random pattern generators, one interfacing with the modulator at the transmitting end and the other interfacing with the demodulator at the receiving end. The generators generate a pseudo-random or pseudo-noise (PN) binary sequence (spreading code) which is impressed on the transmitted signal at the modulator and removed from the received signal at the demodulator.

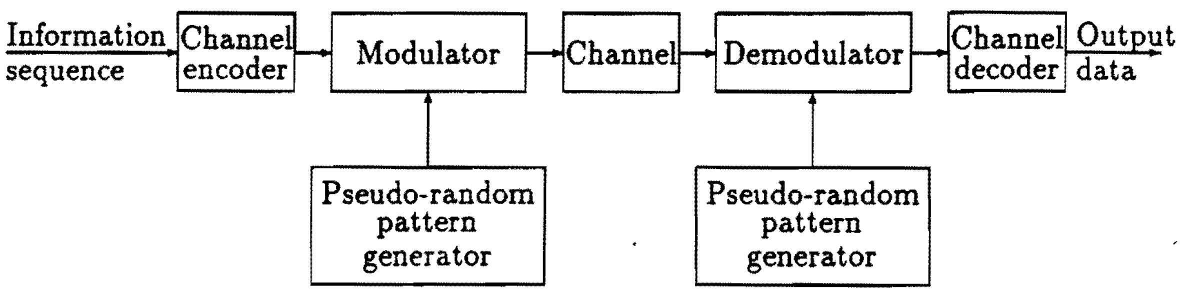


Fig. 1. Block diagram of spread spectrum digital communications system.

There are two main spread spectrum techniques: frequency hopping (FH-SS) and direct sequence modulation (DS-SS). In FH-SS the transmitter repeatedly changes (hops) the carrier frequency in a pattern determined by the spreading code. In DS-SS the transmitter multiplies the information signal by a digital sequence of a much higher bit rate than the information signal. The amplitude of this sequence changes from -1 to $+1$ according to the spreading code.

Consider, for example, a DS system using binary signals [2]. The information signal $b(t)$ is then a rectangular pulse with pulse duration of T_b which has an amplitude of either $+\sqrt{E_s}$ or $-\sqrt{E_s}$, where E_s is the energy of the signal. The pseudo-random sequence $x(t)$ is also a sequence of rectangular pulses; however, each pulse has a duration T_c which is much shorter than T_b. These pulses are referred to as "chips". Figure 2 shows the data and PN signals.

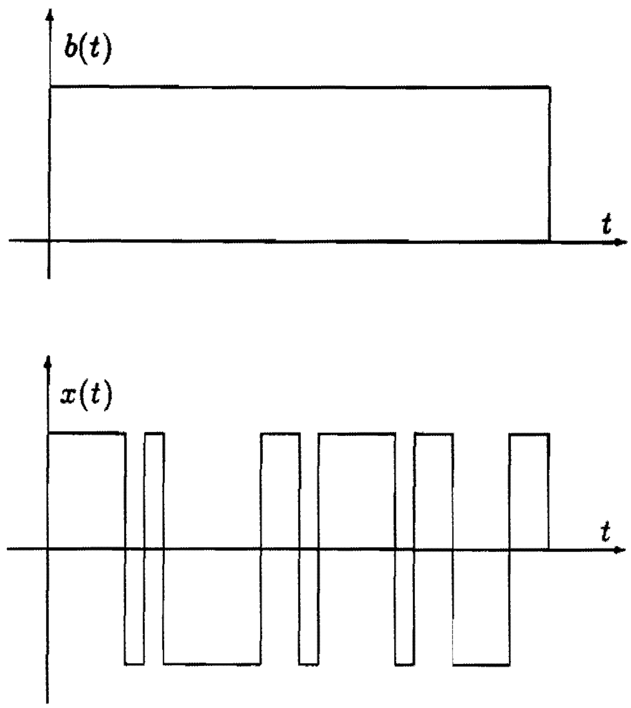


Fig. 2. Data and PN signals.

The transmitted signal $s(t)$ is the multiplicative combination of the data signal, $b(t)$, and the PN signal, $x(t)$: $s(t) = b(t)x(t)$.

The frequency domain representations of the original signal $b(t)$ and the spread spectrum signal $s(t)$ are shown in Fig. 3. The received signal is therefore $r(t) = s(t) + n(t)$, where $n(t)$ is noise. In the receiver, the original signal can be retrieved from $r(t)$ by multiplying by $x(t)$:

$$\begin{aligned} r(t)x(t) &= s(t)x(t) + n(t)x(t) \\ &= b(t)\underbrace{x(t)x(t)}_{=1} + n(t)x(t) \\ &= b(t) + n(t)x(t). \end{aligned}$$

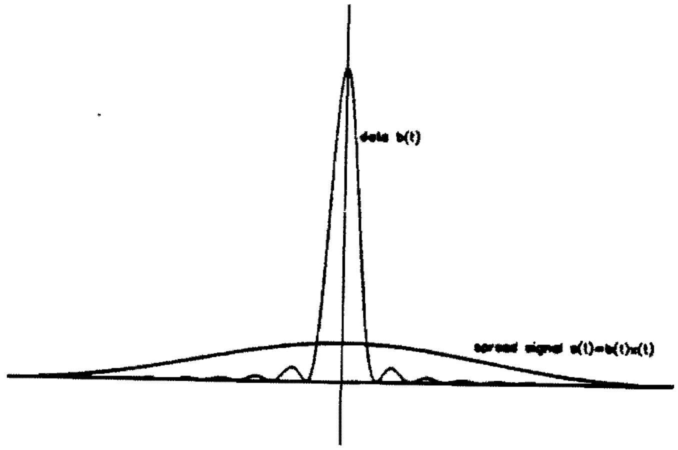


Fig. 3. Power spectrum of data and of spread signal.

Notice that $b(t)$ has been recovered and, as an additional benefit, the noise has been spread thereby reducing its effect. Similarly, when interference signals are present, their effect is also reduced because of the spreading in the receiver. This property is what makes multiple access communications possible.

Various applications of the spread spectrum technique already exist. Here we investigate a possibility of using the technique for a communication system which will provide communication of traffic information between vehicles on a highway and a central traffic control station. Road-side beacons are placed alongside highways. Each vehicle carries its own transmitter and receiver. Each beacon serves only its limited area but is able to simultaneously communicate with the vehicles within that area.

The idea is to use a direct sequence code-division multiple access technique (DS-CDMA) where each vehicle communicating with the road-side beacon in a particular area has its own unique binary code (PN sequence). The goal is to mathematically model such system and to evaluate its error rate performance and power requirements by simulation. A mathematical model of a DS-CDMA system is presented merely as a mathematical description of the communication system preceeded by a discussion on PN sequences. This is given in the second section. Evaluation of the system performance is performed based on the given model for the values of the parameters expected on a highway setting. This is given in the third section. Conclusions are given in the fourth section. The code of the simulation program for the evaluation of error rate performance is given in the Appendix.

B. DS-CDMA System Model

DS-CDMA is a spread spectrum multiple-access technique whereby each user (transmitter – receiver pair) is assigned a unique binary code (PN-sequence). An important issue concerning signal acquisition and interference rejection in the design of CDMA systems is the PN sequences selection problem. Several properties of PN sequences that are relevant for DS-CDMA are discussed below. Two important classes of PN sequences are described as well.

1. Properties of PN Sequences

There are two major conditions that a set of PN sequences should satisfy when used for DS-CDMA [3]:

- each sequence in the set can easily be distinguished from a time shifted version of itself and
- each sequence in the set can easily be distinguished from (a possibly time shifted version of) every other sequence in the set.

These conditions can be checked by examining the autocorrelation and crosscorrelation functions of the sequences in the set.

Consider a set of binary N-periodic sequences $\{\boldsymbol{x}_k\}_{k=1}^{K}$, $\boldsymbol{x}_k = \ldots, x_k^{-1}, x_k^0, x_k^1, x_k^2, \ldots$ whose elements are from the set $\{0, 1\}$. The autocorrelation function of the sequence $\boldsymbol{x}_k$ is defined as

$$\phi_{kk}(j) = \sum_{i=1}^{N}(2x_k^i - 1)(2x_k^{i+j} - 1) \quad 0 \leq j \leq N - 1. \tag{1.1}$$

Ideally, to satisfy the first of the above conditions, a PN sequence should have an autocorrelation function with the property that $\phi_{kk}(0) = N$ and $\phi_{kk}(j) = 0$ for $1 \leq j \leq N - 1$.

The crosscorrelation function of two sequences $\boldsymbol{x}_m$ and $\boldsymbol{x}_k$ is defined as

$$\phi_{mk}(j) = \sum_{i=1}^{N}(2x_m^i - 1)(2x_k^{i+j} - 1) \quad 0 \leq j \leq N - 1. \tag{1.2}$$

Ideally, to satisfy the second of the above conditions, the PN sequences in the set should be mutually orthogonal. However, orthogonality among a number of PN sequences is not easily achieved, especially if the number of the sequences is large.

The selection of a good set of PN sequences is an important problem in the design of CDMA systems. The most commonly considered classes of the sequences are discussed in [3]. Two of these classes, maximal connected sets of m-sequences and Gold sequences, are described below.

2. Sets of PN Sequences

Both maximal connected sets of m-sequences and Gold sequences are based on maximum length shift register sequences. A maximum length shift register sequence, or an m-sequence, of period $N = 2^n - 1$ is any sequence generated by an n-stage shift register with linear feedback connections specified by a primitive polynomial of degree n. Therefore, there are as many different m-sequences of period $N = 2^n - 1$ as there are different primitive polynomials of degree n.

These sequences have the autocorrelation function that takes values $\phi_{kk}(0) = N$ and $\phi_{kk}(j) = 1$ for $1 \leq j \leq N-1$, which is almost ideal concerning the first condition that PN sequences should satisfy. On the other hand the peak value of the periodic crosscorrelation function between any pair of m-sequences of the same period can be relatively large compared with the peak value of the autocorrelation function [3]. Such high values of the crosscorrelations are unacceptable in CDMA. It is, however, possible to construct small subsets of m-sequences with good crosscorrelation properties.

a. Maximal Connected Sets of m-Sequences

Gold [4] and Kasami [5] proved that certain pairs of m-sequences of period $N = 2^n - 1$ have three-valued crosscorrelation function with values from the set $\{-1, -t(n), t(n)-2\}$ where

$$t(n) = 1 + 2^{\lfloor (n+2)/2 \rfloor}. \tag{1.3}$$

These pairs of m-sequences are called *preferred sequences*, and the corresponding pairs of primitive polynomials are called *preferred polynomials.* Preferred polynomials can be obtained from the tables of minimal polynomials of elements in GF(2^n), [6]. If $h_1(p)$ is the minimal polynomial of a primitive element α in GF(2^n) and $h_2(p)$ is the minimal polynomial of the element α^{2^k+1} and $\gcd(n,k) = 1$, then $h_1(p)$ and $h_2(p)$ are preferred polynomials.

A connected set of m-sequences is a collection of m-sequences which has the property that each pair in the collection is a preferred pair. A largest possible connected set is called a *maximal connected set.* Construction of maximal connected sets is described in [3].

Maximal connected sets of m-sequences are useful in those applications which require only a few sequences with excellent crosscorrelation and autocorrelation properties. The number of sequences in the set is usually too small for CDMA applications. It is therefore desirable to obtain larger sets of sequences with good periodic autocorrelation and crosscorrelation properties.

b. Gold Sequences

One important class of periodic sequences which provides larger sets of periodic sequences with the same bound on the peak crosscorrelation as for the maximal connected sets is the class of Gold sequences.

From a pair of preferred sequences $\boldsymbol{u}$ and $\boldsymbol{v}$, a set of N new sequences can be constructed by taking the modulo two sum of $\boldsymbol{u}$ and the N cyclicly shifted versions of $\boldsymbol{v}$ or vice versa. The $N+2$ sequences constructed this way together with $\boldsymbol{u}$ and $\boldsymbol{v}$ are called *Gold sequences.*

The crosscorrelation function for any pair of Gold sequences is three-valued. The possible values are from the set $\{-1, -t(n), t(n)-2\}$, where $t(n)$ is given by (1.3). The off-peak autocorrelation function for a Gold sequence is also three-valued and takes values from the same set. A set of Gold sequences of period N is optimal when N is an odd number considering the Sidelnikov lower bound on the crosscorrelation of two binary sequences of period N in a set of N or more sequences [3].

3. Communication System Description

In the multi user systems [7] that employ the binary PSK direct-sequence spread spectrum modulation transmitted signals are of the form

$$s_k(t) = Ax_k(t)b_k(t)cos(2\pi f_c t + \varphi_k) \quad 1 \leq k \leq K, \tag{1.4}$$

where $b_k(t)$ are binary baseband data signals, $x_k(t)$ are binary baseband spectrum-spreading signals, and K is the number of users. The binary data signals $b_k(t)$ are given by

$$b_k(t) = \sum_i b_k^i p_T(t - iT) \quad 1 \le k \le K, \tag{1.5}$$

where $p_T(t)$ is the rectangular pulse of duration T, and $\boldsymbol{b}_k = \ldots, b_k^{-1}, b_k^0, b_k^1, b_k^2, \ldots$ are the binary data sequences. The spectrum-spreading signals $x_k(t)$ are given by

$$x_k(t) = \sum_j (2x_k^j - 1)\psi(t - jT_c) \quad 1 \le k \le K, \tag{1.6}$$

where $\psi(t - jT_c)$ is a pulse (usually called chip) of duration T_c normalized so that

$$\frac{1}{T_c}\int_0^{T_c} \psi^2(t)dt = 1,$$

and $\boldsymbol{x}_k = \ldots, x_k^{-1}, x_k^0, x_k^1, x_k^2, \ldots$ are periodic binary sequences with period N, as defined in the previous section. The ratio between the data pulse duration and the chip pulse duration is given by $T/T_c = N$. Therefore, the bandwidth expansion from the signal $b_k(t)$ to the signal $x_k(t)b_k(t)$ is N.

As mentioned, in DS-CDMA multiple asynchronous users simultaneously use the common channel for transmission of information. Thus the composite signal present in the channel is of the form

$$r(t) = \sum_{k=1}^{K} s_k(t - \tau_k) + n(t), \tag{1.7}$$

where τ_k is the time delay associated with the k-th signal, and $n(t)$ is additive white Gaussian noise with variance $\mathcal{N}_0$. A particular receiver processes the composite signal in order to extract the transmitted information intended for it. The success of the receiver depends on sequences used, the number of interfering signals and their relative power levels, and the type of processing employed.

Since $n(t)$ is white Gaussian noise, the optimum demodulator can be implemented either as a filter matched to $\psi(t)$ or as a correlator. A possible demodulator structure is shown in Fig. 4.

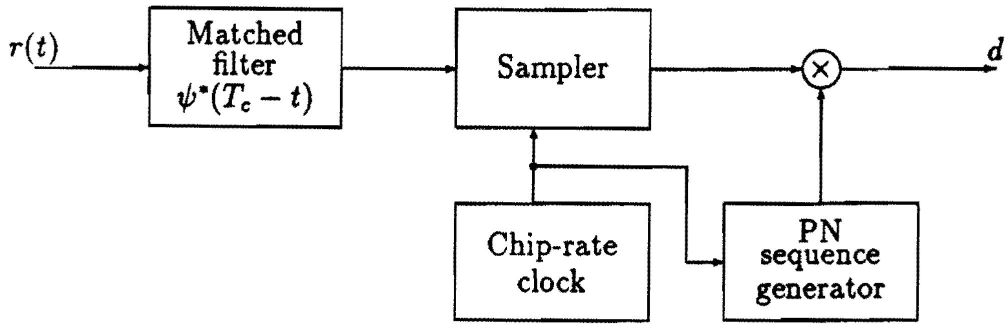


Fig. 4. Demodulator for PN spread spectrum signals.

In order to analyze error rate performance of the m-th receiver we assume $\varphi_m = 0$ and $\tau_m = 0$ since only the relative phase angles and time delays need to be considered. We further assume that $\tau_k = l_k T_c$ for $1 \le k \le K$, where l_k are arbitrary integers, and the receivers are synchronized with the signals they attempt to demodulate. Thus, for the demodulator shown in Fig. 4. with the received signal $r(t)$ given by (1.7), the decision variable d is of the form

$$d = \sqrt{\frac{2E_s}{\mathcal{N}_0}} \frac{1}{N} \phi_{mm}(0) + \frac{1}{N} \sum_{\substack{k=1 \\ k \neq m}}^{K} \phi_{mk}(l_k) + Z, \tag{1.8}$$

where Z is zero-mean, unit-variance Gaussian random variable.

C. DS-CDMA System Performance Evaluation

1. System Requirements

In order to determine the requirements for the communication system, several pieces of data must be obtained. The most important information needed is the maximum number of vehicles that will be in the coverage area at any time. This is determined from the density of cars on the roadway, and the time a vehicle is within the coverage area.

In a study by Urbanik, Hinshaw, and Barnes, traffic densities for 7 locations throughout the Texas were observed [8]. The highest average peak flow rate was just under 2300 vehicles/hour/lane measured across 4 lanes of an 8-line highway. Assuming a worst case scenario, this reduces to 5.11 cars passing a given point every second. Assuming a worst case of this high density traffic moving at 60 miles per hour, each car covers 100 meters every 3.75 seconds. Combining these results, we can assume a worst case scenario of approximately 20 vehicles moving at 60 miles per hour within our 100 meter area of coverage at a time. If the number of vehicles were to increase above this level, we can expect the speed to be reduced which would allow us more time to communicate with each vehicle.

In [1], the authors discuss information transmissions varying from short messages of 100 bytes to images of perhaps 25 kbytes. For our purposes, the ability to communicate 500 kbits (62.5 kbytes) should be sufficient capacity.

2. Error Rate

The bit-error probability in DS-CDMA systems is due to two factors: presence of the noise in the channel and presence of multiple users in the same channel. The effect of the first factor can be decreased by increasing signal to noise ratio (SNR), which means increasing required power. The effect of the second factor, as discussed in the previous section, can be decreased by increasing the length of the used PN sequences, which means increasing required bandwidth. For given SNR, as the number of the simultaneous users increases the bit-error probability increases. For given number of the simultaneous users, as the SNR increases the bit error probability decreases to a certain value determined by the number of the users.

A C computer program has been developed to test the receiver error rate performance by simulation. The simulation was performed to examine the maximum number of simultaneous users for which the bit-error probability can go below 10^{-6} with increasing SNR. The performance is tested for two different sets of Gold sequences.

The first set of sequences used in the simulation was the set of Gold sequences of length $N = 127$. The sequences were generated by the shift register structure shown in Fig. 5.

