THE BELL SYSTEM

TECHNICAL JOURNAL

VOLUME XLIII — NOVEMBER 1964 — NUMBER 6

Effectiveness of Error Control in Data Communication over the Switched Telephone Network

By R. L. TOWNSEND and R. N. WATTS

(Manuscript received September 24, 1963)

This article describes the results of a data communication experiment designed to investigate the effectiveness of error detection and retransmission in providing high-accuracy data transmission over the switched telephone network. Data were encoded into a Bose-Chaudhuri (31,21) error-detecting code and transmitted at 2000 bits per second by a DATA-PHONE data set 201A over a variety of dialed long-distance connections. Transmitted and received data were compared to obtain error data which were analyzed to obtain an estimate of the error performance of the data set and the effectiveness of the code. The results of this analysis are presented.

During the test approximately 6.36×10^7 *31-bit code words or* 1.97×10^9 *bits were transmitted. Of these, 63,002 bits appearing in 29,731 different code words were received incorrectly. Thus, the over-all bit error rate was* 3.19×10^{-5} *and the word error rate* 4.67×10^{-4}*. The decoder was successful in detecting all but two of the erroneous code words, resulting in an average undetected word error rate of* 3.14×10^{-8} *or an average of* 9.85×10^8 *bits between undetected word errors. These results demonstrated that very low undetected error rates can be obtained in practice using an error detection and retransmission system of modest complexity.*

I. INTRODUCTION

Much attention has been focused recently on the problem of transmitting digital data over the switched telephone network with a high degree of accuracy. Selection and evaluation of error control schemes by which the desired high accuracy can be achieved require detailed information about the digital error statistics. Because of the complexity of the switched telephone network, the only feasible way to obtain this information is through the analysis of experimental data.

A method of error control which offers promise for use with telephone facilities is error detection and retransmission. An experiment has been performed to explore the feasibility of this type of error control and to obtain useful statistical information about the switched telephone network. In this experiment, a DATA-PHONE data set 201A,[1] which is a 4-phase unit designed for synchronous operation at 2000 bits per second, was used to transmit data over a variety of connections in the direct distance dialing network. The transmitted data were encoded into the Bose-Chaudhuri[2] (31,21) code described in Appendix A, which had been selected on the basis of a computer study. Transmitted and received data were compared to obtain error data from which digital error statistics were derived.

The over-all results of the test are shown in Table I. This indicates that the decoder was successful in detecting all but two of the 29,731 words received containing transmission errors. These results demonstrate the feasibility of providing high accuracy data transmission over the direct distance dialing network by using an error detection and retransmission system of modest complexity. A description of the error control equipment is given in Appendix A.

A description of the test and an analysis of the numerical error data are presented in the remainder of the article.

II. DESCRIPTION OF THE TEST

The test was conducted between March 13 and August 31, 1962, during which time approximately 1.97×10^9 bits were transmitted. A portable transmitter was used to transmit data from various locations throughout the Continental United States to a stationary receiver located at Murray Hill, New Jersey from March 13, 1962 until May 1, 1962, and then at Holmdel, New Jersey for the remainder of the test. All performance measurements were made at the receiving terminal.

At both Murray Hill and Holmdel three foreign exchange lines were installed, one each to a No. 5 crossbar, No. 1 crossbar, and step-by-step central office. The characteristics of these lines are outlined in Table II. Dialed connections were originated from the receiving terminal, which was so arranged that it could be connected to any of the three lines. At both receiving stations calls were distributed equally, as nearly as pos-

TABLE I — EXPERIMENTAL RESULTS OF DATA COMMUNICATION OVER THE SWITCHED TELEPHONE NETWORK

Number of transmitter locations	28
Number of calls	548
Number of hours of transmission	273
Total bits transmitted	1.97×10^9
Information bits transmitted	1.33×10^9
Words transmitted	6.36×10^7
Number of bits in error (total)	6.30×10^4
Number of words in error	2.97×10^4
Number of undetected word errors	2
Bit error rate	3.19×10^{-5}
Word error rate	4.67×10^{-4}
Undetected word error rate	3.14×10^{-8}
Factor of improvement (word)	1.49×10^4
Average bits between undetected word errors	9.85×10^8

sible, among the three foreign exchange lines. The duration of each call was approximately 30 minutes.

