ESD-TR-66-295

MTR-216

CHARACTERISTICS AND UTILIZATION OF FORWARD-ACTING ERROR CONTROL EQUIPMENT FOR DATA COMMUNICATIONS

AUGUST 1966

L. L. Stine

Prepared for

DIRECTORATE OF AEROSPACE INSTRUMENTATION
ELECTRONIC SYSTEMS DIVISION
AIR FORCE SYSTEMS COMMAND
UNITED STATES AIR FORCE
L. G. Hanscom Field, Bedford, Massachusetts

Distribution of this document is unlimited.

Project 705B
Prepared by
THE MITRE CORPORATION
Bedford, Massachusetts
Contract AF19(628)-5165

ESD-TR-66-295

MTR-216

CHARACTERISTICS AND UTILIZATION OF FORWARD-ACTING ERROR CONTROL EQUIPMENT FOR DATA COMMUNICATIONS

AUGUST 1966

L. L. Stine

Prepared for

DIRECTORATE OF AEROSPACE INSTRUMENTATION
ELECTRONIC SYSTEMS DIVISION
AIR FORCE SYSTEMS COMMAND
UNITED STATES AIR FORCE
L. G. Hanscom Field, Bedford, Massachusetts

Project 705B
Prepared by
THE MITRE CORPORATION
Bedford, Massachusetts
Contract AF19(628)-5165

ABSTRACT

In this report, a mathematical model is developed which provides criteria for utilizing, in an optimal fashion, forward error correction devices for retransmission data communication systems. The devices are specified in terms of delay and error correction capability. System performance is described by the defined terms of error correction improvement factor, information throughput, and channel availability.

REVIEW AND APPROVAL

Publication of this technical report does not constitute Air Force approval of the report's findings or conclusions. It is published only for the exchange and stimulation of ideas.

C. V. Horrigan

C. V. HORRIGAN
Acting Director
Aerospace Instrumentation

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

SECTION I

INTRODUCTION

In digital data communications, error detection and retransmission is one method of obtaining essentially error-free transfer of data over a bursty channel between a data source and a remotely located data sink. The information throughput is the amount of error-free information received by the data sink per unit of time. In an error detection and retransmission system, throughput decreases as the rate of occurrence of errors increases. Forward-acting error-correcting codes can be used to correct errors in the data before the data are examined by the error detector of a retransmission type of communication link. Thus, by decreasing the error rate, the error-correcting code can produce an increase in information throughput.

Unfortunately, use of forward-acting error-correcting codes entails increased delay in transmission of data between the source and sink. This increase in delay can produce an effect which will decrease the throughput. Generally, the error correction capabilities of the code increase as the delay increases. In this report, a mathematical model is developed which provides criteria for utilizing, in an optimal fashion, the forward-acting error correction device, allowing for the tradeoff between delay and the error correction capabilities of the code; and specifying the performance of these devices in terms of delay and capabilities of the code. The model, which enables calculation of the data communications system performance, is described by the defined terms of error correction improvement factor, information throughput, and channel availability.

This report describes: the data communications system; the channel error characteristics; the characteristics of forward-acting error control devices; the mathematical model of the system from which the throughput is calculated; and desirable specifications of forward-acting error control equipment which permit maximum utilization of the data communications system.

All performance curves in this report are used solely to illustrate the method. These curves are not derived from any real channel. When the method described herein is used to evaluate a particular system, the actual performance curves describing the particular channel and error correction equipment should be determined and used for calculations.

SECTION II

THE DATA COMMUNICATIONS SYSTEM

A functional block diagram of a typical data communications system is shown in Figure 1. A data source sends a sequence of digital information via a noisy channel to a remotely located sink. The forward-acting error correction equipment is used to correct errors caused by the channel noise. The sink is assumed to have the capability to detect uncorrected errors in the data at all times. This is a good assumption since very powerful error-detecting codes are available.

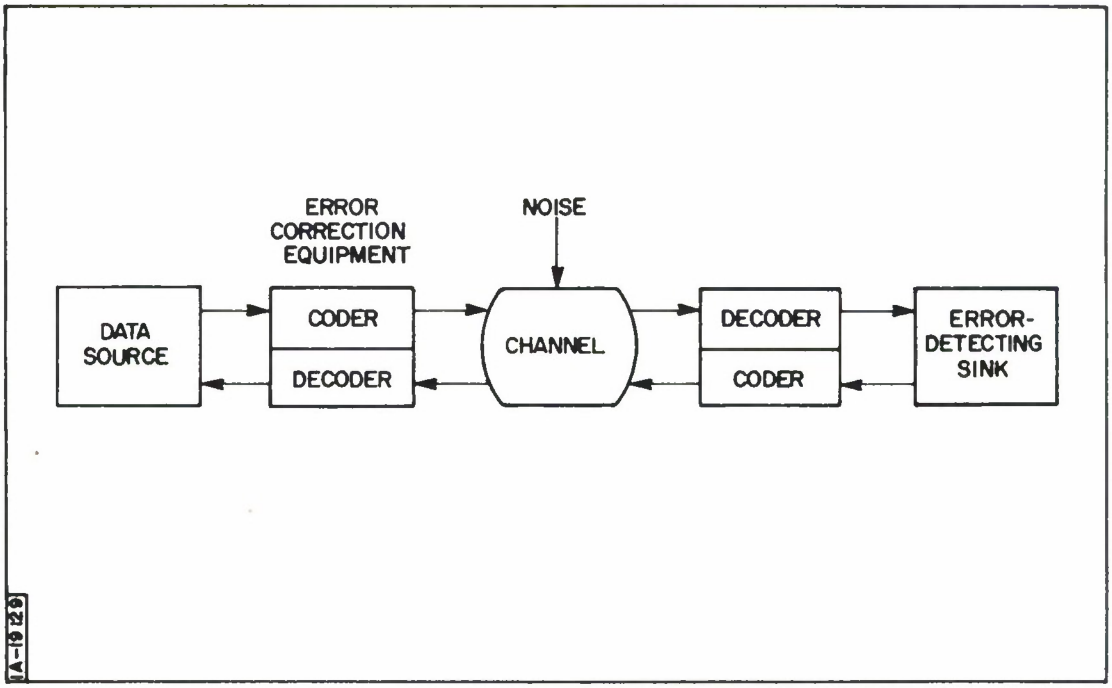

Figure 1. The Data Transmission System

The data are encoded into information blocks called words. A data word is considered in error if one or more errors appear in the data word. When a word in error is detected, the sink may operate under either of two rules of procedure for retransmission of the data in error.