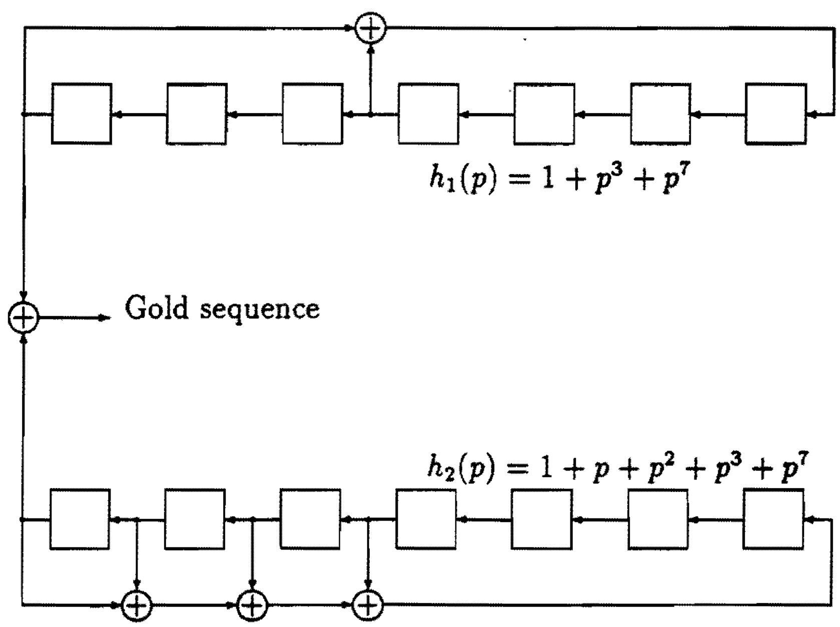


Fig. 5. Generation of Gold sequences of length 127.

The simulation was performed for the number of users $K = 1, 5, 8, 10, 15$, and the results are shown in Fig. 6.

The second set of sequences used in the simulation was the set of Gold sequences

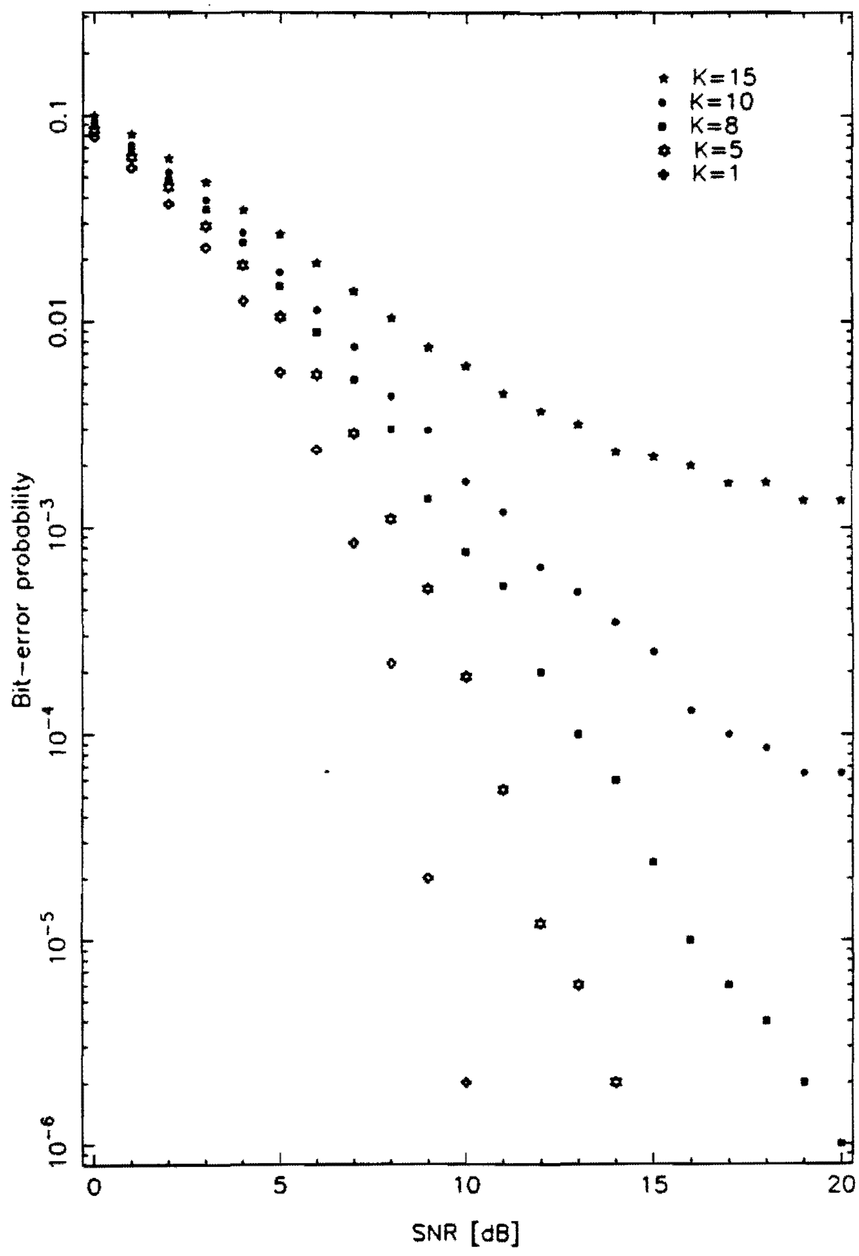


Fig. 6. Probability of error for binary PSK DS-CDMA with Gold sequences of length $N = 127$ for number of users $K = 1, 5, 8, 10, 15$.

of length $N = 511$. The sequences were generated by the shift register structure shown in Fig. 7.

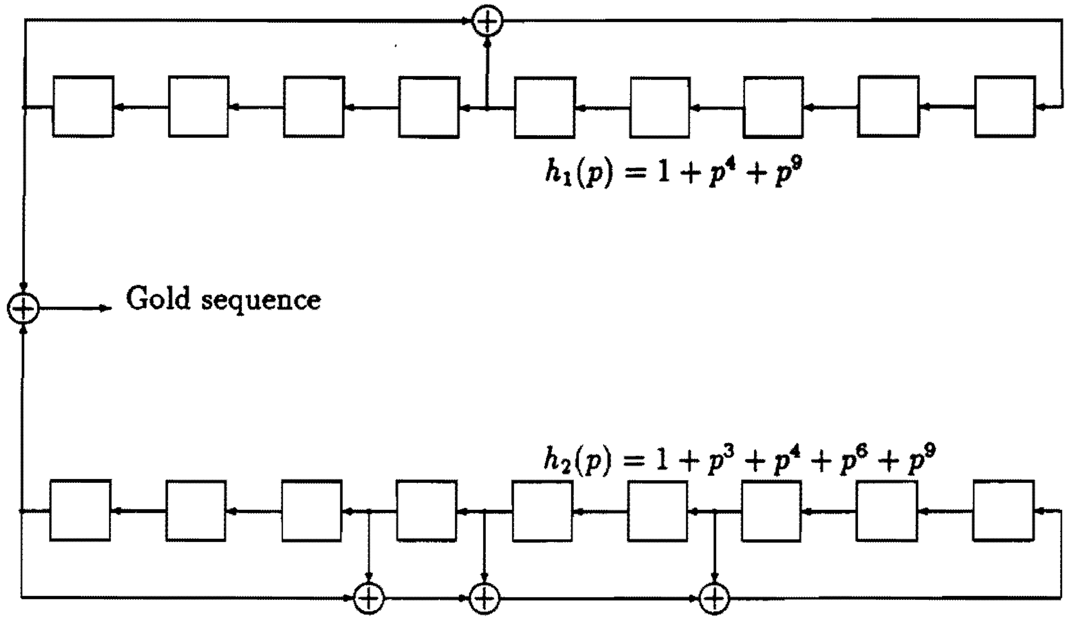


Fig. 7. Generation of Gold sequences of length 511.

The simulation was performed for the number of users $K = 1, 5, 10, 15, 25$, and the results are shown in Fig. 8.

Several conclusions can be made based on the above results. For N=127 Gold sequences, 8 users can be supported at bit error probability Pe=10^{-6} and SNR of 20 dB. For N=511 Gold sequences, 25 users can be supported at bit error probability Pe=10^{-6} and SNR of 20 dB.

3. Power Requirements

In calculating the power requirements for the DS-CDMA system, three factors must be considered: required signal-to-noise (SNR) ratio needed to achieve desired per-

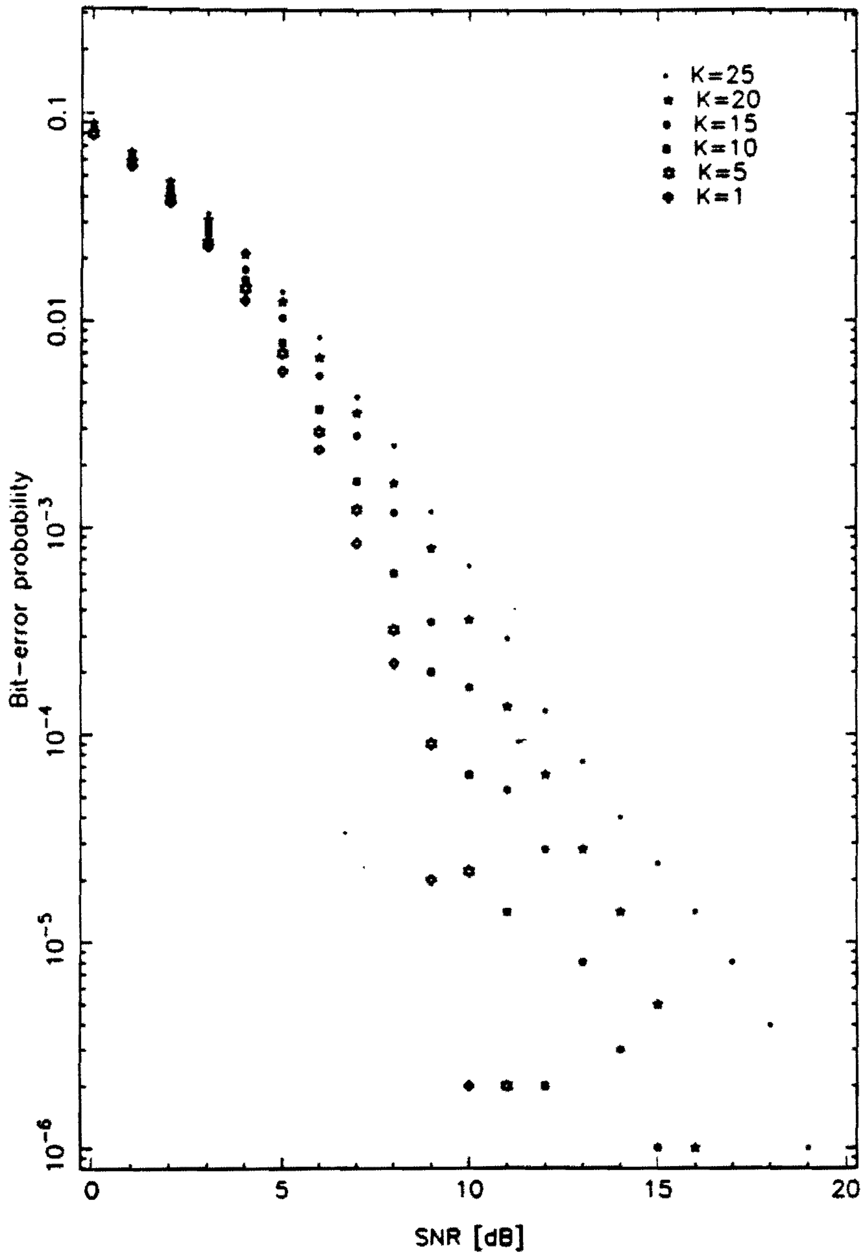


Fig. 8. Probability of error for binary PSK DS-CDMA with Gold sequences of length $N = 511$ for number of users $K = 1, 5, 10, 15, 20, 25$.

formance, attenuation due to distance, and background noise. By using the results of the simulations discussed in the previous section, the required SNR can be determined. Attenuation is determined knowing that the received energy decreases with distance on the power α, where α is a radio propagation parameter that has to be measured for different urban densities. Typical values are in the range 3.3–4, and, to be on the safe side, we take $\alpha = 3.5$. This leaves only the problem of determining the background noise.

Determining the background noise is more difficult and less accurate than calculating the SNR and attenuation. Unfortunately, analytical methods of determining this noise do not exist; however, by taking many empirical measurements in a wide variety of locations and situations, estimates can be made of the average noise. Fortunately, extensive studies have been undertaken and the problem has been reduced to reading the data from charts and tables [9]. So determined, the average background noise, F_a, in the 900 to 950 MHz range is approximately 20 dB relative to kT_0b. F_a is defined as

$$F_a = 10\log \frac{p_n}{kT_0b} \text{ [dB]}, \tag{1.9}$$

where k is Boltzman's constant, T_0 is the reference temperature (310 K), b is the receiver noise bandwidth, and p_n is the average noise power. Notice that F_a was obtained from [9] and that k and T_0 are both constants. We can also calculate b from the equation

$$b = 1.2R \text{ [Hz]}, \tag{1.10}$$

where R is the rate, and 1.2 is an expansion factor. We can now solve equation (1.9) for the only remaining unknown, p_n. The signal power can then be calculated from this average noise power and from the required signal-to-noise ratio. Figure 9 shows the signal power required for 20 users and 25 users over a range of data rates.

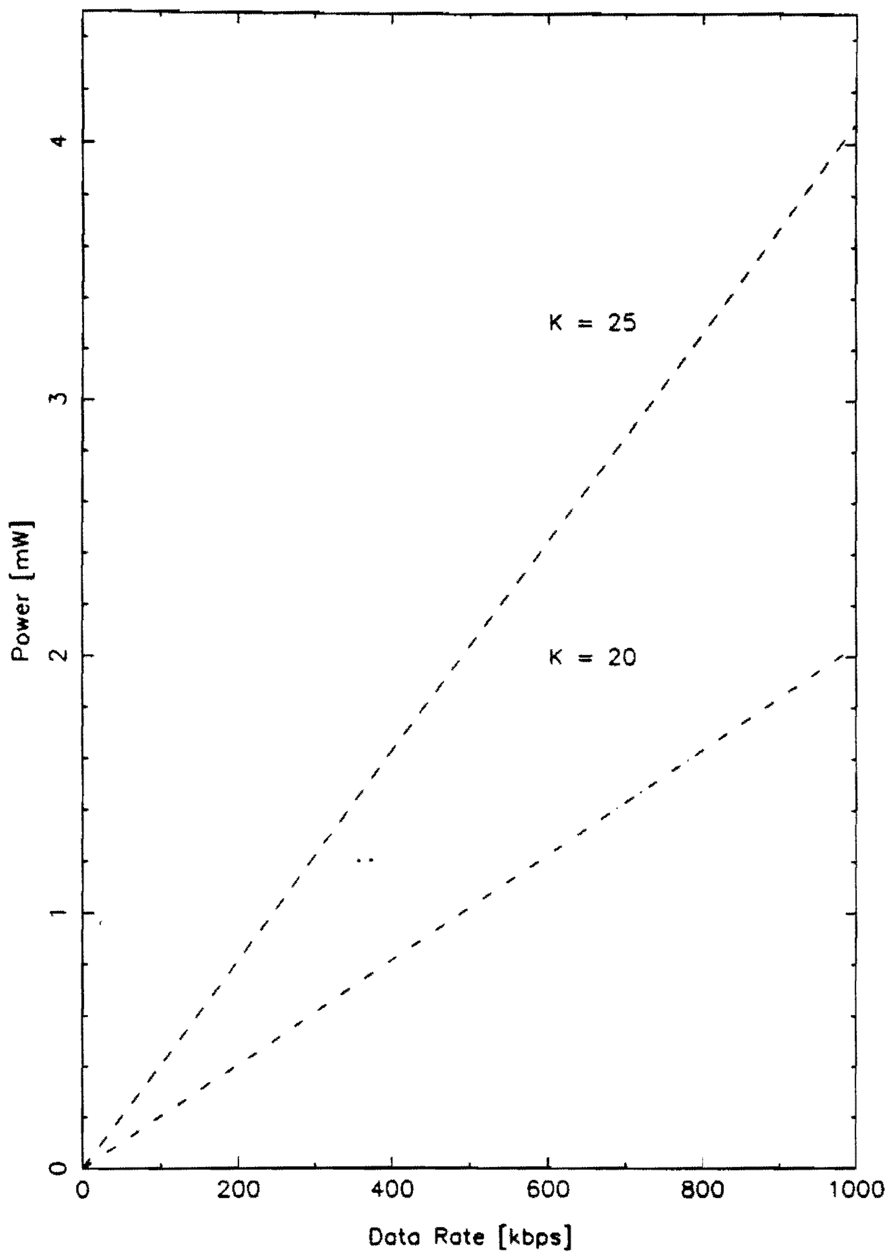


Fig. 9. Power requirements for DS-CDMA for number of users $K = 20, 25$.

4. An Optimization Problem

One problem which must be addressed in designing an intelligent vehicle communication system is in choosing certain parameters. For our system, we want to transmit as much information as possible to as many cars as possible as accurately as possible; however,we would like to use a minimum amount of power and a minimum amount of complexity. Unfortunately, these are not all independent parameters. Gaining an advantage in one area usually requires a trade-off in some other area.

As a starting place, we have previously determined that we pessimistically expect a maximum of 20 vehicles within our coverage area at any one time and we expect them to be within that coverage area for approximately 3.75 seconds. We now have two options for design of our system. We can either have a system which can communicate simultaneously with 20 vehicles, or we can have a system which sequentially communicates to groups of a smaller number of vehicles in order to cover all 20 vehicles within the time frame. Each of these has its advantages and disadvantages. Communicating to several groups of less than 20 requires less complexity, but a higher data rate would be required. Power requirements would also be affected; however, since power increases as data rate increases and decreases as number of users decreases, calculations must be performed to determine whether the effect is beneficial or detrimental. Similarly, communicating with all 20 vehicles simultaneously requires more complexity, but data rate can be decreased.

The previous section provides a method of calculating noise power, and the simulation results provide the signal-to-noise ratios (SNR) required for various combinations of users and performance levels. Using these results, and choosing values for any two other parameters, we can analytically calculate the remaining parameters. We will now examine two cases in which we will choose values for different parameters.

Table I. Optimization Data for Fixed 500 kbits/user

Groups	Users/ group	Time/ group	N	Rate (kb/s)	B/W (MHz)	SNR (dB)	Power (mW)
1	20	3.75	511	133	68	16	0.27
2	10	1.88	511	267	136	12.5	0.24
3	7	1.25	127	400	51	20	0.54
4	5	.94	127	533	68	15	0.87
5	4	.75	127	667	85	14	0.86
7	3	.53	127	933	119	12	0.76
10	2	.38	127	1333	169	11	0.86
20	1	.19	127	2667	339	10	1.37

In both cases, we will set the performance requirement to $P_e = 10^{-6}$, and we will maintain the requirement of communication with 20 users within 3.75 seconds.

Table I shows the results when the number of bits transferred is fixed at 500 kbits/user. The first three columns show the number of groups we have broken the 20 users into, how many users are in each group, and the amount of time allotted for communication with each group. The next column gives the length of the pseudo-random sequence needed for the number of simultaneous users. The remaining columns show the rate required to transmit the 500 kbits/user, the bandwidth required for the sequence length and rate, the SNR required to achieve 10^{-6} error probability, and the signal power needed.

The data shows that the desired configuration is two groups of ten users/group. The information is transmitted at 533 kbits/sec and requires 0.24 mW of power.