The transmitting terminal was moved to the locations listed in Table III. These were selected on the basis of their distance from the receiving terminal, types of connecting facilities and type of end switching office. Since one objective of the experiment was to collect and to analyze data transmitted over typical connections, the locations selected were in or near large metropolitan areas where data traffic is likely to be heaviest.

A pseudo-random sequence generator was used to produce a repetitive pattern of 511 distinct 31-bit code words. These were transmitted serially at 1000 bauds or 2000 bits per second by a DATA-PHONE data set 201A. Received data were demodulated with another data set 201A and then compared with the output of a synchronized, duplicate sequence generator. The output of the receiver and system performance information were recorded on magnetic tape. Error data also were recorded by means of electronic event counters. A test log was kept which

TABLE II — RECEIVING END TEST LINES

Recvr. Location	Type of End Office	Location of CO	Line No.	Line Loss to CO	Type of Line to CO
MH	#5XB	New Providence, N. J.	4641116	3.4 db	H-88
MH	#1XB	Plainfield, N. J.	PL68684	9.8 db	H-88
MH	SXS	Carteret, N. J.	5414054	13 db	H-88
HO	#5XB	Holmdel, N. J.	9464674	5.3 db	H-88
HO	#1XB	Rahway, N. J.	3814270	10.4 db	H-88 & N carrier
HO	SXS	Monmouth Junction, N. J.	DA96550	11 db	H-88 & N carrier

TABLE III—LOCATIONS OF TRANSMITTING TERMINALS

Transmitter Location City	CO Prefix	Date	Office Type	No. of Calls	Transmission Time (Hours)	Total Bits Transmitted
Rahway, N. J.	FU8	3/15/62	#1XBAR	11	5.5	3.96×10^7
Passaic, N. J.	PR8	3/13/62	#1XBAR	9	4.62	3.32×10^7
Paterson, N. J.	MU3	3/22/62	#1XBAR	7	3.22	2.32×10^7
Ridgewood, N. J.	444	3/27/62	#5XBAR	10	4.88	3.52×10^7
Manhattan (N.Y.C.), N.Y.	349	4/18/62	#1XBAR	10	4.83	3.48×10^7
Manhattan (N.Y.C.), N. Y.	HA5	4/23/62	#1XBAR	13	6.5	4.68×10^7
Manhattan (N.Y.C.), N. Y.	LT1	4/20/62	#5XBAR	12	6.17	4.44×10^7
Manhattan (N.Y.C.), N. Y.	RI9	5/11/62	#1XBAR	14	7	5.04×10^7
Manhattan (N.Y.C.), N. Y.	UN1	5/10/62	#5XBAR	9	4.45	3.20×10^7
Brooklyn, N. Y.	JA2	5/16/62	#1XBAR	12	6	4.32×10^7
Queens, N. Y.	445	5/18/62	#5XBAR	12	6	4.32×10^7
Freeport, N. Y.	FR9	5/22/62	#5XBAR	13	6.5	4.68×10^7
Central Islip, N. Y.	CE4	5/21/62	SXS	11	5.67	4.08×10^7
Trenton, N. J.	LY9	3/29/62	SXS	9	4.53	3.27×10^7
Camden, N. J.	WO4	4/2/62	#1XBAR	9	4.63	3.33×10^7
Manahawkin, N. J.	LY7	3/20/62	SXS	10	5.02	3.61×10^7
Atlantic City, N. J.	823	4/6/62	SXS	11	5.64	4.06×10^7
Bridgeton, N. J.	GL1	4/4/62	#5XBAR	7	3.68	2.65×10^7
Hartford, Conn.	247	6/25/62	SXS	10	4.97	3.58×10^7
Washington, D. C.	232	7/20/62	#1XBAR	14	6.75	4.86×10^7
Washington, D. C.	333	7/18/62	#5XBAR	14	7.0	5.04×10^7
Washington, D. C.	392	7/19/62	SXS	12	6.08	4.38×10^7
Washington, D. C.	393	7/17/62	#1XBAR	12	6.0	4.32×10^7
Newton, Mass.	244	6/27/62	#1XBAR	11	5.65	4.07×10^7
Waltham, Mass.	899	6/26/62	#5XBAR	8	3.95	2.84×10^7
Quincy, Mass.	773	6/28/62	#1XBAR	9	4.63	3.33×10^7
South Boston, Mass.	268	6/29/62	#5XBAR	14	7	5.04×10^7
Atlanta, Ga.	231	8/29/62 to 8/30/62	#5XBAR	32	16	11.52×10^7
Atlanta, Ga.	237	8/28/62	SXS	11	5.23	3.76×10^7
Atlanta, Ga.	457	8/31/62	SXS	13	6.42	4.62×10^7
Atlanta, Ga.	521	8/27/62	#5XBAR	13	6.5	4.68×10^7
Atlanta, Ga.	525	8/28/62	SXS	13	6.5	4.68×10^7
Hammond, Ind.	844	8/24/62	#5XBAR	9	4.5	3.24×10^7
Libertyville, Ill.	686	8/21/62	#5XBAR	8	4	2.88×10^7
Lafayette, Ill.	247	8/23/62	#1XBAR	12	6	4.32×10^7
Wabash, Ill.	431	8/22/62	#5XBAR	12	6	4.32×10^7
Superior, Ill.	467	8/20/62	#1XBAR	7	3.25	2.35×10^7
Los Angeles, Calif.	234	8/7/62 to 8/8/62	SXS	22	10.33	7.44×10^7
Los Angeles, Calif.	273	8/6/62	#5XBAR	12	6	4.32×10^7
Los Angeles, Calif.	385	8/8/62	SXS	10	4.91	3.54×10^7
Los Angeles, Calif.	620	8/9/62	#5XBAR	22	11.57	8.33×10^7
Los Angeles, Calif.	655	8/10/62	#5XBAR	6	3.06	2.21×10^7
San Francisco, Calif.	399	8/16/62	SXS	13	6.58	4.74×10^7
San Francisco, Calif.	981	8/17/62	#5XBAR	2	1	0.72×10^7
San Francisco, Calif.	YU1	8/14/62	#5XBAR	11	5.38	3.87×10^7
San Francisco, Calif.	982	8/16/62	#1XBAR	12	5.67	4.08×10^7
San Francisco, Calif.	982	8/15/62	#1XBAR	4	2.06	1.48×10^7
San Francisco, Calif.	982	8/13/62	#1XBAR	11	5.65	4.07×10^7