(1) If an error is detected, the sink immediately requests the source to go back in the sequence to the data word received in error and recommence transmission. The sink then ignores all incoming data until the retransmitted data word is received correctly.

(2) If an error is detected, the sink requests the source to retransmit only the data word or group of data words containing the error. The sink then continues to receive all incoming data. When the retransmitted word or group of words is correctly received, it is inserted into its proper position in the received sequence of data.

The first procedure effectively shuts down data transmission for an interval equal to the round trip transmission time, but requires no insertion of retransmitted information into proper sequence in the received data. Although the second procedure does not interrupt data transmission, it requires additional operations to keep track of the sequence position of each data word in a message.

After the channel characteristics are described (Section III), the error correction equipment operation specified (Section IV), and the mathematical model of the system developed (Section V), the comparative performance resulting from each of these rules of operation will be calculated (Section VI).

SECTION III

CHARACTERISTICS OF THE CHANNEL

Error measurements indicates that communications channels for various communication media are bursty, i.e., the error occurrences are highly correlated. Attempts have been made to model these channels and to list their relevant statistical characteristics [1,2] in order to describe the behavior of the channel error patterns.

A quantity commonly used to describe the channel is average word error rate. This number is found by averaging the number of words in error in a fixed interval of the data stream over the total number of words in the interval. For example, if 10 words were in error in an interval of 10^4 words, the average word error rate would be 10 divided by 10^4 or 10^{-3}. The average word error rate varies from interval to interval. The quantity is usually plotted as in Figure 2, which shows the percent of time the channel average word error rate is less than a value specified on the abscissa. This percent of time is defined as the channel availability.

For a channel with random, uncorrelated error patterns, average word error rate is a good measure of transmission quality. For bursty channels, however, average word error rate is not a very meaningful measure unless additional higher order probabilities describing the channel behavior are given. In this report, however, only average word error rate is used to illustrate the method. Higher order probability terms would vary the calculated results, but not the method.

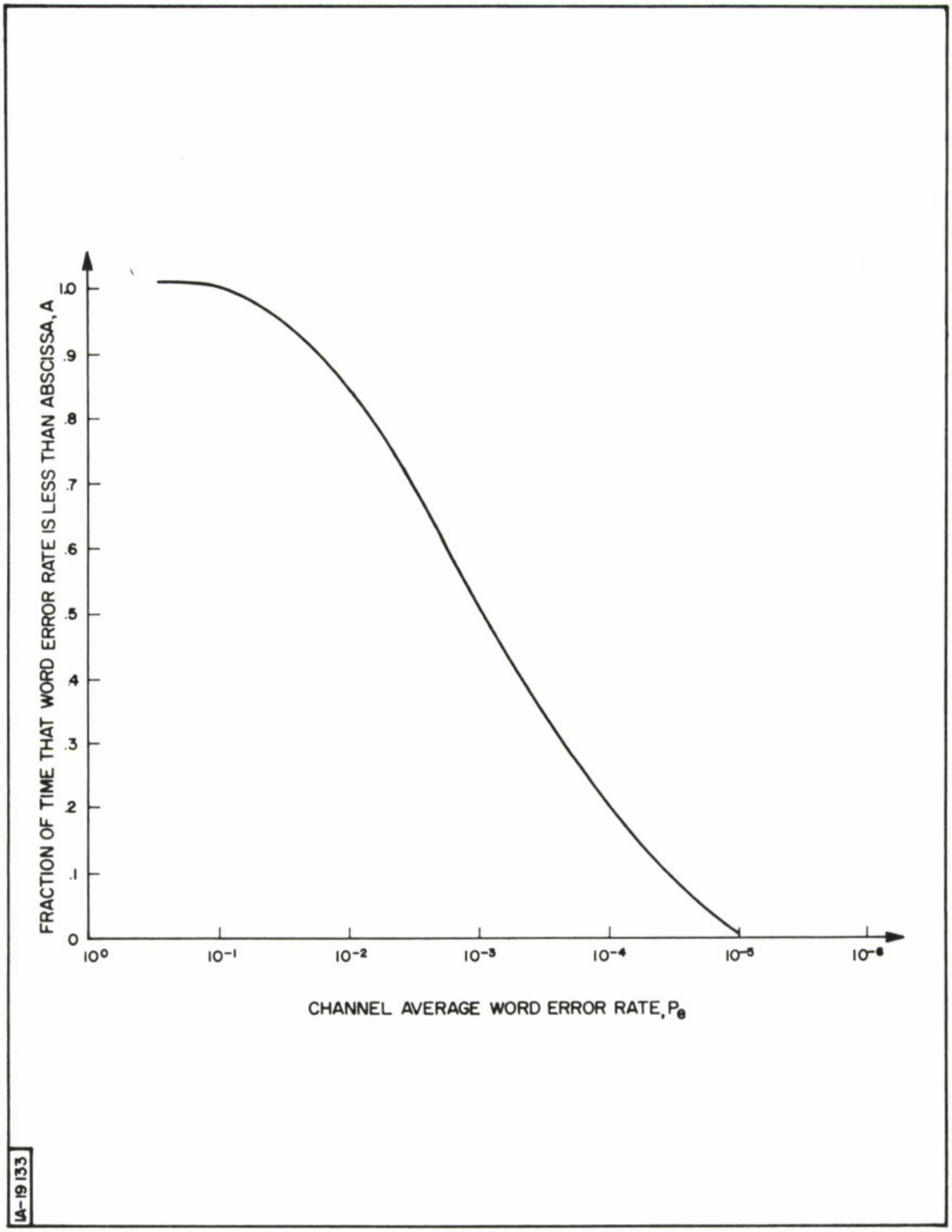

Figure 2. Unimproved Channel Availability, A

SECTION IV

CHARACTERISTICS OF FORWARD-ACTING ERROR CORRECTION EQUIPMENT

Forward-acting error correction is used to decrease the average word error rate, P_e, of the unimproved channel to a new value, P_c. The factor of improvement, F, is given by,

$$F = \frac{P_e}{P_c} \quad . \tag{1}$$

A typical characteristic of the commonly used codes is that the improvement factor increases monotonically as P_e decreases. Hence, as the channel average word error rate decreases, i.e., the channel quality improves, error correction provides greater improvement. An illustrative curve of F versus P_e is shown in Figure 3. Equation (1) expresses P_c as a function of F and P_e. Figure 2 shows the percent of time that the channel error rate before error correction, P_e, is less than a given value. Using Figure 3 and Equation (1), a curve may be derived from Figure 2 to show the percent of time that the word error rate of the improved channel, P_c, is less than a given value. This curve is shown in Figure 4.