Table II. Optimization Data for Fixed 26 MHz Bandwidth

Groups	Users/ group	Time/ group	N	Rate (kb/s)	Bit/user (kbits)	SNR (dB)	Power (mW)
1	20	3.75	511	50.8	190.5	16	0.10
2	10	1.88	511	50.8	95.3	12.5	0.05
3	7	1.25	127	204.7	255.9	20	1.05
4	5	.94	127	204.7	192.4	15	0.33
5	4	.75	127	204.7	153.5	14	0.26
7	3	.53	127	204.7	108.5	12	0.17
10	2	.38	127	204.7	77.8	11	0.13
20	1	.19	127	204.7	38.9	10	0.10

These results require a note of caution however. For this configuration, we must have 136 MHz of available bandwidth. In practice, this might not be available.

In the second case, we examine the situation in which the available bandwidth is limited. Table II shows the results when 26 MHz bandwidth limit is imposed. This table is similar to Table I with the exception of the sixth column which now shows the number of bits/user which can be transmitted during the available time at the highest possible rate.

These results indicate that if the goal is to maximize information transfer, the best configuration is three groups with seven users/groups. Notice that the power levels for these calculations are significantly lower than the calculations in Table I. This is due to the significantly smaller bit rates involved.

These calculations can be performed for any number of variations in the fixed parameters. The ultimate choice of configuration must be based on the exact needs and desires for each individual situation.

REFERENCES

[1] R. Fukui, Y. Noji, and M. Hashizume, "Individual Communication Function of RACS: Road Automobile Communication System," *IEEE Vehicle Navigation and Information System Conference,* 1989.

[2] W. C. Jakes, Jr., *Microwave Mobile Communications,* John Wiley and Sons, New York, NY.

[3] D. V. Sarwate and M. B. Pursley, "Crosscorrelation properties of pseudorandom and related sequences," *Proc. IEEE,* vol. 68, pp. 593-619, May 1980.

[4] R. Gold, "Maximal linear sequences with 3-valued recursive crosscorrelation functions," *IEEE Trans. Inform. Theory,* vol. IT-14, pp. 154-156, Jan. 1968.

[5] T. Kasami, "Weight distribution formula for some class of cyclic codes," University of of Illinois, Urbana, Rept. R-265, April 1966.

[6] S. Lin and D. J. Costello, Jr, *Error Control Coding: Fundamentals and Applications,* Prentice-Hall, Englewood Cliffs, N.J.

[7] M. B. Pursley, "Spread-spectrum multiple-access communications," *Multi-User Communications Systems,* G. Longo, Ed. Springer-Verlag, New York, 1981.

[8] T. Urbanik II, W. Hinshaw, and K. Barnes, "Evaluation of High Volume Urban Texas Freeways," Texas Transportation Institute, Texas A&M University, January 1991.

[9] E. N. Skomal, *Man-Made Radio Noise,* Van Nostrand Reinhold Company, New York, NY.

APPENDIX A

SIMULATION PROGRAM CODE

```
#include <stdio.h>
#include <math.h>

#define N 511
#define Nu 25
#define ns 16

#define snrl 0
#define snru 21
#define I 1000000

#define pi 3.14159265

double drand48();
double erfc(double x);
void srand48(long seed);
double pow(double x, double y);

int sb, rb;

unsigned long as[ns+1];
unsigned long bs[ns+1];
long gs[N+5][ns+1];
int sumus[N+5][N];
int sumu;

bitcount1(p)
unsigned long p;
{
int b;
unsigned long pb;

pb = (p<<1)>>1;
for (b = 0; pb != 0; pb >>=1)
    if (pb & 01)
 b++;
return (b);
}

bitcount(p)
unsigned long p;
{
int b;
```

```
for (b = 0; p != 0; p >>=1)
    if (p & 01)
 b++;
return (b);
}

void coder(cs, fbc)
int cs[ns+1];
int fbc;
{
int srp = 1;
int i, j;

for (j = 0; j < ns; j++)
{
cs[j] = 0;
for (i = 0; i < 32; i++)
{
cs[j] = cs[j] | (((srp&256)>>8) << i);
srp = ((srp<<1) | (bitcount(srp&fbc)&1))&N;
}
}
}
void pn()
{
unsigned long z[ns+1];
int i, j, f, s;

coder(as,272);
coder(bs,308);

for (j = 0; j < ns; j++)
{
z[j] = bs[j];
gs[N][j] = as[j];
gs[N+1][j] = bs[j];
}
z[ns] = z[0];

for (i = 0; i < N; i++)
{
for (j = 0; j < ns; j++)
        gs[i][j] = as[j] ^ z[j];

for (j = 0; j < ns-1; j++)
z[j] = ((z[j+1]&1)<<31) | (z[j]>>1);

z[ns-1] = ((z[ns]&1)<<30) | (z[ns-1]>>1);
z[ns] = z[0];
}
```

```
for (j = 1; j < Nu; ++j)
{
for (i = 1; i < N; ++i)
{
for (s = 0; s < ns; ++s)
z[s] = gs[j][s];

cshift(z,i);

f = 0;
for (s = 0; s < ns-1; ++s)
f = f + 2*bitcount(z[s]^gs[0][s]);

f = f + 2*bitcount1(z[ns-1]^gs[0][ns-1]);
sumus[j][i] = N - f;
}
}
}

cshift(ws, nsh)
unsigned long ws[ns+1];
int nsh;
{
int k, j;
for (k = 0; k < nsh; ++k)
{
for (j = 0; j < ns-1; j++)
ws[j] = ((ws[j+1]&1)<<31) | (ws[j]>>1);

ws[ns-1] = ((ws[ns]&1)<<30) | (ws[ns-1]>>1);
ws[ns] = ws[0];
}
}

void channel()
{
unsigned long z[ns+1];
int j, sg, sh;

sb = 2*drand48();
if (sb == 0) sumu = - N; else sumu = N;

for (j = 1; j < Nu; j++)
{
sg = 2*drand48();
sg = 2*sg - 1;
sh = N*drand48();
sumu = sumu + sg*sumus[j][sh];
}
}
```

```
main()
{
int i, j, snrd;
double snr, snrc;
double a, b , n;
double d, pe;
double s;

s = sqrt(N);
pn();

for (snrd = snrl; snrd < snru; snrd++)
{
srand(1);
snr = pow(10,snrd/10.);
snrc = sqrt(2*snr);
pe = 0;
for (j = 0; j < I; ++j)
{
channel();
a = drand48();
b = drand48();
n = sqrt(-2*log(a))*cos(2*pi*b);
d = snrc*sumu/N+n;

if (d < 0) rb = 0; else rb = 1;
if (sb != rb) pe = pe + 1;
}
printf("%d,", snrd);
printf("%f\n", pe/I);
/*printf("%f\n", 0.5*erfc(sqrt(snr)));*/
}
}
```

CHAPTER II

INFRARED ROAD-AUTOMOBILE COMMUNICATION

A. Introduction

The beacon-automobile communication has recently received much attention not only as a means of providing the traffic information and route guidance but also for receiving and transmitting other information.

The beacon is a stationary transmitter or receiver located along the roadway. As a transmitter, it transmits information signals to the passing automobiles coming within its range of transmission; as a receiver, it receives signals transmitted to passing automobiles. To be able to receive or transmit sufficient information as the automobile is speeding by the beacon, the communication has to be done in high data rates. This requirement renders an infrared communication system a very attractive means of communication.

This study analyzes the feasibility of a beacon-to-automobile infrared communication system. The results of this analysis can be also applied to an automobile-to-beacon communication as well.

B. Proposed System Configuration

The most important source of noise in atmospheric infrared communication is the sun. In order for the communication system to operate efficiently with the minimum amount of power, the location of the transmitter and receiver are very crucial. The sun generates strongest radiation at noon and minimum radiation during early morning and late afternoon. Therefore, if unavoidable, the communication system should be exposed to direct radiation from the sun only during the minimum radiation period.

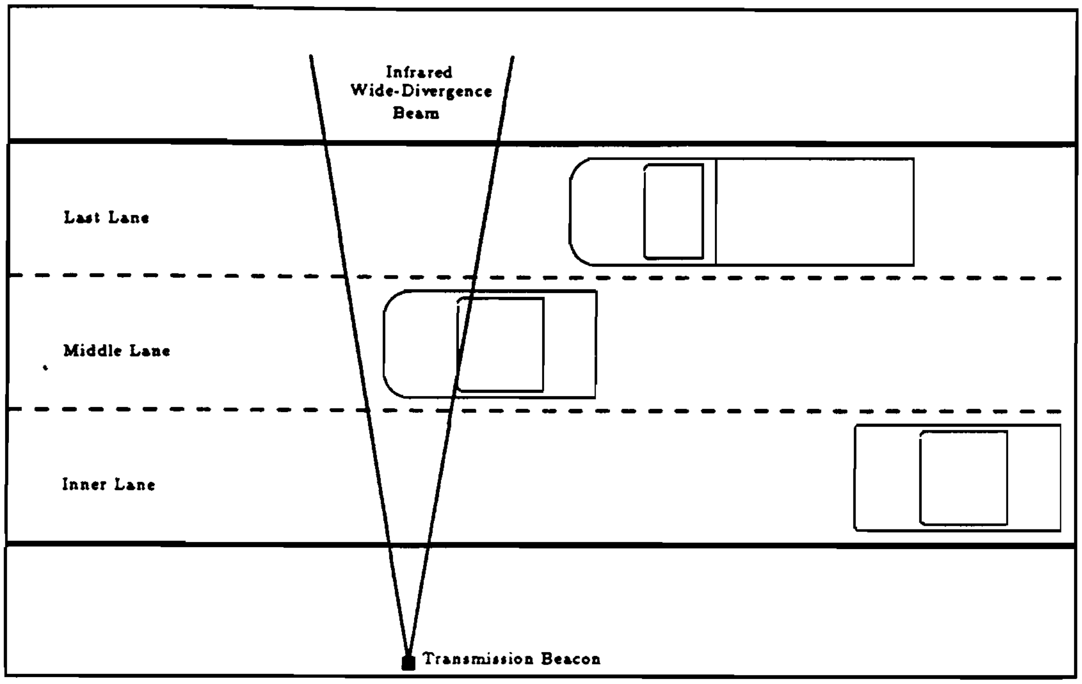


Fig. 10. Proposed infrared wide-divergence beam roadside communication system.

The proposed system analyzed in this study is the infrared wide-divergence beam communication system as shown in Figure 1. The infrared transmitter is located on the side of the road at a height of approximately 1 − 1.2 meters and about 5 meters away from the nearest lane. The transmitting beacon is composed of a semiconductor laser and a lens system which directs the optical energy in a cone-shaped beam toward the roadway. The axis of the radiation cone is parallel to the ground and perpendicular to the road direction. The wide-divergence beam approach is used in order to eliminate the need for tracking between the transmitter and the receiver, and to provide a communication zone for the passing automobiles. In this study we will assume that the roadway has three lanes per traffic direction, and each lane is 12 feet wide. Therefore, the transmission zone for the lane nearest to the transmitter is shortest, and the transmission zone for the out most lane is the longest. In contrast, the nearest lane will receive more transmitted power than the out most lane. This system is designed in such a way that the third lane which is the lane furthest from the beacon has a probability of error of 10^{-6}, and hence the inner lanes should have even better performance but with a shorter transmission zones.

To avoid the possibility of signal obstruction by another automobile, it is the responsibility of the automobile operator to position his/her automobile in the lane nearest to the transmitter or to make sure that no other automobiles are blocking the receiver field of view (FOV) when passing by the beacon.

The receiver is composed of a photodetector in combination with an external energy collecting lens system with a diameter of 5 centimeters and a narrowband optical filter of 10 Angstrom bandwidth which removes some background radiation. A more complicated receiver may have a number of photomultipliers to enhance the received signal. To obtain the maximum transmission zone, the receiver planar angle should equal that of the transmitted signal angle. In addition, we will assume that there are a sun visor and a ground reflection shield to limit the amount of direct sunlight and ground reflection entering the receiver field of view. The receiver is located on the side of the automobile facing the transmitter and should be placed at about the same distance from the ground as the transmitter in order to receive maximum power from the transmitter.

C. Brief Theoretical Background

In this section, we will describe briefly those theories and definitions necessary for the performance analysis of the proposed system.

1. Atmospheric Transmissivity

As an optical signal propagates through the earth's atmosphere, it suffers attenuation due to absorption and scattering of radiation. Atmospheric absorption is caused by molecular particles in the transmission path, such as carbon dioxide, ozone, and water vapor. Wavelength, atmospheric pressure, humidity, and temperature influence the amount of absorption.

Atmospheric scattering is due to particles and gases in the air along the transmission path. If the size of particles is much smaller than the signal wavelength it is treated as Rayleigh scattering [4], if the size is comparable to the wavelength it is treated as Mie scattering, and if the wavelength is much larger that the particles' size then it is non-selective scattering.

There is much theoretical and experimental work done in this area; but we feel that the empirical formulas for calculating absorption transmissivity and scattering transmissivity presented by [2] are relatively not complicated and cover the wavelengths of interest.

The formula to compute absorption transmissivity developed by R.M. Langer and summarized in [2] is an empirical expression based on the number of absorbing molecules encountered by the radiation. Langer's formula is based on the number of precipitable millimeters of water (pr. mm.) which depends on the temperature and relative humidity (R.H.), and it is plotted in Figure 11. The formula can be used to compute the absorption transmissivity of the spectrum in the range of 0.72 to 15 microns which is divided into eight windows as shown in Table III. The empirical expression which can be used to compute the transmittance due to absorption in window i, τ_{ai} is

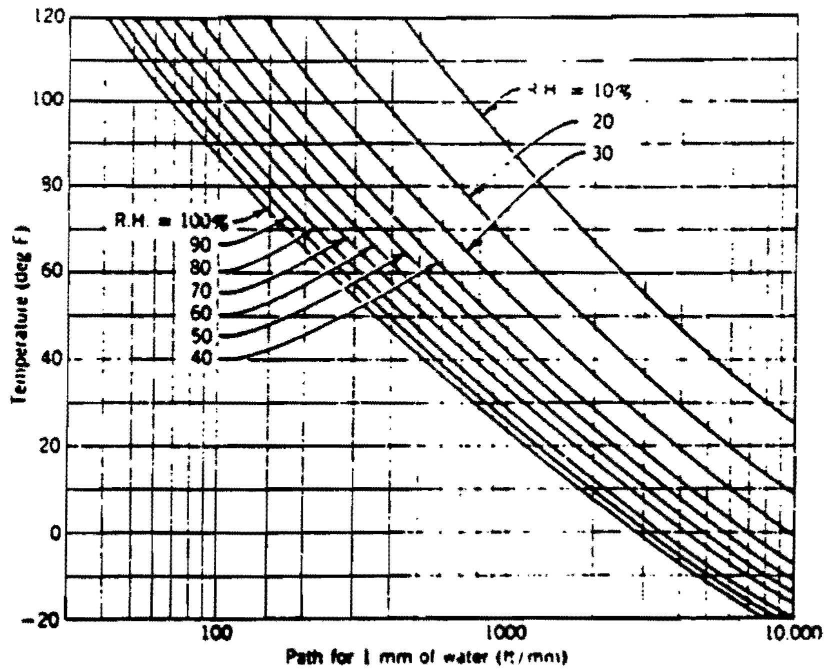


Fig. 11. Path length [ft/mm] to penetrate 1 mm of precipitable water in the atmosphere. (This figure is taken from [2].)

Table III. Wavelength Window

Window No.	Window Boundaries(in micron)
I	0.72 –0.94
II	0.94 –1.13
III	1.13–1.38
IV	1.38–1.9
V	1.9–2.7
VI	2.7–4.3
VII	4.3–6.0
VIII	6.0–15.0

Table IV. Coefficients for Computing Scattering and Absorption Transmissivity

Window	A_i	k_i	β_i	w_i
I	0.0305	0.800	0.112	54
II	0.0363	0.765	0.134	54
III	0.1303	0.830	0.093	2
IV	0.211	0.802	0.111	1.1
V	0.350	0.814	0.1035	0.35
VI	0.373	0.827	0.095	0.26
VII	0.913	0.679	0.194	0.18
VIII	0.598	0.784	0.122	0.165

$$\tau_{ai} = \begin{cases} \exp(-A_i w^{\frac{1}{2}}) & (w < w_i), \\ k_i(\frac{w_i}{w})^{\beta_i} & (w \geq w_i), \end{cases} \tag{2.1}$$

where

w is precipitable millimeters of water in path (pr.mm.),

w_i is is the value of w that causes the absorption in window i to undergo a transition from weak-band to strong-band absorption.

The values of constants k_i, A_i, β_i, and w_i are shown in Table IV., and the value of w can be obtained by calculation based on Figure 11 and the propagation distance.

The empirical equation for calculating scattering transmissivity presented by [2] uses the visual range which is usually reported daily by the Weather Bureau as the calculating parameter. The scattering transmissivity for window i, τ_{si} can be computed by using

$$\tau_{si} = \exp\{-\frac{3.91R}{V}[\frac{\lambda_i}{0.55}]^{-q}\}, \tag{2.2}$$

where

V is the visual range in kilometers,

q is equal to $0.585V^{\frac{1}{3}}$,

R is the propagation distance in kilometers, and

λ_i is the middle wavelength of window i.

The relation $q = 0.585V^{\frac{1}{3}}$ is applicable for visual ranges less than 6 Km. Middleton suggests in [1] that the value of q when seeing conditions are good is 1.6 and on an average day, $q = 1.3$.

The atmospheric transmissivity (also called extinction coefficient), τ_{ei}, corresponding to window i, is the product of the scattering transmissivity and absorption transmissivity as shown below:

$$\tau_{ei} = \tau_{ai}\tau_{si}. \tag{2.3}$$

2. Atmospheric Turbulences

When atmospheric air is warmed heat from the ground, or other heated objects, it rises. The wind tends to break up this rising air mass and mixes it with other air masses of different temperature causing variations in temperature from point to point. These variations in temperature induce random fluctuations in the atmospheric index of refraction. This phenomenon is called turbulence [3]. When an optical beam propagates through the atmosphere with temperature varying from point to point, the beam may partially or totally deviates depending on the beam size and temperature inhomogenity. This interaction between the optical wave and turbulent medium causes random amplitude and phase fluctuations [4]. Turbulence induces many effects on atmospheric optical communications, for example, beam steering, spreading, scintillation, spatial coherence degradation, and polarization fluctuation.

In the case of our infrared wide-divergence beam communication system, scintillation is the most important. Scintillation is power fluctuation at the receiver which is caused by different portions of the wavefront to experience different phase changes when propagating through a turbulent atmosphere [3]. The received power, $P_r(t)$, can be described as

$$P_r(t) = P_{ro}(t)\exp[2\chi(t)], \tag{2.4}$$

where

$P_{ro}(t)$ is the received power with no scintillation effect,

$\chi(t) = \ln(A(t)/A_0)$, is a Gaussian process,

of variance σ_χ^2 and mean $-\sigma_\chi^2$ [14],

$A(t)$ is the received amplitude, and

A_0 is the received amplitude without scintillation effect.

There are many theories trying to deal with turbulence fluctuation. It has been found experimentally that the amplitude fluctuations do not increase with either increasing path length or turbulence strength; rather they saturate to a value of approximately 0.5 for both plane and spherical waves [5].

3. Noise in Detection Process

In atmospheric optical communication there are basically two types of noise: first, thermal noise, and second, quantum noise [6]. Thermal noise originates within the system and is caused by the thermal fluctuation of electrons inside the system. In good systems, the magnitude of the thermal noise is very small and can be neglected.