summarized the results of each call. Included were descriptions of any unusual transmission or operational conditions which caused the test to be interrupted. Appendix A contains a description of the test system. The complete test procedure is given in Appendix B.

During the test 548 calls were completed, each containing approximately 3.6×10^6 bits. The magnetic tape data were reduced and analyzed for the 412 completed calls which contained errors. The other 136 completed calls were error free. Fifty-nine attempted calls were not completed for reasons which are summarized in Appendix C.

III. ERROR RATES

During the course of the test approximately 1.97×10^9 bits were transmitted. Of these, 63,002 bits were received incorrectly, giving an over-all bit error rate of 3.19×10^{-5}. As has been mentioned earlier, data were transmitted in 31-bit code words. Of the 6.36×10^7 code words transmitted, 29,731 were found to contain one or more bit errors. This gives an over-all word error rate of 4.67×10^{-4}. The decoder was successful in detecting all but two of the erroneous code words, thus yielding an average undetected word error rate of 3.14×10^{-8}. This is equivalent to an average of 9.85×10^8 bits or 136 hours of transmission between undetected word errors.

The over-all distribution of bit error rates per call is plotted in Fig. 1. Also plotted in this figure are the corresponding distributions observed by Alexander, Gryb, and Nast[3] for transmission rates of 600 bits per

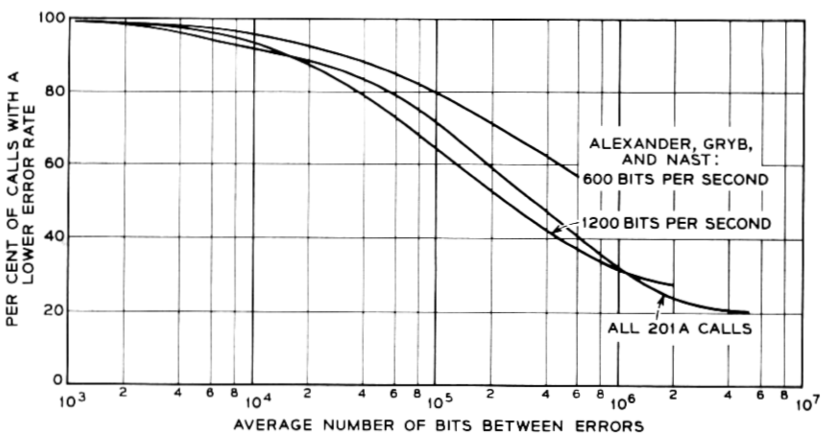

Fig. 1 — Bit error rate distribution for all calls: Alexander, Gryb, and Nast, 600 and 1200 bits/sec.

second and 1200 bits per second in a different test. The distributions for the Alexander, Gryb, and Nast 1200 bits per second data and the 201A data (2000 bits per second) show a remarkable similarity in view of the fact that the two tests employed different types of modulators operating at different speeds.