The error rate improvement always entails increased transmission delay because of encoding and decoding calculations or operations of the equipment which attempt to match the characteristics of the code to the channel. In general, the greater the allowable delay, the more arithmetic operations are possible, the more closely the code can be matched to the channel, and the greater the achievable improvement factor. Figure 3 illustrates the behavior of F versus P_e for two different delays. The delay, D, shown in

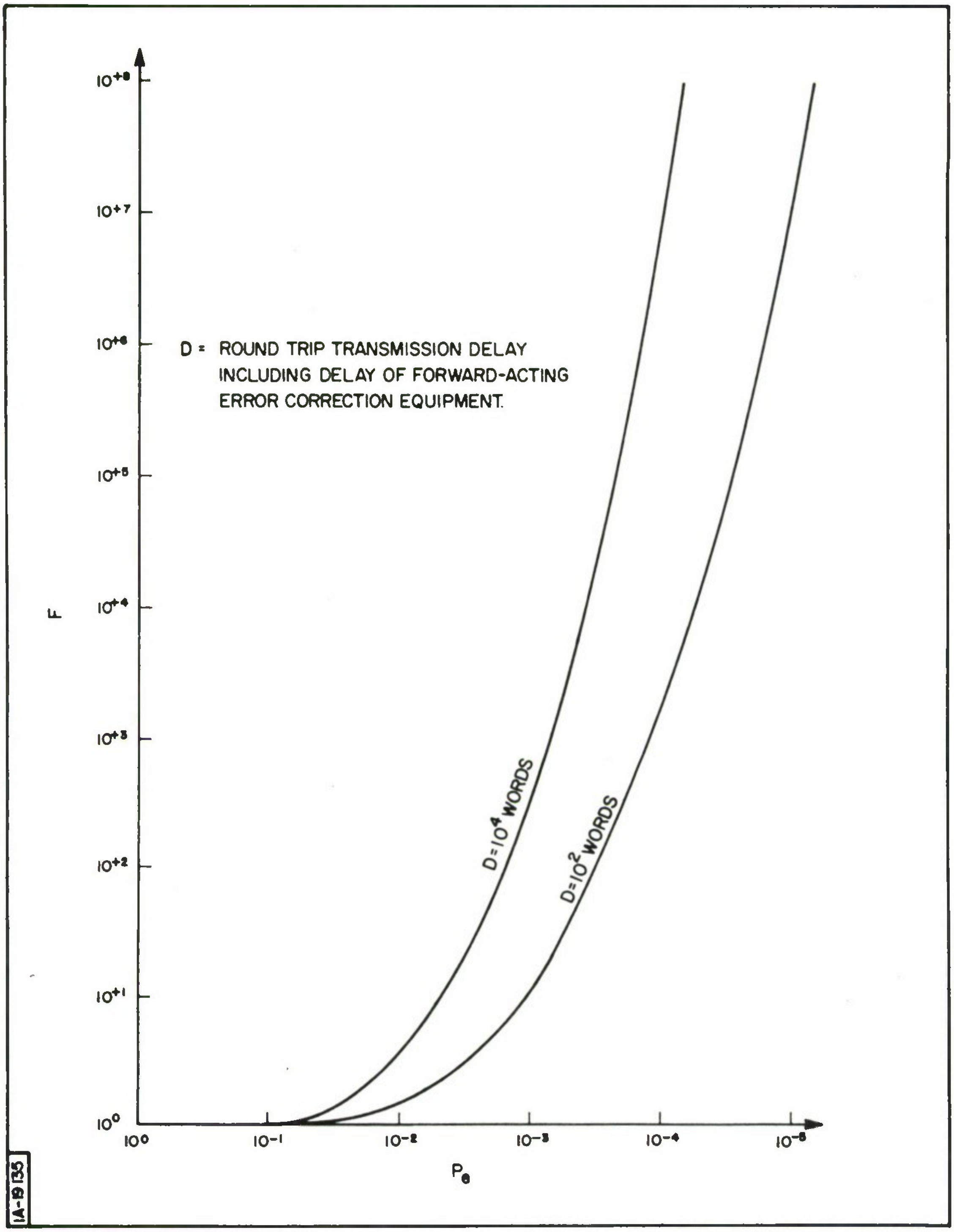

Figure 3. Improvement Factor, F, Versus Unimproved Channel Average Word Error Rate, P_e