Quantum noise is composed of dark current and background radiation. Dark current is the output current flow that appears even in the absence of incidence light and is caused by random emissions from the detector surface due to inherent thermal energy [3]. Dark currents can be modeled at various degrees of complexity. In simplified form, we can model dark current by the Richardson-Dushman equation with blackbody temperature of $T = 300$ K^0, as shown below:

$$I_d = .012T^2 e^{-\frac{hf_c}{kT}} \quad (\text{A.m}^{-2}), \tag{2.5}$$

where

h is Planck 's constant,

f_c is the cutoff wavelength of detector, and

k is Boltzmann 's constant.

It can be represented in power units as:

$$P_d = \frac{hfI_d}{e\eta}, \tag{2.6}$$

where

f is the operating frequency,

e is electron charge, and

η is the photodetector quantum efficiency which is the ratio of detected field power and incident field power.

The dark current obtained by this formula is a conservative (worst-case) result, and is an upperbound to dark current from any real device [7, 8].

Second, background radiation results from emission of optical frequencies from sources whose temperatures are other than absolute zero and also those reflected from them. The important sources of background radiation are the sun, sky, moon, and stars, and earth objects. The role of these sources depends on the atmospheric conditions and their positions relative to the receiver. Radiations from such sources can be modeled approximately by Planck 's law of blackbody radiation [4, 6]. The spectral radiance, N_λ, of an object with temperature T can be expressed as

$$N_\lambda = \frac{2hc^2}{10^6\lambda^5}[\exp(hc/\lambda kT) - 1]^{-1} \quad (\text{W.m}^{-2}.\text{sr}^{-1}.\mu\ \text{m}^{-1}), \tag{2.7}$$

where

h is Planck 's constant,

c is velocity of light,

λ is the optical wavelength used,

k is Boltzmann 's constant, and

T is temperature of background radiation source (in degrees Kelvin).

The spectral radiance is frequency dependent as can be seen from equation (2.7). The peak of spectral radiance of an object with temperature T K^0 occurs at frequency

$$\lambda_m = \frac{2897}{T} \quad (\mu\ \text{m}). \tag{2.8}$$

In general, real background objects are not perfect blackbodies; however, when heated to a temperature T such an object which is referred to as graybody will have a spectral radiance very similar to a blackbody but smaller in magnitude [6].

Background radiation from sunlight can be approximately represented by radiation from a blackbody with temperature 6000 K^o [4]. Daytime sky radiance depends on the position of the sun, the receiver, and atmospheric conditions. Sky radiance is produced by two different mechanisms: scattering of sun radiation and emission of atmospheric constituents [9]. The scattering is significant only in daytime and for those sources with wavelength below 3 microns. The radiance of the sky produced by scattering sunlight for wavelengths in this band can be approximated to be 2×10^{-5} times the amount of the sun radiation at each wavelength [9]. For wavelength longer than 4 microns, emission is important, and sky radiance can be approximated by blackbody at temperature of 300 K^0

Thermal radiation emitted by the earth and its atmosphere can be modeled by blackbody of temperatures between 218 K^o and 288 K^o [6]. This thermal background noise is referred to as earthshine.

In roadside communication, if the automobile is not in the lane nearest to the transmitter, then background radiation might include the radiation from the automobile engine and exhaust system. This radiation can be approximated by blackbody radiation of temperature 400 K^0.

During nighttime, background radiation from the moon and stars is actually reflected radiation from the sun. These radiations can be obtained by modeling them as 6000 K^0 blackbody radiation from the sun; but the radiation intensities are reduced by the absorption of the body. The absorption factor for each of the body is called *albedo*, and it is frequency dependent. More details can be obtained from [4].

The received background radiation power from each of these sources can be expressed as

$$P_b = N_\lambda \Omega_s A_r \tau_r \tau_e B \frac{\cos\psi}{\cos\phi} \qquad \text{(Watts)}, \tag{2.9}$$

where

Ω_s is solid angle subtended by the background source at the receiver,

A_r is the area of the receiver,

τ_a is the transmittance of the atmosphere,

τ_r is the receiver transmissivity,

B is the bandwidth of the receiver filter,

ψ is the the angle of incidence of the radiation from the background source onto the receiver, and

ϕ is the angle of the surface of the background source in cline from the plane normal to the direction of Ω_s [6].

If the field of view is narrow and background radiation is normal to Ω_s, and if we assume that background radiation extends and fills the entire receiver's field of view, then $\Omega_s = \Omega_r = \frac{\pi}{4}\theta_r^2$ (approximation for small field of view); θ_r is the planar angle of receiver field of view . Hence

$$P_b = N_\lambda \pi \theta_r^2 A_r \tau_r \tau_e B \qquad \text{(Watts)}. \tag{2.10}$$

4. Direct-detection Receiver

The receiver employed in the proposed beacon-to-automobile communication system is a direct detection receiver. It is basically a power detecting system which employs the combination of lens and photodetector to collect the instantaneous power of the arriving optical field. From quantum theory, the lowest amount or quantum of electromagnetic radiation is called a *photon* and has an energy equal to hf, where h is Planck's constant and f is the frequency of electromagnetic radiation. Once a photon hits the surface of the photodetector, an electron is released with probability η which is called the quantum efficiency of the photodetector. The quantum efficiency depends on the properties of photosensitive material. Assuming that the bandwidth of the photodetector is large, then the response of the photodetector can be modeled as a simple conditional Poisson process with intensity $(\lambda_s(t)+\lambda_n(t))$. $\lambda_s(t)$ is the intensity of photoelectrons generated by the signal impinging on the photodetector, and $\lambda_n(t)$ is the photoelectron intensity generated by background radiation, dark currents, and thermal noise. The relation between power and intensity is

$$\lambda(t) = \frac{P(t)\eta}{hf}. \tag{2.11}$$

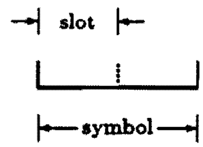


Fig. 12. PPM signal format.

Fig. 13. BPPM signal set.

5. Binary Pulse Position Modulation (BPPM)

BPPM is a signaling scheme in which the signaling interval containing a symbol is subdivided into two equal subintervals called slots as shown in Figure 12. Exactly one optical pulse will be transmitted in one of the slots in each symbol to signify one of two symbols in the signal set as shown in Figure 13.

D. System Analysis

1. Transmission Wavelength Selection

To select the operating frequency of the proposed infrared roadside communication system, we have to consider not only the equipment available, but also the background radiation, and atmospheric attenuation.

From the theoretical background, we know that background radiation at the receiver comes from the sun, sky, stars, objects with elevated temperature, and reflection from these sources. The maximum radiation from the sun occurs at the spectral frequency of .5 microns. The radiation from earth objects including the automobile engine and exhaust pipe have temperatures in the range of 200K° to 400K° and wavelengths in the region of 7 to 15 microns. The frequency range which has minimum radiation effect from the sun and earth objects, both in daytime and nighttime, is in the range of 3 to 4 microns. However, at the present time, to our knowledge there is no semiconductor laser which can be used for our purpose in this frequency range, and a laser operating in the frequency nearer to the natural frequency associated with a blackbody emission at 300 K^0 has higher dark currents. Therefore, we have to chose the transmission frequency at the lower range. From the structure of our communication system, the sun radiation has more effect on the performance of the receiver during early morning and late afternoon. The frequency range in the first window, as shown in Table III, effectively reduces the sun radiation effectively due to scattering, and a semiconductor laser at this frequency is available. Therefore, we decided on .82 microns as the operating frequency.

2. Received Signal Power Analysis

The system structure described previously, in which the roadside beacon radiates power in a cone shape, is shown in Figure 10. The received signal power (neglecting time delay in propagation and scintillation effects) can be expressed as

$$P_r(t) = \frac{A_r \cos(\theta)}{4\pi \sin^2(\frac{\alpha}{4}) R^2} P_t(t) \tau_t \tau_r \tau_e, \tag{2.12}$$

where

$P_t(t)$ is the transmitted power,

A_r is the area of the optical collecting lens at the receiver,

θ is the angle between infrared beam and the receiver plane,

σ is the angle at the apex of the transmitted beam,

τ_t is the transmitter transmissivity,

τ_r is the receiver transmissivity,

τ_e is the atmospheric transmissivity, and

R is the distance from transmitter to receiver.

When scintillation effects are considered, the received power can be expressed as

$$P_r(t) = \frac{A_r \cos(\theta)}{4\pi \sin^2(\frac{\sigma}{4}) R^2} P_t(t) \exp[2\chi(t)] \tau_t \tau_r \tau_e, \tag{2.13}$$

where $\chi(t)$ is as described by equation (2.4).

3. Detection Noise Analysis

Based on the fact that background radiation is stronger during daytime than during nighttime, if the system can perform up to specifications during daytime, then it will be able to perform even better during nighttime. Hence in this analysis we will consider only a daytime situation. To be pessimistic in this analysis, we will assume a high background noise situation.

The received noise power, P_n, is the combination of

$$P_n = P_b + P_d + P_T, \tag{2.14}$$

where

P_n is received noise power,

P_b is background noise,

P_d is power due to dark currents, and

P_T is power due to thermal noise.

As previously mentioned, the background noise, P_b, during daytime comes from the sun both directly and indirectly, earth objects, thermal radiation by the earth surface and earth atmosphere, and radiation reflected from these sources. Therefore, the background noise varies from one location to the other and depends on atmospheric conditions.

For the background radiation generated by the sun, we assume that the sun is in the FOV of the receiver. Based on the receiver position as purposed, the receiver will have the sun in the field of view only during early morning or late afternoon. The spectral irradiance of the sun outside the atmosphere can be approximated by equation (2.7). The propagation distance of the sun radiation at different zenith angles is specified by the number of air masses. The number of air masses in the path of the sun radiation at different zenith angles up to 62 degrees is approximately equal to $\sec(\theta)$, where θ is the zenith angle. For zenith angles greater than 62 degrees, the air mass is no longer approximated by $\sec(\theta)$. The effective air mass must be calculated by tracing the sun rays through the atmosphere. More details can be found in [10]. Each air mass is approximately 8.434 km in distance. The atmospheric extinction coefficient can be computed from equation (2.3), and the sun background power at the receiver can be computed by using equation (2.10).

If the sun does not fill up the receiver FOV then, the additional background radiation comes from sky radiance. The sky radiance can be approximated by 2×10^{-5} of the sun radiant at the operating frequency [9].

The additional background radiations from surrounding objects and terrain that enter the receiver FOV have to be computed individually. This background radiation varies from location to location; therefore, to be pessimistic, we will approximate these radiations by five times the radiation of 332 K^0 based on the temperature of an object heated by the sun in the summer, which is the period of maximum radiation. We will also assume that this radiation fills up the FOV and is right in front of the receiver.

In addition, we will assume that there is radiation from automobile engines and exhausts at a temperature of 120 C^0 which also fills up the FOV, and is right in front of the receiver. In fact, the radiation from automobile engines and exhausts will not affect the receiver if the receiving automobile is in the lane nearest to the transmitter. As previously mentioned, the magnitude of the thermal current is small and is disregarded in this analysis.

It is worth pointing out that there is a possibility that the sun might be reflect from smooth surfaces, like glass, and enter the FOV of the receiver. If this situation happens then the performance of the receiver will worsen; but this situation can be avoided by properly selecting the location of the transmitting beacon.

Since the transmitting distance from beacon to automobile is not large the effect of scintillation should be minimal; however, along the roadway, automobiles run at fast speeds and cause the air mass around them to move rapidly which tends to increase turbulence. There is no experiment performed regarding this matter. Therefore, to be pessimistic, we will assume that the turbulence level is as high as it can be.

4. Binary PPM Performance Analysis

In this study, we decided to use Binary PPM (BPPM) as the format for transmitting data based on the fact that it is easy to implement compared with other transmission schemes. To compute the detection error for BPPM, it is assumed that symbol synchronization exists, i.e. the receiver has a subsystem which continuously locates the symbol locations. Then, the observed process can be modeled as a Poisson process and the sufficient statistic is the number of phtoelectrons counted in each slot in a symbol. The slot in which there is signal present has an intensity $\lambda_s(t)e^{2\chi(t)} + \lambda_n(t)$ in the case of a turbulent channel and $\lambda_n(t)$ otherwise.

The signal and noise intensities can be obtained from the received signal and noise powers derived previously. The relation between power and intensity is shown by equation (2.11).

To find the probability of error, let S be the transmitted symbol, R be the received signal, N_0 and N_1 be the number of photoelectrons counted in the 1^{st} and 2^{nd} slots of a symbol respectively, and T' be the slot duration. Then

$$Pr(\text{error}) = Pr(R = 0|S = 1)Pr(S = 1) + Pr(R = 1|S = 0)Pr(S = 0). \tag{2.15}$$

In BPPM, the probabilities of transmitting a 0 or a 1 are equal. Hence

$$\begin{aligned}
Pr(\text{error}) &= Pr(\text{error}|S = 0) = Pr(\text{error}|S = 1) \\
&= \int_{-\infty}^{+\infty} Pr(error|S = 0, \chi(t) = x) f_{\chi(t)}(x)dx \\
&= \int_{-\infty}^{+\infty} \{ Pr(N_1 > N_0|S = 0, \chi(t) = x) \\
&\quad + \frac{1}{2} Pr(N_1 > N_0|S = 0, \chi(t) = x)\} f_{\chi(t)}(x)dx \\
&= \int_{-\infty}^{+\infty} \{ \sum_{n=0}^{\infty} \{ Pr(N_1 > N_0|N_0 = n, S = 0, \chi(t) = x) \\
&\quad + \frac{1}{2} Pr(N_1 = N_0|N_0 = n, S = 0, \chi(t) = x)\} \\
&\quad . Pr(N_0 = n|S = 0, \chi(t) = x)\} f_{\chi(t)}(x)dx \\
&= \int_{-\infty}^{+\infty} \{ \sum_{n=0}^{\infty} \{ \{ \sum_{j=n+1}^{\infty} Pr(N_1 = j|N_0 = n, S = 0, \chi(t) = x)\} \\
&\quad + \frac{1}{2} Pr(N_1 = n|N_0 = n, S = 0, \chi(t) = x)\} \\
&\quad . Pr(N_0 = n|S = 0, \chi(t) = x)\} f_{\chi(t)}(x)dx.
\end{aligned} \tag{2.16}$$

Due to T' being small, it can be assumed that $\chi(t)$ is constant during that time duration. Then the probability of error for BPPM is

$$
\begin{aligned}
Pr(error) \;=\; & \int_{-\infty}^{+\infty} \Big\{ \sum_{n=0}^{\infty} \Big\{ \Big\{ \sum_{j=n+1}^{\infty} \frac{[(\lambda_s e^{2x} + \lambda_n)T']^j}{j!} e^{-(\lambda_s e^{2x}+\lambda_n)T'} \Big\} \\
& + \tfrac{1}{2} \frac{[(\lambda_s e^{2x}+\lambda_n)T']^n}{n!} e^{-(\lambda_s e^{2x}+\lambda_n)T'} \Big\} \\
& \cdot \frac{(\lambda_n T')^n}{n!} e^{-\lambda_n T'} \Big\} \; f_{x(t)}(x) d_x .
\end{aligned}
\qquad (2.17)
$$

5. Computational Results

In a practical system, there are some energy losses in the modulator and transmitting optical system, and also some losses at the receiving optical system [4]. In this study, we will assume both transmitter and receiver transmissivities to be 80%. The quantum efficiency of the photodetector at the present time is at 60% or better. To be pessimistic, we will choose the minimum value of 60%.

As previously mentioned, to obtain the optimum transmission zone for any transmission planar angle, the receiver planar angle should equal the transmitted signal angle. The transmission zone increases as the transmission angle increases; however the background radiation increases as well. If the transmitted power remains fixed, and the transmission angle as well as receiver FOV increases, then there is a point at which the background radiation will overcome the transmitted signal. To overcome the background radiation, the transmission power has to be increased. Hence, if we want to have a wide transmission zone, we need to transmit the infrared signal at a wider transmission angle and at a higher transmission power. Therefore, there is a trade off between transmitted signal power and transmitted angle as well as receiver FOV.

To verify the performance of the proposed system, we use equation (2.17) to compute the probability of detection error by fixing the transmission power and increasing the transmitted angle one degree at a time. For each transmission angle (which is the same as the receiver FOV), we use the computer to compute the probability of detection error by varying the transmission rate until the probability of detection error for the last lane is approximately 10^{-6}. The total number of information bits that an automobile running in the lane nearest to the transmitter should receive as it traverses through the transmission zone is computed. The computed results are plotted in Figure 14 and 15 for a fixed transmission power of three and five milliwatts respectively. The vertical axis shows the computed number of received bits corresponding to each of the transmission angles (horizontal axis). The transmission rate for each angle is also labeled accordingly. The probability of detection error for the last lane is 10^{-6}. Because the transmitted infrared beam is widely divergent, an automobile running in the middle lane can receive a higher number of transmitted bits under the same transmitted angle than those plotted on the graphs for an automobile running in the nearest lane, and the number of received bits for an automobile running in the last lane can be even higher. Therefore, automobiles running in the first lane will received minimum transmission, with minimum detection error, while the last lane will received maximum number of transmitted bits with the highest probability of detection error.

It is worth pointing out that if the transmitted power is increased, the number of received bits for each transmitted angle will increase as well. This can be verified by comparing Figures 14 and 15.

At the present time, available semiconductor laser can operate at the maximum speed of 10^{12} bps. Hence, all the plotted numbers shown in the graphs should be achievable.