A question of major interest is "What factors have the greatest effect on error rate?" An attempt to answer this question was made by sorting the call error rates by all of the known parameters of the call, such as types of central offices at the transmitting and receiving ends, time of day, day of week, etc. Since none of the calls was actually traced, factors such as types of carrier systems in the circuit, types of intermediate central offices, etc., for any given call were generally not known. The only call parameter examined which showed a clear relationship with error rate was distance over which the call was made. Although none of the other parameters showed a definitive effect on error rate, this does not necessarily imply that these other parameters do not affect performance. It is likely that the data recorded did not allow adequate separation of the effects of these parameters.

Calls were classified into exchange, short-haul, and long-haul categories. Following the definitions used by Alexander, Gryb, and Nast, exchange calls are those within a single dialing area; short-haul calls are interarea calls between points separated by an airline distance of 400 miles or less; and long-haul calls are those exceeding 400 airline miles. Distributions of bit and word error rates per call for these categories are shown in Figs. 2 and 3. Again the bit error rate distributions of these

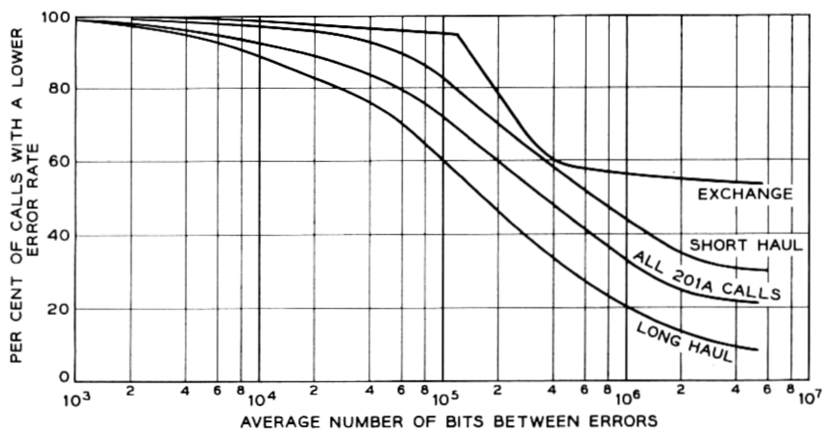

Fig. 2 — Bit error rate distribution for all calls: exchange, long-haul, and short-haul.

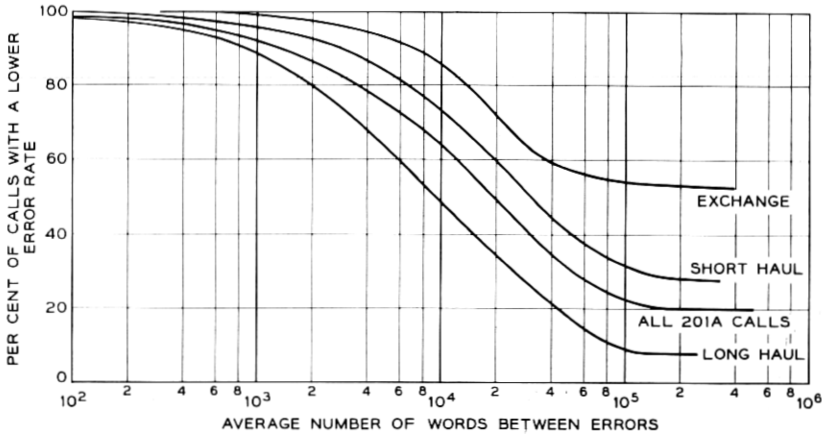

Fig. 3 — Word error rate distribution for all calls: exchange, long-haul, and short-haul.

three categories are very similar to the corresponding distributions of Alexander, Gryb, and Nast. As a rule, word error rates varied quite uniformly with bit error rate, indicating that the parameters studied had little effect on the density of error bits in an error word.