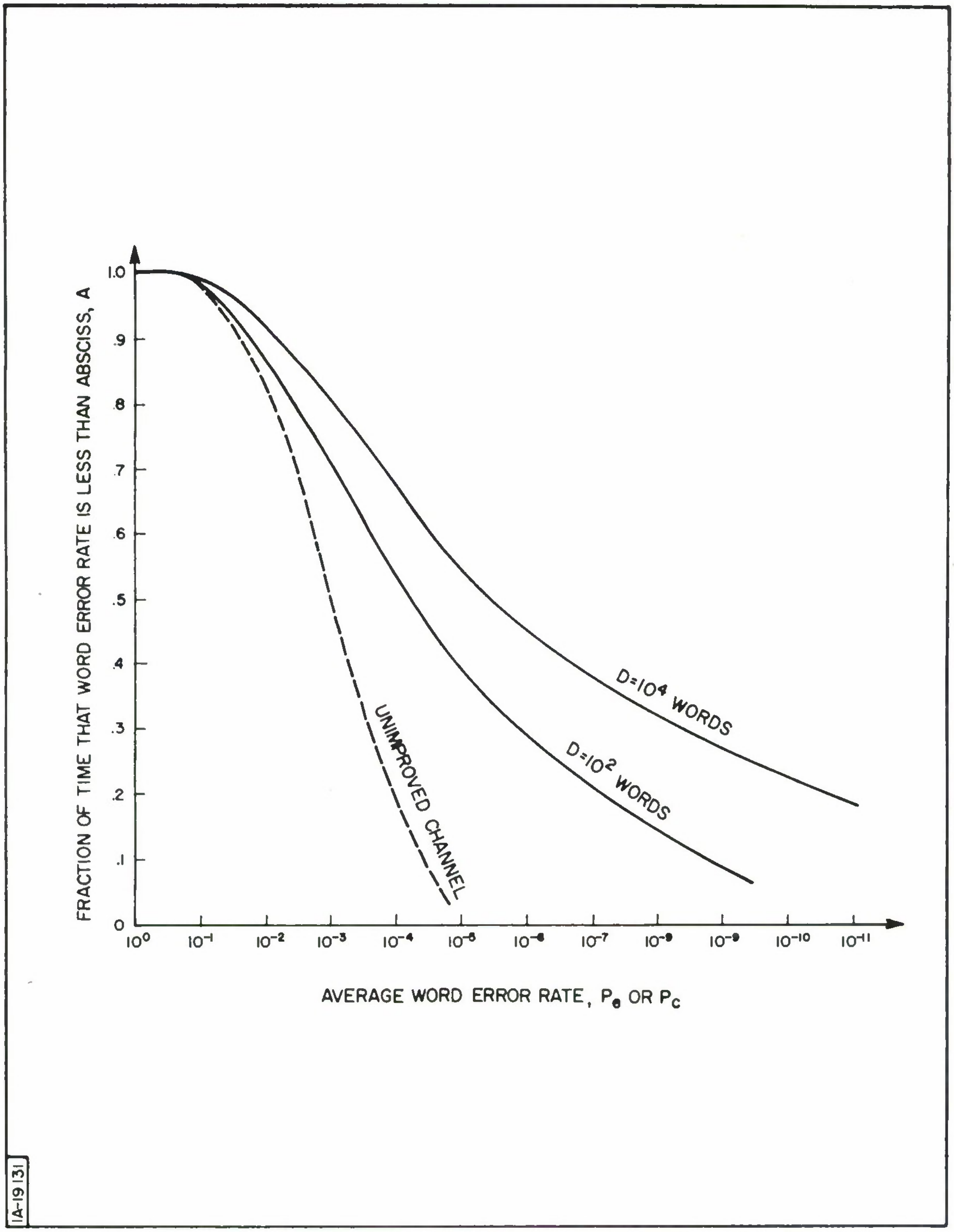

Figure 4. Availability of Improved Channel, A

the figure includes the channel propagation delay and the delay due to error correction. The one-way delay is $D/2$, and the round trip delay is D.

The coding equipment determines the improvement factor, F, and delay, D. These, in turn, can be related to the channel word error rate, P_e, in order to calculate throughput and availability of the data transmission system described in Section II.

SECTION V

THE MATHEMATICAL MODEL

INTRODUCTION

In this section, expressions are derived for: the average length of an error-free sequence of data words; the throughput rates for the two types of retransmission schemes; and the availability of the channel for the two retransmission schemes. All calculations assume that channel errors occur randomly and independently. Actually, with and without error correction, the channel is bursty, and the results for real channels should be modified accordingly. In order to describe the method and its use, however, the complications arising from the assumption of a bursty channel are omitted.

AVERAGE LENGTH OF AN ERROR-FREE SEQUENCE

Let P be the probability that a word is received by the sink with an error. Let $\overline{P}$ equal 1-P. From the theory of combinatorial analysis, it can be shown that the probability of a sequence of n words being transmitted without any errors, L(n), is given by

$$L(n) = P\,\overline{P}^{n}, \quad n > 1 \tag{2}$$

and

$$L(1) = 1, \quad n = 1.$$

The average sequence length, $\overline{n}$, is given by

$$\overline{n} = \sum_{n=0}^{\infty} n\,L(n)\,. \tag{3}$$

Substituting Equation (2) into Equation (3) gives

$$\overline{n} = 1 + \sum_{n=0}^{\infty} n\,P\,\overline{P}^{n}\,. \tag{4}$$

Equation (4) reduces to

$$\bar{n} = 1 + \frac{\bar{P}\,P}{(1 - \bar{P})^2} = 1 + \frac{P}{\bar{P}} . \tag{5}$$

Therefore, the average length of the errorless sequence is $P/\bar{P}$ words.

THROUGHPUT OF THE INTERRUPTED TRANSMISSION

A retransmission caused by the occurrence of an error and the termination of a sequence of errorless words results in a transmission interruption interval of D words, which includes the word which is retransmitted by request. The average time of transmission without interruption is $\bar{n}$ words. The average transmission interval without interruption is $\bar{n}$, and the average interruption duration is $(\bar{x} + 1)\,D$, where $(1 + \bar{x})$ is the average number of transmissions needed to receive without error the requested word. The throughput, T_I, is defined as the ratio of $\bar{n}$ to the sum of $\bar{n}$ plus $(\bar{x} + 1)\,D$. The average number of retransmissions needed to receive without error the word which was in error after the first transmission is $\bar{x}$. Thus,

$$T_I = \frac{\bar{n}}{\bar{n} + (1 + \bar{x})\,D} \qquad D \geq 1 . \tag{6}$$

The average interruption duration is D times the average number of times $(\bar{x} + 1)$ that a word received in error must be transmitted in order to be received without errors. From the mathematics of combinatorial analysis, it can be shown that the probability of x transmissions before an error-free retransmission, K (x), is given by

$$K(x) = \bar{P}\,P^{x} . \tag{7}$$

The average number of transmissions before an error-free retransmission is given by

$$\bar{x} = \sum_{x=0}^{\infty} x\,K(x) \;. \tag{8}$$

Substituting Equation (7) into Equation (8) yields

$$\bar{x} = \bar{P} \sum_{x=0}^{\infty} x P^{x} \;. \tag{9}$$

Equation (9), which is similar to Equation (4), reduces to

$$\bar{x} = \frac{P}{\bar{P}} \;. \tag{10}$$

Therefore, the average number of retransmissions including the error-free retransmission is $1 + \bar{x}$, or, since $\bar{P}$ plus P equals unity, $1/\bar{P}$. The average retransmission shutdown time including an error-free repeat of the word originally in error is $D/\bar{P}$ word intervals.