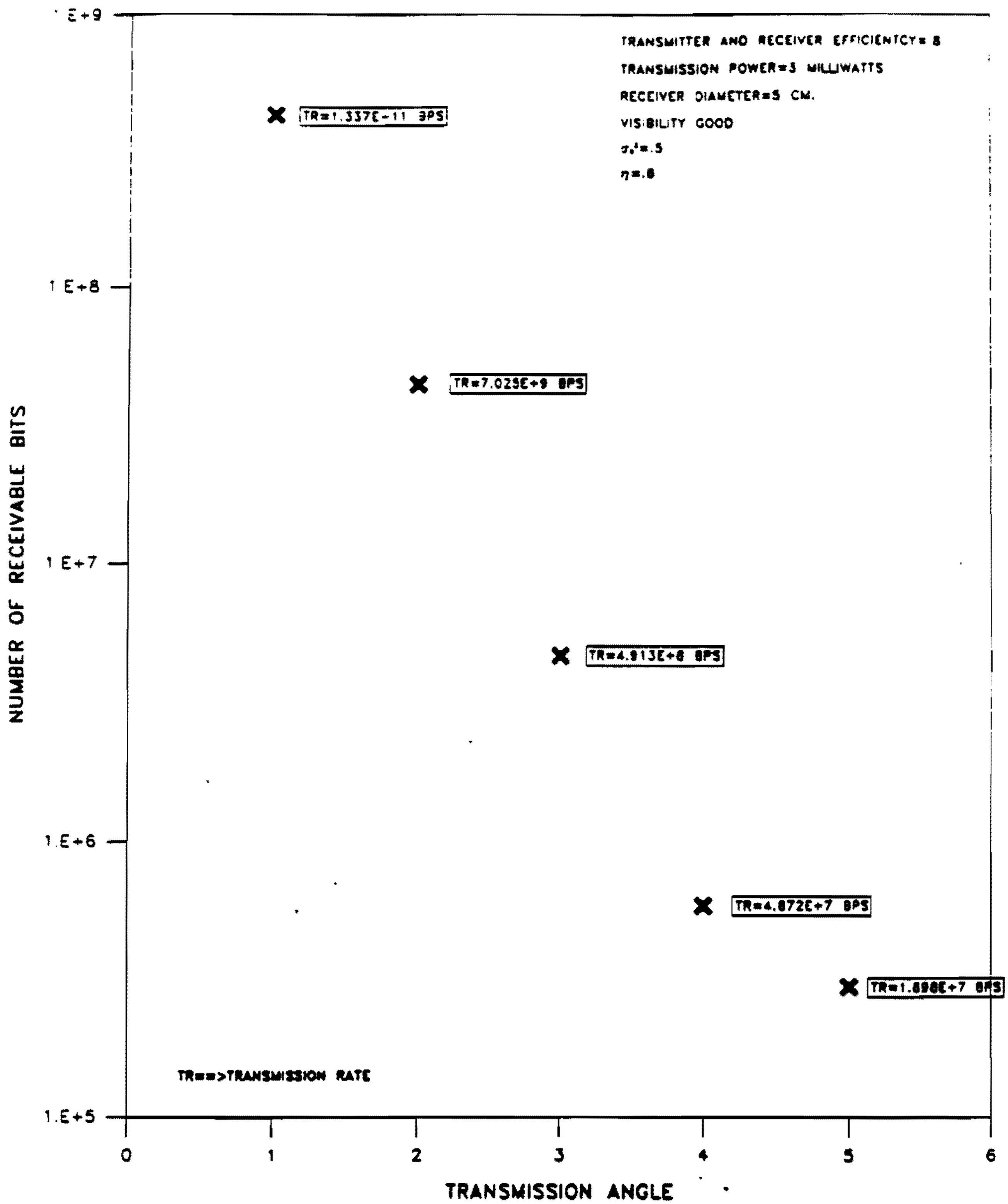


Fig. 14. Plot of transmission angle vs. minimum number of receivable bits, with transmission ower of 3 milliwatts.

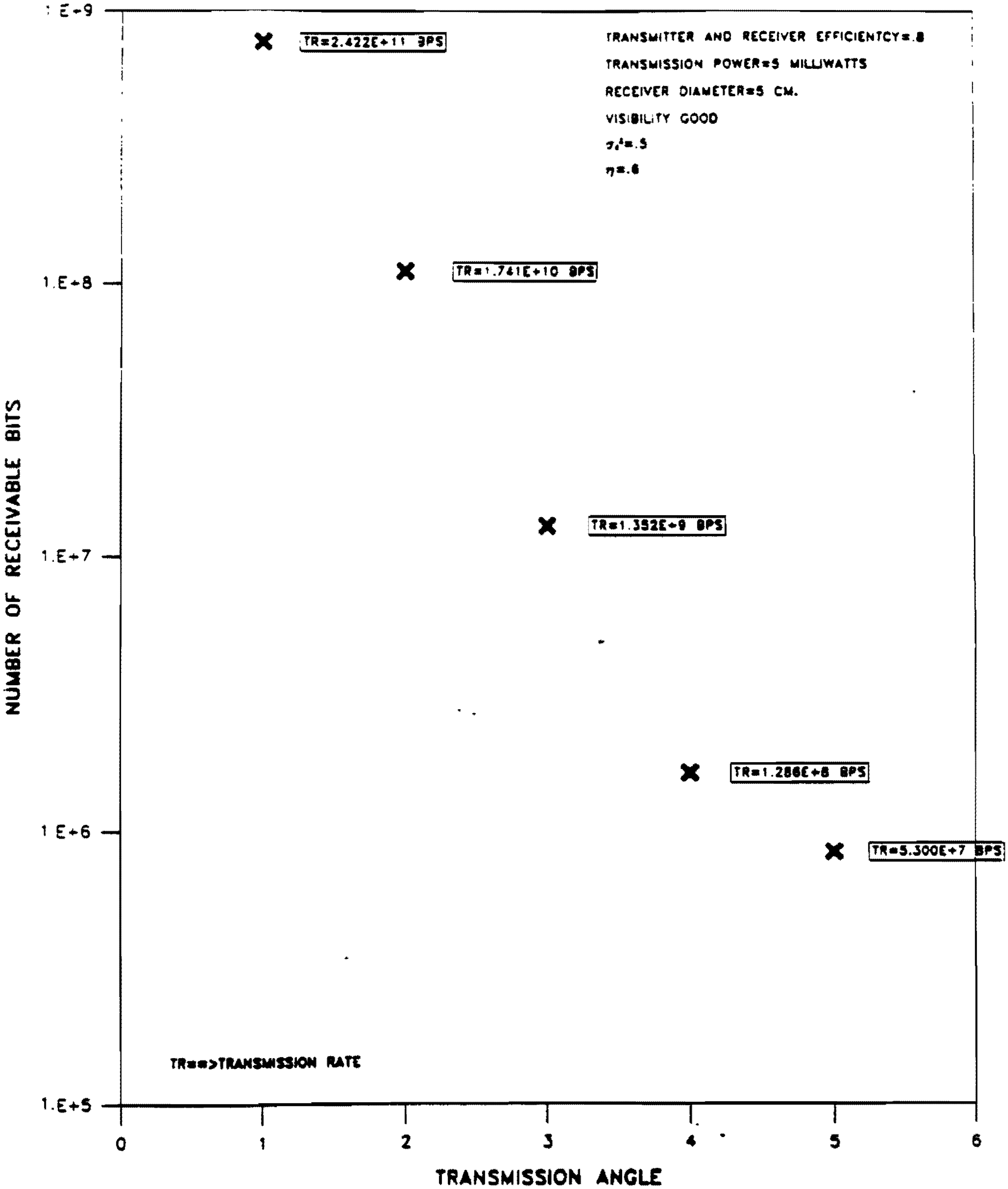


Fig. 15. Plot of transmission angle vs. minimum number of receivable bits, with transmission power of 5 milliwatts.

E. Conclusions

We have proposed an infrared wide-divergence beam roadside communication system and discussed some theoretical background relating to the performance of the proposed system. We have also presented computational results on the performance of the system at various transmission angles and transmitted powers. We would like to caution that the theoretical results which we have presented are based on some proposed constraints and on worst-case conditions. There may be some unforeseen circumstances which may hinder the performance of the system. Therefore, these presented results should be considered as a theoretical possibility; but the actual performance has to be confirmed through an experimental system.

REFERENCES

[1] R.E. Middleton, *Vision Through the Atmosphere*, University of Toronto Press, 1952.

[2] P.W. Kruse and L.D. McGlauchlin and R.B. McQuistan, *Infrared Technology*, New York: John Wiley & Sons Inc., 1962.

[3] S. Karp et.al., *Optical Channels*, New York: Plenum Press, 1988.

[4] W.K. Pratt, *Laser Communication Systems*, New York: John Wiley & Sons Inc., 1969.

[5] H.T. Yura, "Physical Model for Strong Optical-amplitude Fluctuations in a Turbulent Medium ," *Journal of the Optical Society of America*, Vol. 64, No. 1, p. 59-67, January 1974.

[6] N.S. Kopeika and J. Bordogna, "Background Noise in Optical Communication Systems ," *Proceedings of the IEEE*, Vol. 58, No. 10, p. 1571-1577, October 1970.

[7] R.H. Kingston, *Detection of Optical and Infrared Radiation*, New York: Springer-Verlag, 1978.

[8] H.R. Zwicker, "Photoemissive Detector ," *Optical and Infrared Detector*, R.J. Keyes editor, Topic in Applied Physics Vol 19, New York: Springer-Verlag, 1980.

[9] E.E. Bell et.al., "Spectral Radiance of Sky and Terrain at Wavelengths between 1 and 20 Microns. II. Sky Measurements ," *Journal of the Optical Society of America*, , Vol. 50, No. 12, p. 1313-1320, December 1960.

[10] P.R. Gast, et.al., "Solar Irradiance ," *Handbook of Geophysics and Space Environments*, S.L. Valley editor, Cambridge, MA: Air Force Cambridge Research Lab, Section 16, 1965.

[11] R.W. Boyd, *Radiometry and the Detection of Optical Radiation*, New York: John Wiley & Sons Inc., 1983.

[12] R.D. Hudson Jr., *Infrared System Engineering*, New York: John Wiley & Sons Inc., 1969.

[13] W.L. Wolfe and G.J. Zissis, *The Infrared Handbook*, prepared by the Environmental Research Institute of Michigan for the Office of Naval Research, Department of the Navy, Washington, DC, 1978.

[14] F.M. Davidson and Y.T. Koh, "Interleaved Convolutional Coding for the Turbulent Atmospheric Optical Communication Channel ," *IEEE Transactions on Communications*, , Vol. 36, No. 9, p. 993-1003, September 1988.

CHAPTER III

AUTOMOBILE POSITIONING

A. Introduction

Since the dawn of the automobile era, there has been a need for automobile positioning systems. Position can be obtained easily if one drives in familiar areas; however, if one drives in new surroundings or in uncharted areas, to obtain positioning may not be easy. Therefore, an automobile positioning system becomes necessary in these situations. Real time positioning systems are also needed in automobile route guidance systems which provide real-time route guidance through congested areas. There are many systems for obtaining automobile position; in this study, we will classify those methods and discuss their advantages and disadvantages.

B. Automobile Positioning Systems

Automobile positioning can be classified roughly into four methods.

1. Dead Reckoning
2. Radio Navigation
3. Proximity System
4. Map-Matching

1. Dead Reckoning

Dead reckoning is the method of obtaining a vehicle's position through computing the relative movement of an automobile from a known starting location, heading, and distance traveled. The main equipments used by this method are a magnetic compass to find heading and an odometer to obtain the distance traveled. This method of positioning accumulates a distance error and therefore requires reinitialization intermittently. More accurate systems may include a differential odometer which is used to measure relative heading changes, a digital flux-gate compass which is a new type of compass that can compensate for and reduce the effects of anomalies of earth magnetic fields and heading-dependent induced fields caused by the automobile itself, and a yaw rate sensor which is typically a conventional rate gyro, etc. [1, 2, 3].

Advantages

- Inexpensive.
- Easy to implement.

Disadvantages

- Accumulation of errors from distance measurement. Needs to be reinitialized periodically.
- Magnetic abnormalities cause error in heading.

2. Radio Navigation

There are two types of radio navigation:

1. Ground Based.
2. Space Based.

Ground based systems, in general, employ the concept of time or phase difference between two radio signals from two different radio transmitters to obtain the locus of all possible positions with the same time difference. Therefore, to compute the location, three different radio transmitters are needed to generate the intersection between two locus lines, which determine the location of the receiver. There are many ground based radio systems, for example, Omega, Decca, and Loran-C. Only Loran-C is potentially useful for automobile navigation in the United States. Loran-C can provide approximate location accuracy to within 660 ft. At present, Texas does not have complete coverage [4, 5].

Advantages

- Does not need to know the starting point like dead reckoning.
- Location can be updated in a matter of seconds.

Disadvantages

- At present, no complete coverage in Texas.
- Performance affected by multipath propagation, interference, and terrain obstruction.

Space based systems, in general, use high frequencies to transmit coded signals from satellites orbiting in space for positioning. Each satellite continuously broadcasts its unique pseudo-random code. To obtain location, the receiver employs the concept of time delay of signals transmitted from the satellites to compute the distance from satellite to the receiver. Note that the location of the satellite is known to the receiver at all times. Due to the absence of synchronization between the receiver and satellite clocks, the receiver needs four signals from different satellites are needed to identify the location, instead of only three. The first United States satellite navigation system was TRANSIT which eventually will be replaced by GPS (Global Positioning System) which is still not fully operational at the present time. For civilian uses, the average positioning error is within 50 meters [6, 7].

Advantages

- Similar to Loran-C but more accurate.

Disadvantages

- Receiver costs more than that for Loran-C.
- Performance deteriorates because of multipath and obstruction.
- The system is designed for civilian uses but is still under the control of the U.S. Department of Defense.
- Not fully functioning at this time.

3. Proximity System

These systems rely on placing short-distance transmitters, which transmit their location codes along the road and require vehicles to have knowledge of these codes. The transmitter can be a radio with an antenna either at the roadside or an inductive loop buried in the ground. Infrared, microwave, and millimeter wave transmitters are also being used [3, 8, 9, 10].

Advantages

- This system can be incorporated into the road information system and can provide two-way communication.
- Good accuracy can be obtained.

Disadvantages

- High cost of installation due to the number of transmitters to be installed to make the system work.

- Radio interference and bandwidth problems.
- Transmitter and receiver must be in line of sight.

4. Map-Matching

This technique relies on pattern-recognition techniques based on the movement of a vehicle on a known roadway. The accuracy of this method depends on the information stored in a digitized map data base [11, 12].

Advantages

- Systems mentioned above need map data in some form; therefore, map-matching will augment those systems in positioning.
- Reduces the cost of sensors required by those systems.

Disadvantages

- Expensive due to complex database.
- Hard to keep up with road changes.
- Cannot be used under off-road conditions or when no map data is available.
- Performance depends on the accuracy of the map.

C. Conclusions

We have classified the automobile positioning systems and discussed their advantages and disadvantages. Radio navigation systems, and proximity systems are the suitable systems to be used in real time positioning. In route guidance applications, the combination of a radio navigation, a proximity system and map matching should work best. We would like to caution that at present time radio navigation employing Loran-C does not completely cover Texas, and GPS is not fully operational.

REFERENCES

[1] S.K. Honey and W.B. Zavoli, "A Novel Approach to Automotive Navigation and Map Display ," *IEEE Transaction on Industrial Electronics*, vol. IE-34, No. 1, p. 40-43, February 1987.

[2] M.R. Foster, "Vehicle location and Navigation in The Urban Environment ," 6^{th} *International Conference on Automotive Electronics*, p. 152-155, October 12-15, 1987.

[3] E. Ryerson and R.L. French, "Vehicular Navigation and Route Guidance Systems: Technology and Application ," *Proceedings of the* 57^{th} *Anual Meeting: Institute of Transportation Engineers*, Washington, D.C., p. 203-208, 1987.

[4] R.E. Menick, "Loran-C and the Land-Mobile Environment ," 36^{th} *IEEE Vehicular Technology Conference Record*, Dallas, Texas, p. 381-386, May 20-22, 1986.

[5] R.C. Sagers, "A Loran-C Based Receiver for Automatic Vehicle Location ," 36^{th} *IEEE Vehicular Technology Conference Record*, Dallas, Texas, p. 387-391, May 20-22, 1986.

[6] R.A. Dork, "Satellite Navigation Systems for Land Vehicles ," *IEEE Aerospace Electronic Systems Magazine*, vol. 2, no. 5, p. 2-5, May 1987.

[7] T. Itoh et. al., "Navigation Systems Using GPS for Vehicles ," *SAE Technical Paper Series*, SAE-861360, p. 1-13, September 22-25, 1986.

[8] R. Fukui and Y. Noji and M. Hashizume, "Individual Communication Function of RACS: Road Automobile Communication System ," *1989 IEEE Vehicle Navigation and Information Systems Conference Record*, Toronto, Ontario, Canada, p. 206-213, September 11-13, 1989.

[9] A.J.R. Pawley, "In-Vehicle Information Systems Using Roadside beacons ," 6^{th} *International Conference on Automotive Electronics*, p. 149-151, October 12-15, 1987.

[10] M. Shibata, "Development of a Road/Automobile Communication System ," *Transportation Research*, vol. 23 A, no. 1, 1989.

[11] R.L. French, "Automobile Navigation: Where Is It Going? ," *IEEE AES Magazine*, p. 6-12, May 1987.

[12] C.B. Harris et. al., "Digital Map Dependent Functions of Automatic Vehicular Location Systems ," *1988 IEEE Position Location and Navigation Symposium Record*, p. 79-87, 1988.

æ

CHAPTER IV

OPTICAL FIBER COMMUNICATIONS

A. Introduction

We start this chapter with a brief introduction in an attempt to expose and clarify the major concepts behind optical fiber communications. Before the adoption of the fiber link as a valid medium for data transmission, communication at optical frequencies was restricted mostly to unguided channels such as the free-space channel and the atmospheric one. However, clouds, rain and atmospheric turbulence (resulting in distortion and scattering) adversely affected the reliability of optical communication systems. Thus the need arose for a more convenient way to transmit the information, yet keeping the advantages associated with communicating at optical frequencies. Fortunately, an alternative was soon available and optical communications via a guided dielectric was proposed in 1966 by Kao and Hockham.

The fiber was viewed as a potential replacement for the conventional coaxial cable. Upon the fiber's early appearance, difficulties were encountered, the most damaging one being the unacceptably high attenuation involved. Nevertheless, within the space of a few years, optical fiber communication became a commercial reality, with the advent of the virtually lossless glass fibers. An important development in the evolution of optical fibers is the increasing interest in automotive communications applications. Fiber optics has been in use in automobiles since the late 1960s. Its primary uses have been illumination and sensors. The Toyota Century (1982) used fiber optics to communicate information between the control modules located in the doors and a central processing unit. The Mitsubishi Debonair also used fiber optics to transmit switch information from a rear seat remote radio control unit to the radio chassis. Since the late 1970s, the number of microprocessor based systems in automobiles has increased substantially, thus rendering the area of on board data communications more accessible.

In this study, we investigate the general communications requirements in an Intelligent Vehicle Highway System (IVHS) environment. The ultimate goal is to establish a reliable flow of information both from and to passing vehicles. The services that the traffic management center could provide to the drivers include signals that allow an on-board processor to estimate spatial positioning, messages related to travel conditions in the neighbourhood of the vehicle, etc. Reports may be also generated by the network concerning safety warning (collision, skid,) night vision, convoy aid, route/area safety, speed advisory, an estimate of travel time, knowledge of conditions not visible to drivers, hotel information, parking information or any other general information service.

Communication can also take place from the vehicle to the network. On the driver's request, the system may be able to plan a route for the vehicle using real-time traffic information. The driver may wish to transmit requests for emergency service, specifying the nature of the problem, his location and what kind of assistance is sought. By transmitting the geographic location, the type and possibly the destination of the vehicles, cooperative drivers will help the traffic management center to better predict traffic conditions, by using the vehicles in question as traffic probes for various road measurements.

Concerning the physical realization of such a network, there are many possible architectures that could be implemented in an automotive system. In this study, we propose the installation of a fiber optic network under major highways. Communication of informational data will occur from one point in the fiber link to another destination point. The information obtained from vehicles as well as from other sources (accident reports or traffic information from the traffic management center) will be distributed through this fiber optic network.

It is crucial that the traffic management center be in position to process, compile and distribute the data in real-time. An immediate and easy way must be provided to process the continuous stream of traffic flow data. For this matter, the use of beacons would be in order. Road-side beacons can be conveniently placed on the side of the highway to establish communications to and from passing vehicles. The information provided by each of these beacons will be further integrated into the transmitted traffic information stream.

The driver interface will be a combination of displays, a synthesized voice, detailed maps, information about locations buildings, parks and other public accommodations. In the following sections, the fiber network system is approached in a more quantitative way. The different parameters that determine the data rates that the system can accommodate are considered in detail. A program is used (and appended) to compute these rates and some values actually achieved through experimentation are disclosed.