IV. CORRELATION BETWEEN ERRORS

It is well known that digital data errors in telephone circuits tend to be bunched together,[3] but little is known about the exact nature of their correlation. One measure of the degree of correlation between errors is the autocorrelation function of the bit error sequences of the calls. Here we shall define the sequence $\{X_{ji}\}$, $i = 1,2, \cdots ,N_j$ of call j to be the binary sequence having 1's in positions corresponding to the positions of bits incorrectly received, and 0's in positions corresponding to error-free bits. The number N_j of terms in the sequence is equal to the number of bits transmitted in the call. We shall define the normalized autocorrelation function $\varphi(k)$ of the bit error sequences of any collection M of calls to be:

$$\varphi(k) = \frac{\sum_{j \epsilon M} \sum_{i=1}^{N_j - k} X_{ji} X_{ji+k}}{\sum_{j \epsilon M} \sum_{i=1}^{N_j - k} X_{ji}} .$$

As the number of terms in the above expression becomes large, $\varphi(k)$ will converge to the conditional probability that, given an error bit, the bit k positions later will also be in error.

The normalized bit error autocorrelation function for all calls containing errors is plotted in Fig. 4. This curve shows the presence of a very strong periodic component. The fact that the oscillations occur at a rate of exactly 120 cycles per second strongly suggests 60-cycle power line interference with the circuit. This periodic component was traced to three calls from a single location. The bit error autocorrelation function for all calls except the three containing a 120-cycle component is shown in Fig. 5. The general shape of this curve is similar to that observed in other studies.[4] The three periodic calls are excluded from the remaining distributions.

The autocorrelation function was also tabulated individually for each call. Efforts to find relationships between the autocorrelation and the known parameters of the calls were generally unsuccessful. It was noticed, however, that the initial shapes of the autocorrelation functions were similar for most calls, but the sizes of the tails varied widely. For most individual calls the autocorrelation function decreased considerably more rapidly than did the autocorrelation for all calls, which

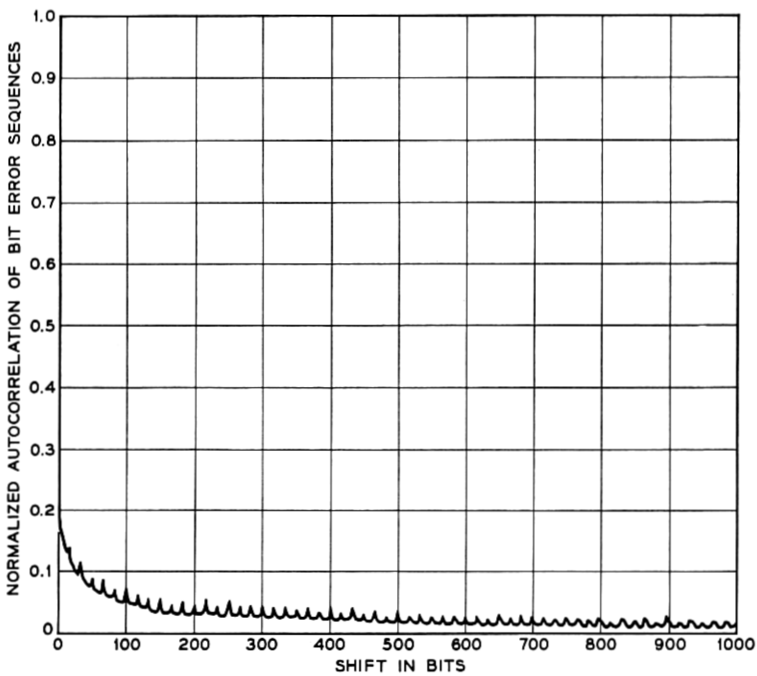

Fig. 4 — Bit error autocorrelation for all calls.