Substituting Equations (10) and (5) into Equation (6) yields

$$T_{I} = \frac{1/P}{1/P + D/\bar{P}} \;. \tag{11}$$

for the throughput. Equation (10) may be rewritten as

$$T_{I} = \frac{\bar{P}}{\bar{P} + DP} \;. \tag{12}$$

THROUGHPUT OF THE CONTINUOUS TRANSMISSION

In a continuous transmission system, errors do not cause transmission interruptions. If a word or group of words is received by the sink with errors, this block of data is retransmitted out of sequence and, if received correctly, reinserted in proper sequence at the sink. Hence, each word or group of words carries identification information so it may be retransmitted out of sequence and inserted into correct sequence position at the data sink. The retransmission shutdown interval, D, is one word or one group of words. Thus, from Equation (12) where D equals 1, the throughput becomes

$$T_c = (1 - R) \frac{\bar{P}}{\bar{P} + P} , \tag{13}$$

where R is the fraction of information units per word or group of words devoted to identification. Thus, the transmission delay does not affect throughput, but it does determine how many words separate the retransmitted word from its proper place in the sequence.

In order to reduce the fraction of information devoted to identification, R, the data may be transmitted and examined for errors N words at a time. When data are transmitted in groups of N words, all N words must be received without errors or a retransmission is requested. The probability of successful transmission $\bar{P}(N)$ is given by

$$\bar{P}(N) = \bar{P}^N . \tag{14}$$

The probability that the N word block contains at least one error P(N) is given by

$$P(N) = 1 - \bar{P}^{N} . \tag{15}$$

Let $\bar{n}_N$ be the average number of blocks transmitted without any errors and $1 + \bar{x}_N$ be the average number of transmissions of N word blocks in order to receive correctly a block which was in error after the first transmission. From combinatorial analysis similar to that of Section V,

$$\bar{n}_N = \frac{\bar{P}\,(N)}{P\,(N)} , \tag{16}$$

and

$$\bar{x}_N = \frac{P\,(N)}{\bar{P}\,(N)} . \tag{17}$$

The throughput, $T_c(N)$, of the continuous transmission when data are transmitted in N word blocks is given by

$$T_c\,(N) = \frac{\bar{n}_N}{\bar{n}_N + (\bar{x}_N + 1)} . \tag{18}$$

Using Equations (14), (15), (16), and (17), the throughput becomes

$$T_c\,(N) = \frac{\bar{P}^{N}}{\bar{P}^{N} + (1 - \bar{P}^{N})} . \tag{19}$$

COMPARATIVE THROUGHPUTS OF THE OPERATION

The throughput for the interrupted transmission case described in Section V is given in Equation (12) as a function of P_c and D. The throughput for the continuous transmission case described in Section V is given in Equation (19) as a function of P_c and N. Figure 5 graphically illustrates throughput versus channel average word error rate for D equal to 10^4 and 10^2 words and N equal to 16 and 64 words.

For these particular examples, the curves show that, as the channel word error rate decreases, the throughput of the continuous transmission case has a much faster rise to its maximum than the interrupted transmission case. Also, while the shorter block requires a high percentage of identification information, the throughput is higher for error rates below 10^{-3}.

The effect of transmission delay in reducing the throughput of the interrupted transmission case is also shown in Figure 5. Thus, to increase the throughput of this type of operation, the delay should be as short as possible to attain a given channel error rate improvement factor. The optimal tradeoff between delay and improvement factor is found in the Appendix. The results show that, for the interrupted transmission case, optimal throughput is obtained when the improvement factor increase is directly proportional to the increase in delay.