B. The Communication System

Like any other communication system, the optical fiber one consists of a transmitter, a channel over which information is carried and a receiver at the destination. At the sending end, we have generally an analog message signal that we seek to faithfully transmit. Because of efficiency considerations, digital transmission of information is more suitable for our fiber optic system than its analog counterpart. Therefore, before entering the modulator, the original information source signal is mapped onto an electrical digital one. The modulator then converts the signal available at its input into a suitable form for propagation over the transmission medium. This is usually achieved by modulating a carrier from the optical band of the electromagnetic spectrum. Some possible modulation schemes are OOK, PPM, PAM, BPSK and FSK (these formats will be defined and studied in greater detail in the following sections).

We note that when a laser is used as our light source (which is often the case,) a laser drive circuitry is included to trigger the laser according to the adopted modulation format. Hence, a digital optical signal is launched into the optical fiber cable.

An optical fiber is an extremely thin waveguide in that it confines light rays and guides them down its length, much like the metallic conductors that carry electric signals in the conventional coaxial cable. The basic optical fiber construction consists of an inner core made of high-quality glass (or plastic) over which the optical beam can be propagated since the glass exhibits a minimal attenuation to the optical wave transmission. Mechanical strength, shielding and core isolation from external perturbations are provided by a cladding layer, usually constructed of similar material but with a different refractive index. Light that is inserted into the fiber follows a straight path until it hits a bend. At this point, it deviates from the core's center towards the core-cladding junction. Due to the relatively small incidence angles, along with the difference in refractive indices between the core and the cladding, the beam is reflected back toward the core's center. This process is repeated until the light reaches the other end of the fiber. We note that, as for any other communication media, losses are incurred. Impurities in the glass, such as iron, copper, etc., usually cause attenuation of the optical fields. On the other hand, signal dispersion within the fiber limits the rate at which information can be transmitted. Both the attenuation coefficient α and the dispersion coefficient β will be dealt with in a later chapter.

As the light beam hits the photosensitive surface at the receiving end, photoelectrons are generated, possibly amplified and processed according to some optimal detection principle to retrieve, after demodulation, the original digitally transmitted signal. A transducer will finally convert it back to its analog, often non electrical form.

The photodetection process will be briefly explained later, including the two types: direct-detection and heterodyne detection. We also note that our choice of the optical receiver in the direct-detection case is the Avalanche Photodetector (APD) whose characteristics are established in a subsequent section.

C. Advantages of Fiber Optic Communication

Here, we emphasize the advantages that a fiber optic communication system would be able to provide and we make some useful comparisons with the fiber main competitor, the conventional coaxial cable.

1. Huge Potential Bandwidth

Using an optical carrier frequency yields a much wider potential transmission bandwidth than metallic cable systems. With fiber optic systems, communication at tens of GHz over many kilometers is possible. Therefore, the capacity to carry information in optical fiber systems is far superior to the copper-based systems where the heavier losses incurred restrict the transmission distance to only a few kilometers at a bandwidth over 100MHz. We note that the bandwidth that is available in fiber optic systems is not fully exploited yet.

2. Size and Weight

Since the fiber core diameter is in the order of microns, we can qualify it as a "thread" of transmission path over which light pulses travel. (In fact, the diameters involved are often comparable to the diameter of a human hair.)

If we consider also the protective coating around the inner cable core, the whole fiber cross section would be in the order of millimeters, which makes the optical fiber a far smaller and much lighter communication channel than the copper cables. The possibility of packing appreciable amounts of modulated data over such a small spatial area is a considerable advantage for the fiber optic implementation.

3. Electrical Isolation

One of the common faults in the automobile is a short circuit to ground. Since optical fibers are made of glass or plastic polymer (dielectric nature) they are electrical insulators. Thus, unlike their metallic competitors, optical fiber cannot be electrically shorted and does not exhibit arcing, sparking or earth loop problems. This means that they are particularly useful in environments of electrically hazardous nature.

4. Interference Immunity

Since an optical fiber is a dielectric waveguide, it is free from three major sources of interference: electromagnetic interference (EMI) (so no shielding is required), radiofrequency interference (RFI) and switching transients giving rise to electromagnetic pulses (EMP). Thus, with fiber optics, the fiber can be routed anywhere without causing a problem in another system or having another system interfere with its operations. This property is often referred to as electromagnetic compatibility (EMC), and its significant importance becomes rapidly obvious if one considers the many electric/electronic systems that are usually packaged in the close neighborhood of an automobile.

Another desirable feature is the near perfect isolation of the fiber from lightning strikes if used overhead (rather than buried under the ground). Add to this the weather immunity that characterizes the fiber and, in fact, the optical link is quite unaffected by rain or fog.

5. Crosstalk Immunity

Unlike communication via electrical conductors with eventual interferences between cables, there is virtually no crosstalk among optical fibers, even when many fibers are bundled together.

6. Low Transmission Loss

The interest raised by fiber optics as an attractive means of communicating energy from one spot to another over the last 20 years resulted in the production of fibers exhibiting very low attenuation or transmission loss in comparison with the most improved copper conductors. Fibers with losses as low as 0.1 dB/km have been fabricated. This translates into the possible implementation of the communication system with a minimum number of repeaters (intermediate electronic circuitry for amplification purposes) thus reducing the complexity of the system as well as the overall cost.

7. Ruggedness and Flexibility

Optical fibers may be manufactured with appreciably high tensile strengths. (It is a fact that a multiple fiber cable is able to withstand more pulling tension than the coaxial cable.) Another surprising feature (for a glassy substance) is that fibers may be bent, or twisted without damage and are much more flexible than their copper-based counterparts. Size, weight, flexibility, compactness and ruggedness make the optical fiber cables superior in matters of storage, transportation, handling and installation to the copper cables with comparable durability.

8. Cost

At present, the optical fiber cable is reasonably competitive with the coaxial one. Although the fiber communication system is estimated to be somewhat more costly, it is often the case that the advantages offered by the optical implementation are deemed to be worth the higher cost.

In general, the initial cost will be the combination of the fiber ends, transmitter-receiver pairs and fiber length. Additional costs will be applied towards assembly, warranty and service costs.

In the future, the expectations are that new technology advancements will render the optical fiber more and more competitive in efficiency and cost, thus attracting a larger user population.

CHAPTER V

DIRECT-DETECTION SYSTEMS

A. Key Parameters in Fiber Optics System Design

In this section, we define the quantities encountered in the appended program (as well as some other important parameters) and explain each one's effects on our fiber optics communication system.

1. Error Probability

A very important quantity in digital communications is the probability of error. Whether the transmission link includes the atmosphere, a transmission line or an optical fiber, attenuation, distortion, dispersion, as well as noise signals generated in the media, are introduced, often causing errors to occur. Plainly speaking, the ultimate objective of any communication system is to be able to transmit the information reliably to its final destination. A communication system that often fails to reproduce at the receiving end a faithful replica of the actually transmitted message, cannot meet the basic user demands. Adopting such a system is a hazardous decision, even risky. Thus, it is quite crucial to reduce the number of error occurrences, or, in more scientific terms, reduce the error probability. In fact, in a broad sense, the transmitter (modulator, encoder) and the receiver (demodulator, decoder) as a pair are specifically designed to combat the hostile effects of the channel on information transmission and consequently improve the performance of the system.

2. Parameter K_s

K_s is the contribution to the average photon count from the information signal. A digitally transmitted optical field will be processed at the receiving end to retrieve the embedded information with a minimal probability of committing a decision error. A receiver that implements such processing of the incoming optical beam is known as an optimal receiver. Often, the optimal receiver has to count the number of photons detected and compare this number (or some related quantity) to some threshold established by the modulation format we choose. Since communication of information is basically a probabilistic process, various probability computations will be performed by the receiver. Since the number of photons received is a random variable, the average count K_s, plays an important role in the decision making process. We note at this point that a higher K_s means more energy is available for detection, and thus the task of the receiver is alleviated in making correct decisions, which translates into a laser error probability. K_s can be increased by increasing the transmitted laser power.

3. Parameter K_b

The contribution to the average photon count from the background noise field is referred to as K_b. The background noise includes all kinds of undesirable optical beams emanating from light sources other than the one sending the information. During the detection process, the corresponding photo-electrons emitted will be processed along with the ones that carry the information, as the receiver has no a priori knowledge to discriminate the noise terms from the signal ones. Since the photodetector has no estimation of how much noise is embedded in the received waveform, it may favor a certain signal when a large part of the energy detected could actually be due to noise.

Such receiver decision would eventually result in errors committed, and the higher the noise level, the more these errors are likely to occur. In other words, for a fixed signal average energy K_s, the probability of error increases (i.e the system performs poorly) as K_b gets larger. If no background noise is present, K_b equals zero. In general, however, the number of photons detected at the receiver is a random variable with mean $K_s + K_b$. Values of K_b would normally lie in the range 1 to 5 photons/pulse, for practical data rates.

4. Thermal Noise

Thermal noise is another spurious disturbance that masks the received signal. Thermal noise is the spontaneous emission of photo-electrons due to thermal interactions at the atomic level between the free electrons and the vibrating ions in a conducting medium. Thus, the inherent thermal energy generates an output flow of electrons even in the absence of incident light. These currents are referred to as dark currents. Since the dark current is a thermally induced phenomenon, it usually can be reduced by detector cooling. As far as the detection process is concerned, thermal noise has a deleterious effect on the performance. In addition to the photo-electrons generated by some unwanted external sources (K_b), the receiver also has to deal with dark currents. Additional uncertainties are introduced as to which photons bear the information and which ones do not. This automatically translates into enhanced chances of making wrong decisions and thus increased error probability. Finally, we note that since the dark current process is basically a thermal one, increasing the ambient receiver temperature will adversely affect the system performance.

5. Optical Field Detection at the Receiving End

At the receiving end, there are basically two possible ways for detecting the incoming optical field: by direct-detection or by heterodyne-detection. The term direct-detection refers to the type of communication where information is conveyed by the energy of the light signal. The optical field is photodetected directly to produce an electrical signal at the output of the photodetector (since an optical detector is a photosensitive surface that reacts to incident light by generating photoelectrons). This signal is further processed to retrieve the embedded information. Finally, we emphasize that in order to be able to detect the photons at the receiver in direct detection systems, only energy detection modulation schemes can be used (OOK, PPM, PAM).

On the other hand, we have the heterodyne (photomixing) receiver. Prior to photodetection, an optical field from a local oscillator is mixed with the incident field to produce a waveform centered at an intermediate frequency (IF), usually a microwave frequency. This waveform is subsequently processed to get the information originally carried by the incident field. Despite the fact that heterodyne systems may outperform direct detection systems in matters of receiver sensitivity and frequency selectivity, they have in general some major drawbacks. In fact, it is still difficult to implement heterodyne receivers, mainly because the locally generated field requires a high tolerance in spatial coherence with the incident wave. We note that phase-shift-keying (PSK) and frequency shift keying (FSK) are two possible modulation schemes for heterodyne systems.

6. Operating Frequency

Another important quantity is the frequency at which the transmitted channel waveform oscillates. When the medium used for data communication is an optical link, the frequency band is one selected from the optical range of the electromagnetic spectrum (10^{12} to 10^{16}Hz). For optical fiber communications, a suitable band within the optical range is one that covers 0.8 *microns to* 1.7 *microns*. In this range, attenuation can be reduced most. We note that a major advantage of optical carrier frequencies over their radiofrequency counterparts is the much greater associated potential bandwidth. Since the rate of signal transmission is in direct relation with the available bandwidth, communication at optical frequencies can accommodate higher data rates.

7. Transmitted Power P_t

Transmitted power is the power that we inject at the input of the fiber link. Although it can be directly related to the other fundamental quantities involved in the optical fiber communication system design, it is independently chosen by the user. There are several light sources (laser, LED, etc.) but the device most commonly used is the laser. It has many advantages and in general is able to provide the necessary amount of energy needed for the proper operation of the system. A typical range for the laser output power suitable for our communication network is 1 mW to about 10 mW (even though power values as high as 150 mW have been occasionally encountered).

8. Modulation Schemes

In a communication system the information transfer is achieved by superimposing or modulating the information onto an electromagnetic wave which acts as a carrier for the information signal. This modulated carrier is then transmitted to the desired destination for reception and the original message is obtained by demodulation. The carrier may be modulated using either analog or digital signaling. In our study, we investigate only digital systems. In this chapter, we concentrate on direct-detection receivers and consider the relevant modulation schemes OOK, M-ary PPM and M-ary

PAM.

On-Off-Keying (OOK): This is a two-level intensity modulation scheme, with one level at zero. Under OOK, a "0" is sent by switching the laser off during the τ second duration of a symbol and a "1" is sent by turning the laser on and sending a constant intensity signal for τ seconds.

Pulse-Position Modulation (PPM): Another modulation scheme of interest is pulse-position modulation. Under PPM, a symbol interval of duration τ seconds is further subdivided into M subintervals (referred to as slots) of equal duration, and exactly one optical pulse is inserted in one of the M slots at a fixed intensity, resulting in an orthogonal signal set.

Pulse-Amplitude Modulation (PAM): Under PAM, the laser is pulsed during a symbol period of fixed length but with different levels of intensity. In general we will have M intensity levels.

9. Parameters Characterizing the Receiving Device

We now consider some of the important quantities that affect directly the performance of the photodetector.

a. Quantum Efficiency, η

The quantum efficiency η is defined as the fraction of incident photons which are absorbed by the photodetector and generate electrons which are collected at the detector terminals. The quantum efficiency of a photosensitive surface is defined as

$$\eta = Detected\ field\ power / Incident\ field\ power.$$

It measures the ability of the photodetector to detect incoming optical power by converting it into a current. η differs from material to material. Most detectors' efficiency value ranges from 0.5 to 0.7.

b. Gain G of the Avalanche Photodetector

From the four basic types of photodetectors, we choose here the avalanche photodetector (APD). The APD is a semiconductor device that utilizes gap doping to trigger further free-electron regeneration from an initial primary electron (avalanching) to produce gain. APDs are small in size. Their useful gain values are limited to the range of 50 to 200. Since the number of regenerated electrons from a primary released photoelectron is random, the gain G of photomultiplying devices is a random variable having a prescribed probability density. The mean gain is a very important quantity that enters all the expressions for conditional probability.

c. Ionization Coefficient γ for APDs

In the structure of the APD, there exists a high field region in which holes and electrons can acquire sufficient energy to excite new electron-hole pairs. This process is known as impact ionization and is the phenomenon that leads to avalanche breakdown in ordinary reverse biased diodes. A quantitative measure of this effect is given by the ionization coefficient γ. Typically, γ has a value of 0.01 to 0.1 for silicon detectors and 0.1 to 0.5 for germanium ones.

d. Load Resistor R_L

A typical photodetector circuit model is one that features a resistive load R_L so that the released electrons from the incident light are caused to flow through the load, producing the detector output voltage $v(t)$. Variations in the field intensity cause similar variations in the number of released electrons, thereby producing corresponding voltage changes across R_L. A typical value for this load is 50 ohms.

e. Temperature T

A crucial factor in determining the error probability of the system is the ambient receiver temperature. As the surrounding temperature increases, the thermal agitation at the molecular and atomic levels is enhanced and more dark currents are produced in the process. This causes a degradation of the overall system performance.

10. Parameters Characterizing the Optical Fiber Link

a. Attenuation Coefficient α

[A very important parameter that determines how much of the launched power will actually reach the different points along the fiber length.] A number of factors are responsible for the signal attenuation within optical fibers, the major ones being the material composition, the chemical purification technique, the waveguide structure, etc. Naturally, we are interested in small fiber attenuation since this will reduce the required power at the fiber input. In fact, optical fiber communications became particularly attractive when the transmission losses of fibers were reduced below those of the competing metallic conductors. Nowadays, fibers with attenuation as low as 0.1 to 0.2 dB/Km are available.

b. Dispersion Coefficient β

In studying propagating guided fields, one is concerned not only with power loss through attenuation but field dispersion as well. Dispersion is a phenomenon characterized by the fact that the fiber core material causes the various frequencies in a mode waveform to propagate at different velocities within the mode (for single mode). For multiple modes, we have another kind of dispersion: two modes traveling at different angles at the same speed will reach a point along the fiber at different times. Dispersion mechanisms within the fiber cause broadening of the transmitted optical pulses as they travel along the link. Each pulse broadens and overlaps with its neighbors, thus creating accurate reception difficulties. This means that more time spacing should be provided between the different pulses to compensate for the widening incurred, and consequently less data rates will be accommodated by the channel. We note that a quantitative measure of dispersion is given by the coefficient β (5E-8 to 5E-10 sec/Km is an appropriate range).

11. Data Transmission Rate R (Throughput)

Throughput is a fundamental quantity of major importance for the communication system engineer. It is always beneficial to be able to communicate at high data transmission rates. A system that requires more time to communicate certain amounts of information is simply less appealing than one that can accommodate higher rates. In general, there are two major limitations that prevent the achievement of arbitrarily high rates. The first one is the power limitation. We cannot expect to accommodate a desirable transmission rate without providing the adequate amount of required power. Once the power injection into the system is fixed, the corresponding upper power limit on R is established. On the other hand, power is not the sole entity that specifies the data rate. Dispersion of the fiber plays an equally limiting role on R. As we argued in the previous section, excessive overlapping between adjacent light pulses (due to accentuated dispersion mechanisms) will eventually cause an appreciable reduction in the possible rates. Thus, even in the presence of an infinite amount of energy (which is not the case anyway) data transmission rates will be upper bounded by the fiber dispersive effects. Finally, since there is no theoretical necessity for the rate threshold due to a power limitation to be higher (or lower) than that due to a dispersion limitation, this threshold is taken, in general, to be the minimum of both quantities.

B. An Example and Some Figures

To illustrate the concepts reviewed above, a numerical example will be given next. We note that a self-explanatory program is included in the appendix. The program guides the user through its steps in a quite clear way. It asks the user to input the required information, hinting to the relevant ranges of values. The final output will inform the user how much data rate the system described can accommodate. The different sources for the equations used in the program can be found in [1].

The following example gives a flavour of the values involved in a practical system:

We are given an Avalanche Photodetector (APD) with the following characteristics:

- Load resistance = 50 Ohm.
- Mean gain = 100.
- Ionization coefficient = 0.028.
- Quantum efficiency = 0.5.

Let the operating frequency be 3E+14 Hz and the ambient temperature at the receiver 300 degrees Kelvin.

As a modulation scheme, we choose the OOK format and we assume a background noise K_b of 1 photon/pulse.

Some of the values specified above, together with a required error probability of, say, 10E-06, yield a value of K_s equal to about 275 photons/pulse.

We still have to specify the power launched at the input end of the fiber. Let this power be 10 mW. The combination of all the quantities encountered so far leads to a number, Q, that takes into account all power parameters. (We remind the reader that there are two fundamental limitations to arbitrarily high data rates: the power limitation and the dispersion one.)

In this example, Q is computed to be 1.8293E+14.

Finally, we need the necessary information about the fiber link. Let our fiber have the following specifications:

- Fiber length: 50 Km.

- Attenuation Coefficient: 0.1 dB/Km.

- Dispersion Coefficient: 5E-10 Sec/Km.

Now that all the relevant quantities are specified, we can determine the following data transmission rates:

R_b = 3.8788E+13 bits/sec (power limitation).

R_b = 282842712 bits/sec (dispersion limitation).