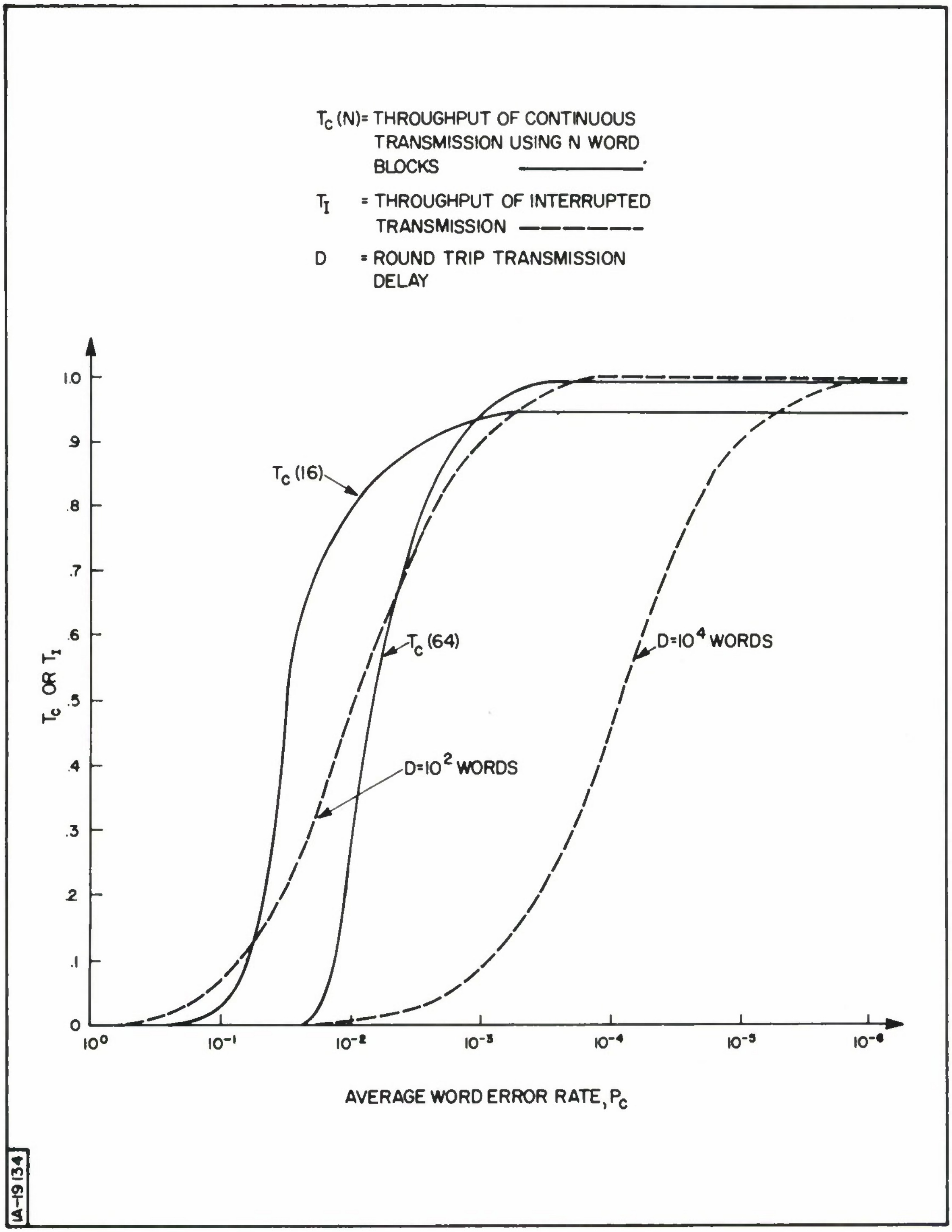

Figure 5. Throughput, T, Versus Average Word Error Rate, P_c

SECTION VI

AVAILABILITY OF THROUGHPUT

Figure 2 shows the percent of time that the unimproved channel average word error rate, P_e, is less than a given value. This quantity is defined as the channel availability, A, for a given level of error rate, P_e. Figure 4 shows the availability of the improved channel. Figure 5 plotted throughput, T, as a function of channel word error rate. Using Figures 4 and 5, it is possible to plot the channel availability, A_T, for a given level of throughput, T. These curves are shown in Figure 6.

Thus, the availability of throughput of the channel, A_T, is the percent of time the information throughput is greater than a given value T. This measure of performance enables the system designer to calculate how often a given level of throughput performance will be obtained.

Examination of Figure 6 reveals that the continuous transmission case has a sharp threshold in throughput availability as well as throughput. The 64-word block continuous transmission operation is generally worse than the 16-word block, but it crosses over to a higher throughput depending upon the improvement factor.

An additional curve (also plotted in Figure 6), which provides some useful information, is the plot of throughput availability of the interrupted transmission for D equal to 10^2 words. This curve uses the unimproved channel availability curve of Figure 2. For this case, it is interesting to note that for the unimproved channel the throughput availability is generally better than for the improved channel if the delay is too large (e.g., 10^4) for the extra improvement in channel error rate availability. This illustrates

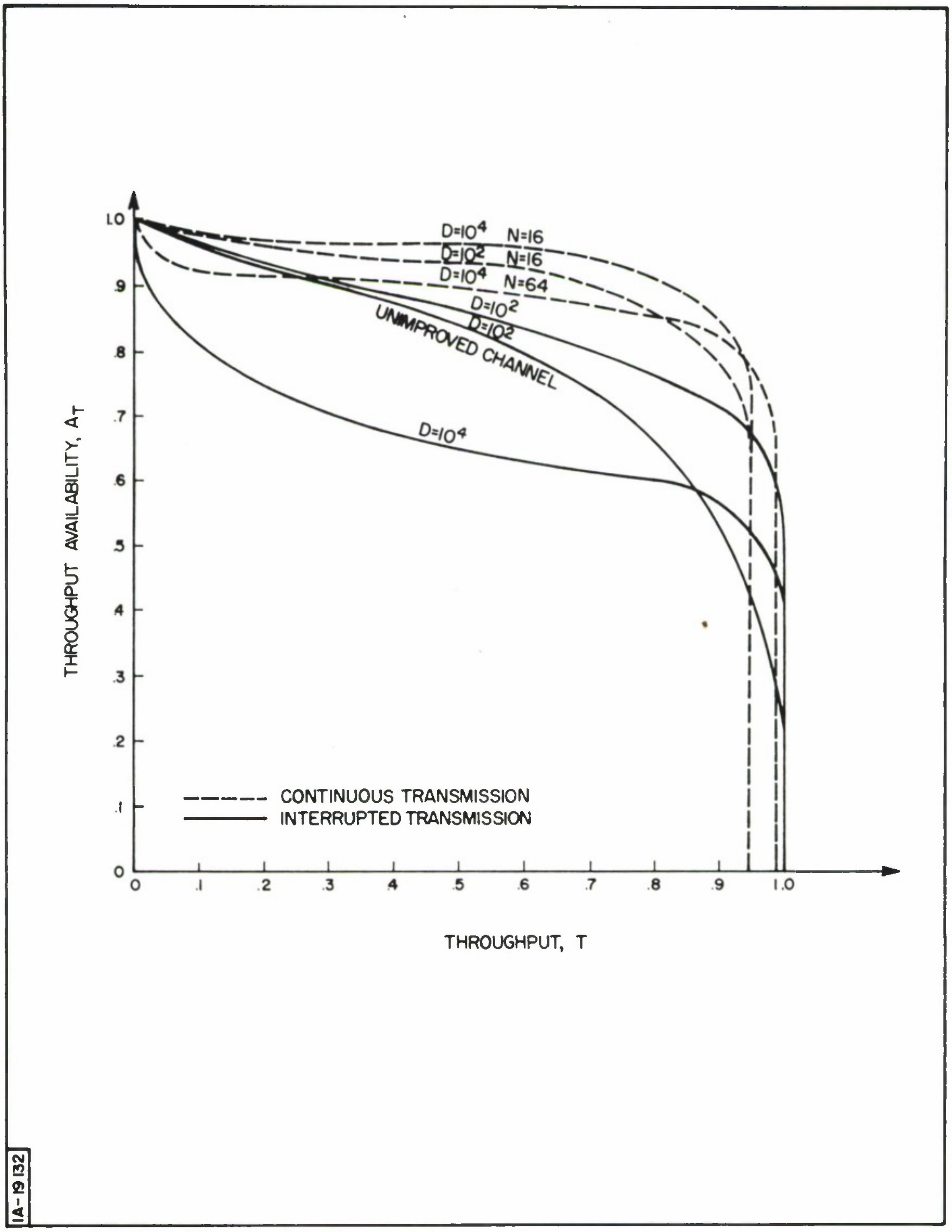

Figure 6. Availability of Throughputs, A_T

again the tradeoff between delay and improvement factor for the interrupted transmission case previously discussed. Delay does not affect throughput of the continuous transmission case, however, and there is no tradeoff possible for maximizing throughput.

SECTION VII

CONCLUSIONS

A method of evaluating the performance of communication systems using retransmission error control is to calculate throughput availability. Two different systems are compared taking into account error rate, message overhead, and transmission delay. Delay has no effect on the continuous transmission system, but the necessity for block identification can decrease the maximum obtainable throughput. For the interrupted transmission system, extra transmission delay to obtain a decrease in error rate may not increase throughput availability if certain derived limits are exceeded.

APPENDIX

THE TRADEOFF RULE BETWEEN P_c AND D FOR THE INTERRUPTED TRANSMISSION CASE

Error correction is used to decrease the channel average word error rate, P_c. The reduction in P_c always entails delay: the greater the allowable delay, the greater the achievable reduction in P_c. By inspecting Equation (12), it is obvious that the throughput will increase if an increase in delay results in a proportionally greater decrease in error probability.

The effect is graphically illustrated in Figure 7. Let the optimal operating point (P_c^*, D^*) be that point that results in a maximum throughput, T_I^*, which is that point where

$$\frac{dT_I^*}{dD} = 0 \; . \tag{20}$$

To simplify the calculation, assume P_c is very small so that $\bar{P}_c$ approaches unity. Thus,

$$P_c \approx 0, \bar{P}_c \approx 1 \; . \tag{21}$$

Equation (12) thus becomes

$$T_I = \frac{1}{1 + P_c D} \; . \tag{22}$$

Taking the derivative of T_I with respect to D from Equation (22) results in the expression

$$\frac{dT_I}{dD} = \frac{-P_c - D \dfrac{dP_c}{dD}}{(1 + P_c D)^2} \; . \tag{23}$$

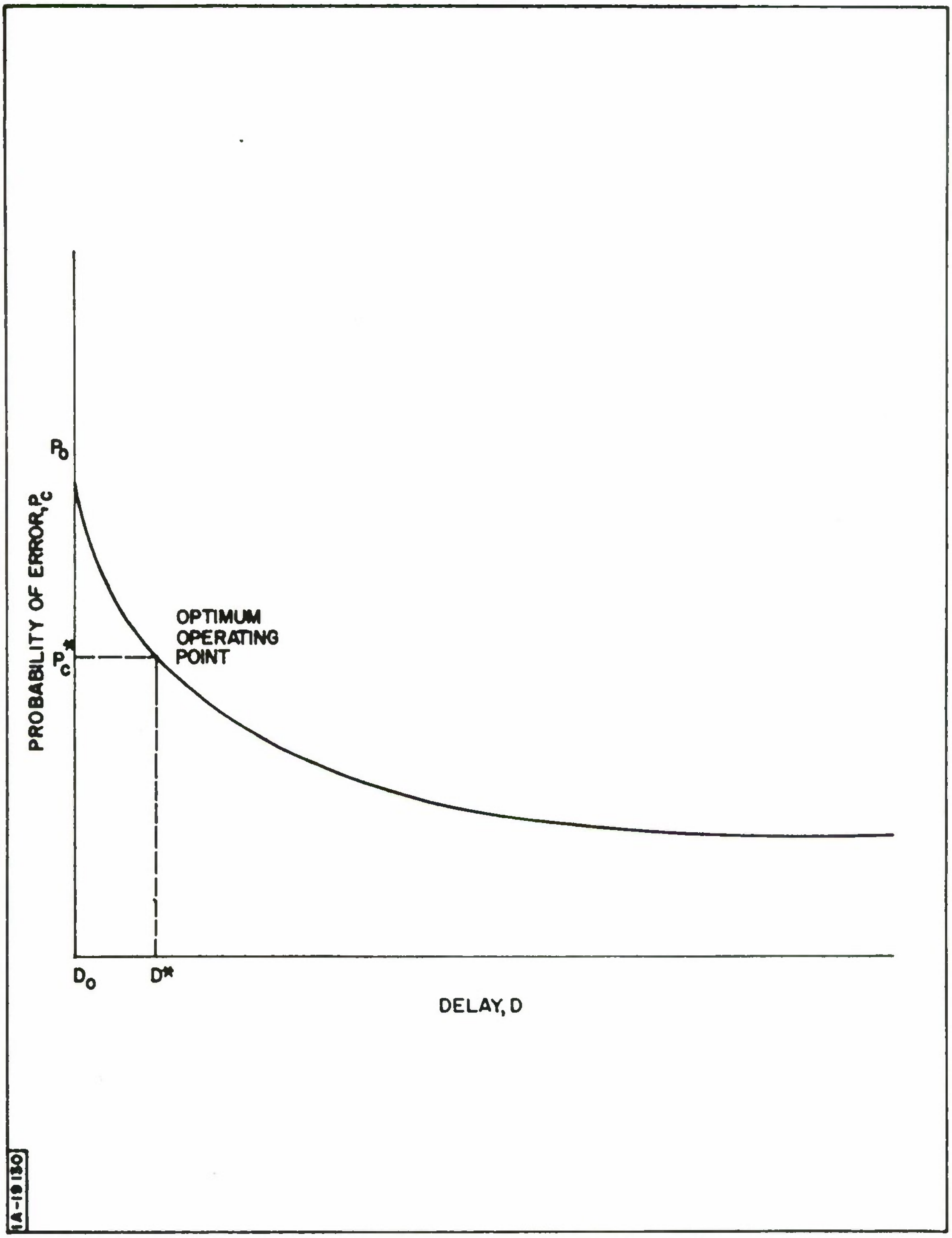

Figure 7. Probability of Word Error, P_c, Versus Delay, D

From Equation (22),

$$\frac{-P_c^* - D^* \dfrac{dP_c^*}{dD} = 0}{(1 + P_c^* D^*)^2} \quad . \tag{24}$$

Equation (24) reduces to

$$\frac{P_c^*}{D^*} = \frac{dP_c^*}{dD} \quad . \tag{25}$$

For any channel for which a P_c versus D curve can be determined, the optimal operating point is at those points which satisfy Equation (25).