Taking the minimum of the two values, we get the final result:

The system described can accommodate 282842712 bit/sec.

Before moving to the coherent detection, we present some figures that may help clarify the concepts discussed earlier.

We note first that for all these figures, the quantum efficiency is 0.5, the frequency is 3E+14 Hz, the mean gain is 100, the ionization coefficient is 0.028, the load resistance is 50 Ohm and the background noise is 1 photon/pulse.

In Figure 16, we consider the OOK format as our modulation scheme, and we assume the temperature in the vicinity of the receiver to be 300 degrees Kelvin. The graph shows the variation of the quantity Q (described in the previous example) with the transmitted power P_t, for different values of error probability. It is clear that Q is in direct proportion with the power independent of the probability of error. We inform the reader that a higher value of Q translates into a higher data rate, a goal that is normally quite desirable. All this means that the larger the injected power at the fiber input, the larger the associated data transmission rates (if there is no penalty due to the dispersion limitation). We also learn from this figure that a bigger probability of error means a larger Q (and thus a higher rate) for a fixed launched power. This is to be expected because of the tradeoff that exists between how much information could be sent and how reliably this information would be retrieved.

In Figure 17, the error probability is fixed to 10^{-06} and the temperature to 300 degrees Kelvin. the plot shows the response of Q to the transmitted power for the different modulation formats, namely: OOK, binary PPM, 4-ary PPM, 8-ary PPM, 4-ary PAM and 8-ary PAM. Again, Q increases with the power irrespective of the signal modulation. For a fixed P_t, Q (and thus the rate R) is the largest for the 8-ary PPM and smallest for the 8-ary PAM. Despite the increase in the value of Q when the modulation is of the PPM type, we particularly note that higher rates could be achieved with OOK as compared to Binary PPM.

Figure 18 conveys the same information as the previous one concerning the variation of Q with P_t, except now we have an additional variable: the temperature. At any fixed power and any modulation scheme, the value of Q is higher for lower temperatures. This is also to be anticipated since a higher ambient temperature at the receiving side of the fiber causes greater thermal agitation at the molecular level and more dark currents are generated. The excessive presence of dark currents makes the demodulation process more difficult and thus communicating at high rates may not be reliable.

Figure 19 shows the transmission rate R versus the fiber length, with the attenuation coefficient, α, the dispersion coefficient, β, and Q as parameters. As soon as Q is specified (we saw in the example and the three previous figures the relevant quantities that determine Q,) the rate R can be computed, with the help of the three fiber characteristics, α, β and the length. As we noted earlier a higher Q leads to a higher R, for a fixed α. On the other hand, for a fixed Q, the rate decreases substantially with high attenuation coefficients. For $\alpha = 10$ dB/Km, the rate vanishes for about 10 Km of fiber length. The other limitation to arbitrarily high rates is the one due to dispersion. It is clear from the plot that a higher dispersion coefficient beta means a lower rate, irrespective of power considerations. To determine R for a fixed fiber length L, we take the minimum of the two rates, i.e the one that satisfies both the power and the dispersion constraints.

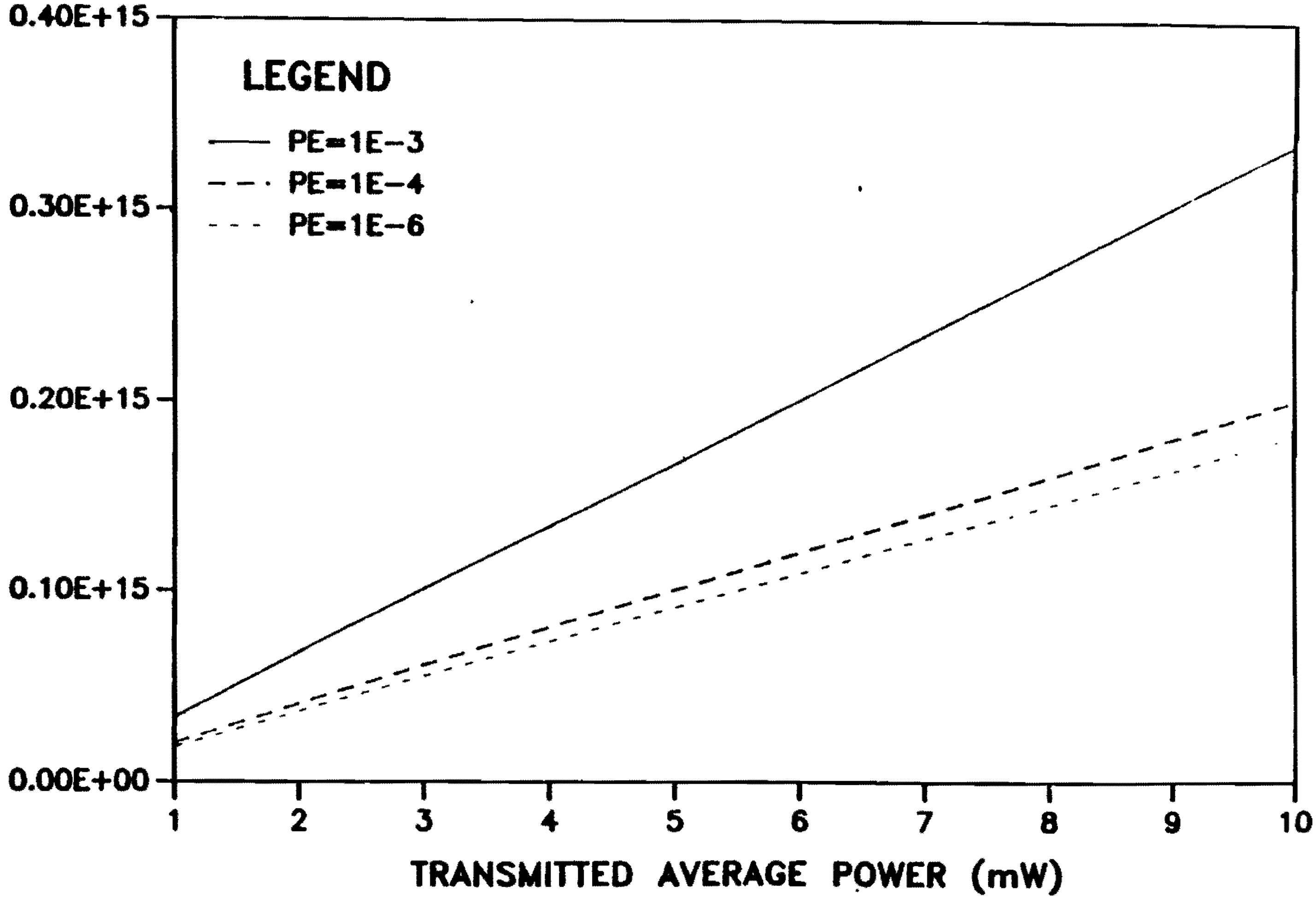


Fig. 16. *Q* function of launched power for OOK (T=300K).

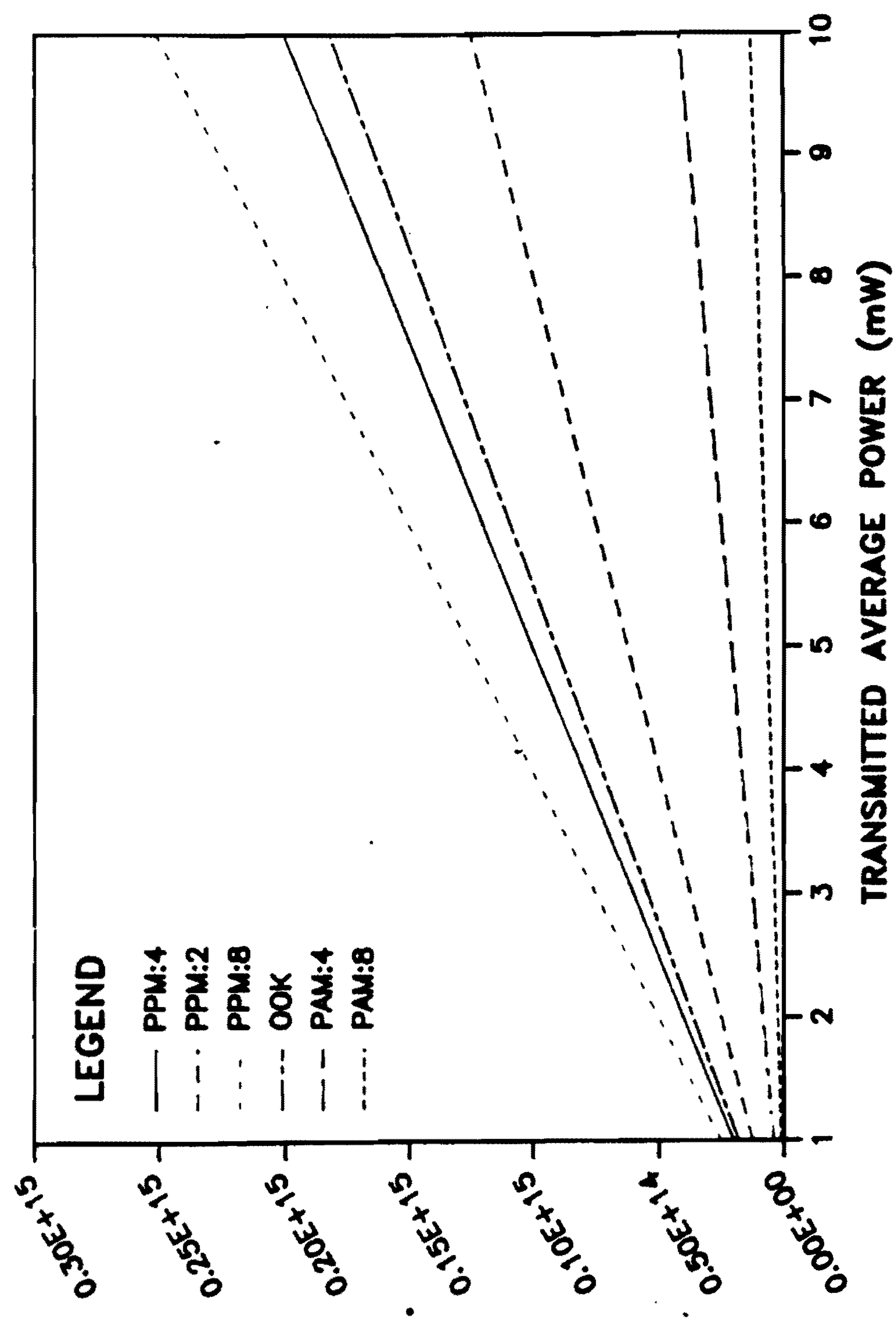


Fig. 17. Q function of launched power for different schemes (PE=1E-6,T=300K).

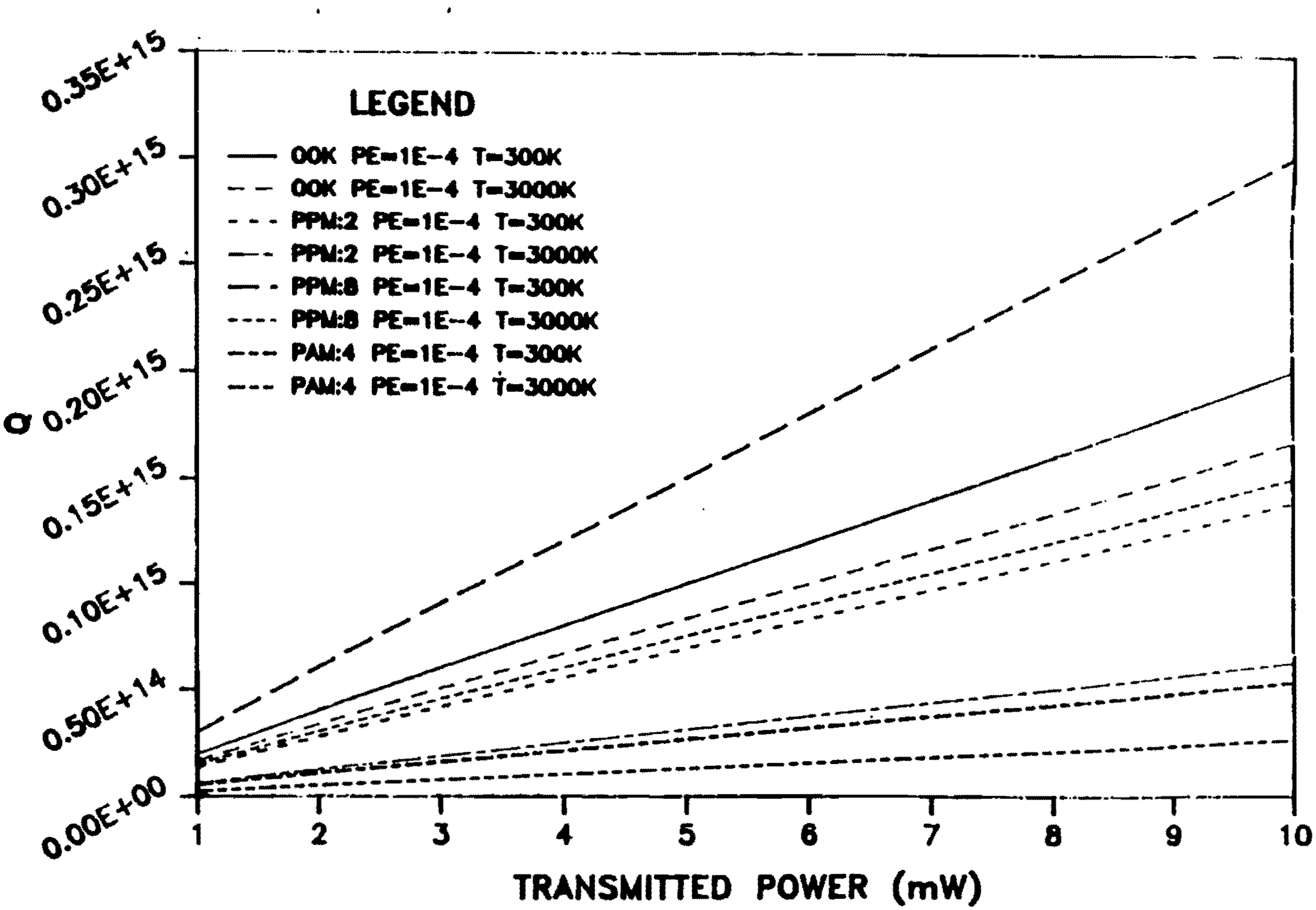


Fig. 18. Q function of launched power for different schemes.

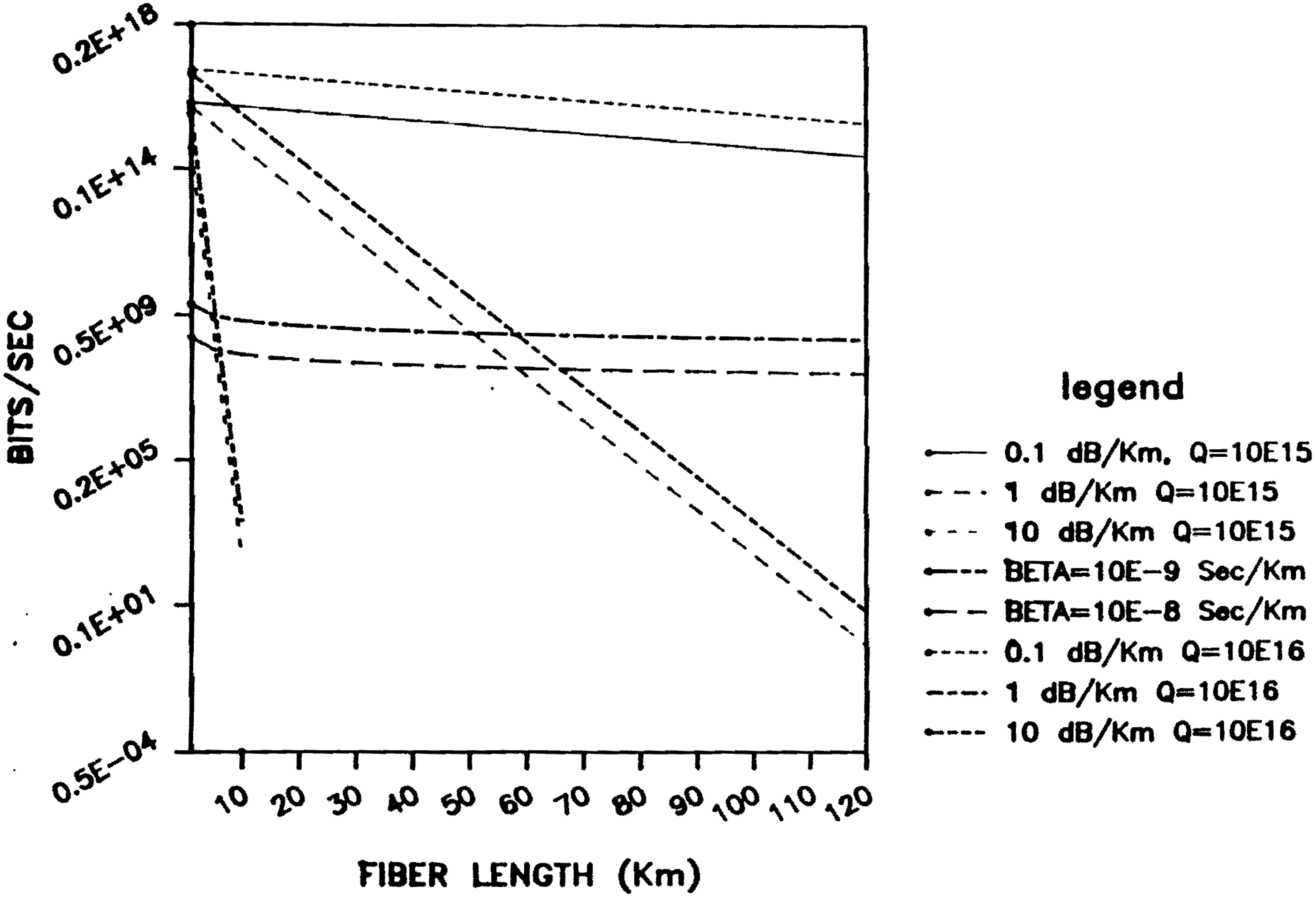


Fig. 19. Bit rate versus fiber length (OOK).

CHAPTER VI

COHERENT HETERODYNE DETECTION

In Chapter V we considered the direct, or noncoherent detection of the received optical field. We found that, for such systems, suitable modulation schemes would be OOK, PPM etc.

In this chapter, we focus on an alternative method for field detection, namely, the heterodyne one. In heterodyne detection, the receiver operates by mixing (heterodyning) the incoming signal with a local optical oscillator, prior to photodetection. Since the optical frequency of the received laser beam is different from the optical frequency of the locally generated laser field, the result of combining these two fields is a signal centered at an intermediate frequency (IF), usually a microwave frequency, which contains the modulating information signal. Standard decoding techniques may be subsequently used by the receiver to process the intermediate frequency waveform in order to extract the embedded information. Since the addition of two electromagnetic fields requires spatial alignment of the fields, the use of heterodyne detection is often called spatial coherent detection.