REFERENCES

1. E. Berkovits, et al., A Model for Digital Error Distributions, First IEEE Annual Communications Convention, June 7-9, 1965, Boulder, Colorado.

2. S. Sussman, Communication Channel Attributes as Related to Error Control, First IEEE Annual Communications Convention, June 7-9, 1965, Boulder, Colorado.

Security Classification

DOCUMENT CONTROL DATA - R&D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author) The MITRE Corporation Bedford, Massachusetts	2a. REPORT SECURITY CLASSIFICATION Unclassified 2b. GROUP

3. REPORT TITLE

Characteristics and Utilization of Forward-Acting Error Control Equipment for Data Communications

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)

N/A

5. AUTHOR(S) (Last name, first name, initial)

Stine, Larry L.

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
August 1966	29	2

8a. CONTRACT OR GRANT NO. AF 19(628)-5165 b. PROJECT NO. 705B c. d.	9a. ORIGINATOR'S REPORT NUMBER(S) ESD-TR-66-295 9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report) MTR-216

10. AVAILABILITY/LIMITATION NOTICES

Distribution of this document is unlimited.

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY Directorate of Aerospace Instrumentation; Electronic Systems Division L.G. Hanscom Field, Bedford, Massachusetts

13. ABSTRACT

In this report, a mathematical model is developed which provides criteria for utilizing, in an optimal fashion, forward error correction devices for re-transmission data communications systems. The devices are specified in terms of delay and error correction capability. System performance is described by the defined terms of error correction improvement factor; information throughput, and channel availability.

DD FORM 1 JAN 64 1473

Security Classification

14. KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Data Transmission						
Error Control						
ARQ						
Forward-Acting Error Correction						

INSTRUCTIONS

1. ORIGINATING ACTIVITY: Enter the name and address of the contractor, subcontractor, grantee, Department of Defense activity or other organization (*corporate author*) issuing the report.

2a. REPORT SECURITY CLASSIFICATION: Enter the overall security classification of the report. Indicate whether "Restricted Data" is included. Marking is to be in accordance with appropriate security regulations.

2b. GROUP: Automatic downgrading is specified in DoD Directive 5200.10 and Armed Forces Industrial Manual. Enter the group number. Also, when applicable, show that optional markings have been used for Group 3 and Group 4 as authorized.

3. REPORT TITLE: Enter the complete report title in all capital letters. Titles in all cases should be unclassified. If a meaningful title cannot be selected without classification, show title classification in all capitals in parenthesis immediately following the title.

4. DESCRIPTIVE NOTES: If appropriate, enter the type of report, e.g., interim, progress, summary, annual, or final. Give the inclusive dates when a specific reporting period is covered.

5. AUTHOR(S): Enter the name(s) of author(s) as shown on or in the report. Enter last name, first name, middle initial. If military, show rank and branch of service. The name of the principal author is an absolute minimum requirement.

6. REPORT DATE: Enter the date of the report as day, month, year; or month, year. If more than one date appears on the report, use date of publication.

7a. TOTAL NUMBER OF PAGES: The total page count should follow normal pagination procedures, i.e., enter the number of pages containing information.

7b. NUMBER OF REFERENCES: Enter the total number of references cited in the report.

8a. CONTRACT OR GRANT NUMBER: If appropriate, enter the applicable number of the contract or grant under which the report was written.

8b, 8c, & 8d. PROJECT NUMBER: Enter the appropriate military department identification, such as project number, subproject number, system numbers, task number, etc.

9a. ORIGINATOR'S REPORT NUMBER(S): Enter the official report number by which the document will be identified and controlled by the originating activity. This number must be unique to this report.

9b. OTHER REPORT NUMBER(S): If the report has been assigned any other report numbers (*either by the originator or by the sponsor*), also enter this number(s).

10. AVAILABILITY/LIMITATION NOTICES: Enter any limitations on further dissemination of the report, other than those imposed by security classification, using standard statements such as:

(1) "Qualified requesters may obtain copies of this report from DDC."

(2) "Foreign announcement and dissemination of this report by DDC is not authorized."

(3) "U. S. Government agencies may obtain copies of this report directly from DDC. Other qualified DDC users shall request through

_______________________________________."

(4) "U. S. military agencies may obtain copies of this report directly from DDC. Other qualified users shall request through

_______________________________________."

(5) "All distribution of this report is controlled. Qualified DDC users shall request through

_______________________________________."

If the report has been furnished to the Office of Technical Services, Department of Commerce, for sale to the public, indicate this fact and enter the price, if known.

11. SUPPLEMENTARY NOTES: Use for additional explanatory notes.

12. SPONSORING MILITARY ACTIVITY: Enter the name of the departmental project office or laboratory sponsoring (*paying for*) the research and development. Include address.

13. ABSTRACT: Enter an abstract giving a brief and factual summary of the document indicative of the report, even though it may also appear elsewhere in the body of the technical report. If additional space is required, a continuation sheet shall be attached.

It is highly desirable that the abstract of classified reports be unclassified. Each paragraph of the abstract shall end with an indication of the military security classification of the information in the paragraph, represented as *(TS)*, *(S)*, *(C)*, or *(U)*.

There is no limitation on the length of the abstract. However, the suggested length is from 150 to 225 words.

14. KEY WORDS: Key words are technically meaningful terms or short phrases that characterize a report and may be used as index entries for cataloging the report. Key words must be selected so that no security classification is required. Identifiers, such as equipment model designation, trade name, military project code name, geographic location, may be used as key words but will be followed by an indication of technical context. The assignment of links, rules, and weights is optional