The most common modulation formats for heterodyne systems are phase-shift-keying (PSK) and frequency-shift-keying (FSK). In binary PSK (BPSK), for example, the carrier is phase shifted between two phase states 180 degrees apart to represent a data one or zero. On the other hand, in binary FSK the carrier is frequency shifted between two frequencies to represent each bit. Thus, in PSK the information is carried by the phase, whereas it is reflected in the different frequency tones if FSK is used.

A. Some Heterodyne Detection Features

Recently we witnessed an increased interest in heterodyne optical communication systems for fiber optic applications. One of the incentives is that coherent detection offers significant improvements (compared with direct-detection) in receiver sensitivity and wavelength selectivity. For a system operating at a certain error probability, this translates into an appreciable saving in transmitter power. We note here that coherent techniques promise to replace the noisy APDs with stable p.i.n photodetectors. (p.i.n photodiodes have a simple structure with a p-n junction and a depleted region). Since a small reverse voltage is needed to bias the diode, the background noise current is negligible. In order to take advantage of the wideband characteristics of the p.i.n photodiode, a combination with a low-noise field-effect transistor (FET) amplifier has been investigated to replace the APD).

Another advantage of the coherent heterodyne systems resides in the fact that while direct-detection systems are generally limited by background radiation and photodetector dark currents, their heterodyne counterparts are almost immune to background noise. In fact, the background radiation is largely noncoherent and thus does not heterodyne significantly with the local laser oscillator beam.

A potentially very important feature of coherent reception is that it allows the use of electronic equalization as a compensation for the deleterious effects of optical pulse dispersion in the fiber.

B. Engineering Challenges Associated With Heterodyne Detection

Several problems arise by virtue of heterodyne detection:

Since both the incident field and the local oscillator beam have spatial extend, proper wavefront alignment at the face of the photodetector is of primordial importance. Any misalignment could cause excess noise to join the output of the photodetector and thus degrade the performance. In practice, it is an extremely difficult task to have ideal heterodyning where the fields are perfectly aligned, and effects of misalignment and local field distortion must be seriously considered.

A major nuisance source that cannot be neglected in the design of a reliable receiver is the phase noise. Laser phase noise is a random process arising from the spontaneous emissions within the laser cavity, which cause the phase of the optical output wave to deviate randomly from the value it would have had in the absence of such emission. Since undesired phase fluctuations in the received wave impair the demodulation process, a poor performance is to be expected in the presence of large phase noise. Thus, it is clear that the phase noise phenomenon sets fundamental limitations on coherent optical communications.

We note finally that in coherent communication systems, the frequency of the transmitter and that of the local laser oscillator must be synchronized within a few gigahertz. This means that the receiver must precisely know the frequency of the distant transmitter, a task that is quite difficult in practice.

C. Some Actual Heterodyne System Implementations

In this section, we expose several of the ideas and improvements that have been suggested and physically implemented in order to take full advantage of the benefits promised by heterodyne systems.

a) A no-repeater fiber transmission experiment over 308 Km is reported [2]. The modulation format is optical continuous-phase FSK (CPFSK). The total injection current is 240 mA and the fiber input power of the signal 8.8 dBm. At a wavelength of $1.554\mu m$, the fiber loss was 54.3 dB. Its dispersion is 1.8ps/Km/nm. The received signal was combined with a local oscillator light by a directional fiber coupler. The total local oscillator power is 6.22 dBm. The intermediate frequency was recorded to be 5.0 GHz. At a sensitivity of 67 photons/bit at 10^{-9} BER (bit error rate) a throughput rate of 2.488 Gbit/s was successfully achieved. We note that, to the knowledge of the designers, the 308 Km distance is longer than any previously reported (as of March 1990) optical transmission length without an optical repeater.

b) Very recently the bit rate of PSK synchronous coherent systems was increased to 2 Gbit/s [3]. At the receiving side, a p.i.n/100 ohm photodiode was used and a 1320 nm diode-pumped miniature Nd:YAG laser as a local oscillator. Estimation or tracking of the phase noise of the incoming optical field (so that it can be used to coherently demodulate PSK) is accomplished at the receiver by an optical phase-locked loop. The phase modulator requires a drive voltage of 5.6 V peak-to-peak for a π phase deviation. The intermediate frequency (IF) was set to 4 GHz. At 2 Gbit/s, after 50 Km of single-mode fiber transmission, the receiver exhibited the sensitivity of 225 photons/bit at 10^{-9} BER.

c) Another recent success is the FSK coherent optical transmission over 2223 Km at 2.5 Gbit/s using Erbium-doped fiber amplifiers [4]. It was shown that 25 repeaters placed at approximately 80 Km intervals offer a total gain of more than 440 dB. Erbium-doped fiber amplifiers offer polarization-insensitive gain and small coupling loss. The low-noise optical balanced receiver consisted of an InGaAs dual-pin photodiode and a HEMT transimpedance preamplifier. The pumping power coupled to the erbium-doped fibers was around 50 mW. The optical fibers used were single-mode fibers whose average loss and dispersion at $1.554\mu m$ were 0.21 dB/km and less than 0.1 ps/km/nm respectively.

d) We report next an experiment where rates up to 8 Gbit/s were achieved

[5]. The modulation scheme is again optical CPFSK and the transmission distance is 202 Km. Here the fiber chromatic dispersion was successfully compensated by a wideband microstrip line delay equalizer and a wideband optical balanced receiver. The experiment was conducted at the wavelength of $1.55\mu m$. The fiber input power of the signal was 8.5 dBm, that of the local oscillator, 7 dBm. The fiber total loss was 39 dB and its dispersion was 16 ps/Km/nm. The IF center frequency was 12 GHz.

e) Finally, recalling from the previous section that accurate synchronization may represent a major practical limitations, we report here the presence of a system that has been designed and tested in the laboratory to combat this difficulty. In fact, it was demonstrated that a $1.5\mu m$ FSK heterodyne detection system can start communication without manual adjustment to synchronize the frequencies of transmitter and local oscillator lasers. It is simple and easy to use [6].

D. Conclusion

To conclude this study, we remind the reader of the typical advantages of fiber optics over its coaxial cable competitor, namely in matters of size, weight, bandwidth, electrical isolation, interference immunity, crosstalk immunity, low transmission loss, ruggedness and flexibility. A potential reduction in cost is anticipated in the future since technology improvements are constantly in progress. All this suggests that future automotive communications needs will be adequately met with the use of fiber optic systems. The choice remains open between direct-detection systems and their coherent heterodyne counterparts. In general, the technology associated with direct-detection systems is relatively less sophisticated, which makes the practical implementation of such systems easier to undertake. However, large benefits are promised by heterodyne systems and the future may tilt accordingly. Recently, many coherent systems have been actually tested and the following results are reported: over a distance of 50 Km of single mode fiber transmission, a rate of 2 Gbit/sec was achieved using PSK modulation. Another successful experiment is one where 2.5 Gbit/s rate was attained using FSK coherent optical transmission. The transmission length is 2223 Km with 25 repeaters placed at 80 Km one from another. We also report a case where a huge rate increase was achieved to transmit 8 Gbit/s with CPFSK over 202 Km.

For a more concrete example, we note that over a distance of about 190 miles (Controlled access Highways Houston-Waco, Houston-San Antonio, Austin-Corpus Christi, Austin-Dallas, Austin-Fortworth, etc.) a throughput of 2.488 Gbit/sec can be sustained at 10^{-9} bit error rate and no optical repeaters.

We finally note that, despite the excellent results obtained so far and that have been illustrated in the examples above, a host of practical difficulties remain associated with the implementation of an ideal heterodyne detection system. The realization of such a system requires further research and development of the optical system components. The potential benefits of these coherent heterodyne fiber optic communication systems ensure that interesting improvements may be expected in the near future. The challenge is big but so is the stimulus to further research.

REFERENCES

[1] S. Karp, R. Gagliardi, S. Moran, and L. Stotts, *Optical Channels,* 1988 Plenum Press, 233 Spring Street, New York, N.Y. 10013

[2] T. Imai, N. Ohkawa, Y. Ichihashi, T. Sugie, and T. Ito. "Over 300 Km CPFSK Transmission Experiment Using 67 photons/bit Sensitivity Receiver at 2.5 Gbit/s," *Electronics letters,* Vol.26, No.6, pp.357-358, March 1990.

[3] D. A. Atlas and L. G. Kazovsky. "2 Gbit/s PSK Heterodyne Communication System Using Optical Phase-Locked Loop ," *Electronics letters,* Vol.26, No.14, pp.1030-1032,July 1990.

[4] S. Saito, T. Imai, T. Sugie, N. Ohkawa, Y. Ichihashi, and T. Ito. "Coherent Transmission Experiment Over 2223 Km at 2.5 Gbit/s Using Erbium-Doped Fiber Amplifiers ," *Electronics letters,* May 1990, Vol.26, No.10, pp.669-671,May 1990.

[5] N. Takachio, K. Iwashita, and K. Nakanishi. "8 Gbit/s 202 Km Optical CPFSK Transmission Experiment Using 1.3μm Zero Dispersion Fiber ," *Electronics letters,* Vol.26, No.8, pp.506-508, April 1990.

[6] Y. C. Chung and R. W. Tkach. "Cold-Start 1.5μm FSK Hetrodyne Detection Experiment ," 1989 Optical Fiber Communication Conference.

APPENDIX B

SIMULATION PROGRAM CODE

```
C This program computes the data rates that a fiber
C  optic direct-detection system can accommodate when all
C the relevant parameters are specified.
C The receiver features an avalanche photodetector (APD) with the
C following characteristics:
C -Load Resistance         : 50 ohms.
C -Mean Gain               : 100
C -Ionization Coefficient : 0.028
C The background noise is taken to be 1 photon/pulse.
C We note that all of these values are practical ones .

C       ***********************************************************
IMPLICIT REAL *8 (A-H,O-Z)
REAL *8 S(8)
REAL *8 KS
10 H=6.626D-34
WRITE(*,*) 'DESIRED MODULATION SCHEME ?'
WRITE(*,*) 'INPUT 1 FOR OOK.'
WRITE(*,*) 'INPUT 2 FOR BINARY PPM '
WRITE(*,*) 'INPUT 3 FOR 4-ARY PPM '
WRITE(*,*) 'INPUT 4 FOR 8-ARY PPM '
WRITE(*,*) 'INPUT 5 FOR 4-ARY PAM '
WRITE(*,*) 'INPUT 6 FOR 8-ARY PAM '
READ(*,*) N
WRITE(*,*) 'DESIRED ERROR PROBABILITY ?'
WRITE(*,*) 'INPUT 3 FOR ERROR PROBABILITY = 10**-3 '
WRITE(*,*) 'INPUT 4 FOR ERROR PROBABILITY = 10**-4 '
WRITE(*,*) 'INPUT 6 FOR ERROR PROBABILITY = 10**-6 '
READ(*,*) M
WRITE(*,*) 'AMBIENT TEMPERATURE ?'
WRITE(*,*)'INPUT 1 FOR TEMP CLOSER TO 300 DEGREES KELVIN THAN TO 3000 '
WRITE(*,*)'INPUT 2 FOR TEMP CLOSER TO 3000 DEGREES KELVIN THAN TO 300 '
READ(*,*) I
IF(N.EQ.1) GOTO 100
IF(N.EQ.2) GOTO 200
IF(N.EQ.3) GOTO 300
IF(N.EQ.4) GOTO 400
IF(N.EQ.5) GOTO 500
IF(N.EQ.6) GOTO 600
100 IF(M.EQ.3.AND.I.EQ.1) KS=150.0D0
IF(M.EQ.4.AND.I.EQ.1) KS=250.0D0
IF(M.EQ.6.AND.I.EQ.1) KS=275.0D0
IF(M.EQ.3.AND.I.EQ.2) KS=200.0D0
```

```
IF(M.EQ.4.AND.I.EQ.2) KS=300.0D0
IF(M.EQ.6.AND.I.EQ.2) KS=500.0D0
GOTO 1000
200 IF(M.EQ.3.AND.I.EQ.1) KS=150.0D0
IF(M.EQ.4.AND.I.EQ.1) KS=180.0D0
IF(M.EQ.6.AND.I.EQ.1) KS=200.0D0
IF(M.EQ.3.AND.I.EQ.2) KS=300.0D0
IF(M.EQ.4.AND.I.EQ.2) KS=400.0D0
IF(M.EQ.6.AND.I.EQ.2) KS=550.0D0
GOTO 1000
300 IF(M.EQ.3.AND.I.EQ.1) KS=185.0D0
IF(M.EQ.4.AND.I.EQ.1) KS=220.0D0
IF(M.EQ.6.AND.I.EQ.1) KS=250.0D0
IF(M.EQ.3.AND.I.EQ.2) KS=400.0D0
IF(M.EQ.4.AND.I.EQ.2) KS=480.0D0
IF(M.EQ.6.AND.I.EQ.2) KS=700.0D0
GOTO 1000
400 IF(M.EQ.3.AND.I.EQ.1) KS=200.0D0
IF(M.EQ.4.AND.I.EQ.1) KS=250.0D0
IF(M.EQ.6.AND.I.EQ.1) KS=300.0D0
IF(M.EQ.3.AND.I.EQ.2) KS=450.0D0
IF(M.EQ.4.AND.I.EQ.2) KS=500.0D0
IF(M.EQ.6.AND.I.EQ.2) KS=750.0D0
GOTO 1000
500 IF(M.EQ.3.AND.I.EQ.1) KS=1000.0D0
IF(M.EQ.4.AND.I.EQ.1) KS=1500.0D0
IF(M.EQ.6.AND.I.EQ.1) KS=1900.0D0
IF(M.EQ.3.AND.I.EQ.2) KS=2000.0D0
IF(M.EQ.4.AND.I.EQ.2) KS=3000.0D0
IF(M.EQ.6.AND.I.EQ.2) KS=4300.0D0
GOTO 1000
600 IF(M.EQ.3.AND.I.EQ.1) KS=6800.0D0
IF(M.EQ.4.AND.I.EQ.1) KS=8500.0D0
IF(M.EQ.6.AND.I.EQ.1) KS=10000.0D0
IF(M.EQ.3.AND.I.EQ.2) KS=10000.0D0
IF(M.EQ.4.AND.I.EQ.2) KS=13000.0D0
IF(M.EQ.6.AND.I.EQ.2) KS=16000.0D0
1000    IF(N.EQ.5.OR.N.EQ.6) GOTO 1100
WRITE(*,*) 'THE CORRESPONDING VALUE OF KS =',KS
GOTO 1500
1100 IF(N.EQ.6) GOTO 1200
WRITE(*,*) 'WE HAVE 4 INTENSITY LEVELS'
WRITE(*,*) 'THE PEAK INTENSITY LEVEL =',KS
WRITE(*,*) 'THE OPTIMAL CHOICE FOR THE 3 REMAINING LEVELS IS:'
DO 1150 J=1,3
S(J)=(((J-1D0)/3D0)**(2D0))*KS
WRITE(*,*) J,'           ',S(J)
1150 CONTINUE
GOTO 1500
1200 WRITE(*,*) 'WE HAVE 8 INTENSITY LEVELS'
WRITE(*,*) 'THE PEAK INTENSITY LEVEL =',KS
```

```
WRITE(*,*) 'THE OPTIMAL CHOICE FOR THE 7 REMAINING LEVELS IS:'
DO 1250 J=1,7
S(J)=(((J-1D0)/7D0)**(2D0))*KS
WRITE(*,*) J,'            ',S(J)
1250 CONTINUE
1500    WRITE(*,*) 'INPUT THE OPERATING FREQUENCY IN Hz'
        WRITE(*,*) 'A SUITABLE RANGE: 1.765D14 Hz TO 3.75D14 Hz'
READ(*,*) F
        WRITE(*,*) 'INPUT THE RECEIVER EFFICIENCY'
        WRITE(*,*) 'A SUITABLE RANGE: 0.5 TO 0.7'
READ(*,*) ETA
        WRITE(*,*) 'INPUT THE TRANSMITTED POWER IN mW'
        WRITE(*,*) 'A SUITABLE RANGE: 1 mW to 10 mW'
READ(*,*) PO
IF(N.EQ.1)  GOTO 6543
IF(N.EQ.2)  GOTO 6544
IF(N.EQ.3)  GOTO 6545
IF(N.EQ.4)  GOTO 6546
IF(N.EQ.5)  GOTO 6547
IF(N.EQ.6)  GOTO 6548
6543    Q=2.0D0*ETA*PO*(10.0D0**(-3.0D0))/(H*F*KS)
WRITE(*,*) 'Q=',Q
GOTO 7000
6544 Q=ETA*PO*(10**(-3.))/(H*F*KS)
WRITE(*,*) 'Q=',Q
GOTO 7000
6545 Q=2.0D0*ETA*PO*(10**(-3.))/(H*F*KS)
WRITE(*,*) 'Q=',Q
GOTO 7000
6546 Q=3.0D0*ETA*PO*(10**(-3.))/(H*F*KS)
WRITE(*,*) 'Q=',Q
GOTO 7000
6547 Q=3.2D0*ETA*PO*(10**(-3.))/(H*F*KS)
WRITE(*,*) 'Q=',Q
GOTO 7000
6548 Q=5.33333333D0*ETA*PO*(10**(-3.))/(H*F*KS)
WRITE(*,*) 'Q=',Q
7000 WRITE(*,*) 'INPUT THE FIBER ATTENUATION IN dB/Km '
WRITE(*,*) 'ATTENUATION AS LOW AS 0.1 dB/Km ARE POSSIBLE NOWADAYS'
READ(*,*) AL
WRITE(*,*) 'INPUT THE FIBER DISPERSION IN Sec/Km '
WRITE(*,*) '5E-8 TO 5E-10 Sec/Km MAY BE AN APPROPRIATE RANGE'
READ(*,*) BE
WRITE(*,*) 'INPUT THE LENGTH OF THE FIBER IN Km'
READ(*,*) DIS
RBP=Q*DEXP(-(AL*0.1551D0*2D0*DIS))
WRITE(*,*) 'THE RATE THAT CAN BE ACHIEVED UNDER '
WRITE(*,*) 'THE POWER LIMITATION CONSTRAINT =',RBP
RPD=1D0/(BE*DSQRT(DIS))
IF(N.EQ.1) RBD=RPD
IF(N.EQ.2) RBD=RPD
```

```
IF(N.EQ.3) RBD=2D0*RPD
IF(N.EQ.4) RBD=3D0*RPD
IF(N.EQ.5) RBD=2D0*RPD
IF(N.EQ.6) RBD=3D0*RPD
WRITE(*,*) 'THE RATE THAT CAN BE ACHIEVED UNDER '
WRITE(*,*) 'THE DISPERSION LIMITATION CONSTRAINT =',RBD
IF(RBP.LT.RBD) R=RBP
IF(RBP.GT.RBD) R=RBD
WRITE(*,*) '*********************************************'
WRITE(*,*)' THEREFORE THE SYSTEM YOU DESCRIBED CAN ACCOMMODATE'
WRITE(*,*)' A RATE EQUAL TO ', R, 'BIT/SEC'
WRITE(*,*)'                                               '
WRITE(*,*)'                                               '
WRITE(*,*)'                                               '
WRITE(*,*)'                                               '
WRITE(*,*)' INPUT 1 TO EXIT'
WRITE(*,*)' INPUT 2 FOR ANOTHER SYSTEM DESCRIPTION'
READ(*,*) NMBR
IF(NMBR.EQ.2) GOTO 10
STOP
END
}
```