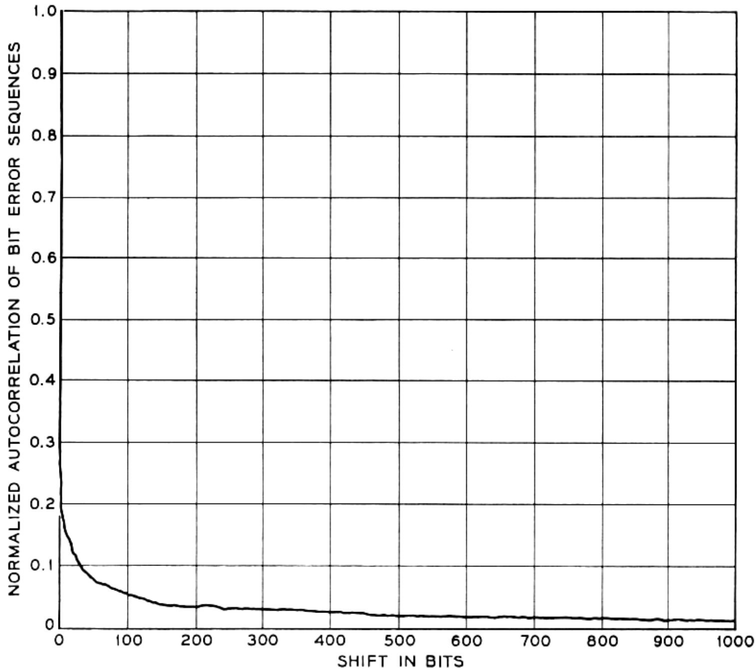

Fig. 5 — Bit error autocorrelation for all calls except those with a 120-cycle component.

is shown in Fig. 5. As one would expect, those calls whose autocorrelations had large tails were found to contain short periods of very high error rates. Surprisingly, there was not a very strong correlation between the call error rate and the size of the tail, but there was an apparent relationship between the variance of the error rate over one-minute intervals within the call and the size of the tail. This suggests that a long tail on the autocorrelation function probably was due to short dropouts or very noisy periods which were more or less independent of the over-all error rate.

The autocorrelation of the word error sequences was similarly computed. Here the autocorrelation is defined analogously, with error bits being replaced by error words. The word autocorrelation for all error calls except the three previously mentioned 120-cycle calls is plotted in Fig. 6. As one would expect, this curve is very much flatter than the corresponding error bit autocorrelation.

Further insight into the nature of the bunching of the errors can be obtained from the distribution of error-free bits between errors. The

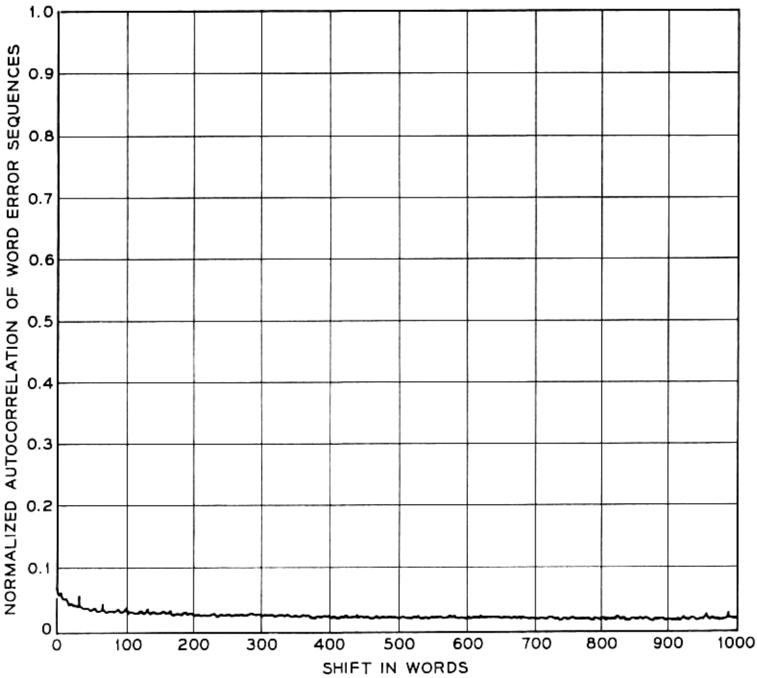

Fig. 6 — Word error autocorrelation for all calls except those with a 120-cycle component.

empirical cumulative probability distribution function of the number of error-free bits between bit errors is shown in Fig. 7. In this curve the ordinate gives the fraction of the total number of bit errors whose proximity to the previous bit error was equal to or less than the value given on the abscissa. As the number of occurrences becomes large, the empirical probability distribution function will converge to the true probability distribution function. It is interesting to note that the curve has a rather sharp knee at about ten bits on the abscissa and levels off to an ordinate value of approximately 0.65. This suggests that roughly one-third of the bit errors are separated by at least 200 bits (100 ms) from the previous error and that the remaining two-thirds of the errors are usually separated by not more than ten good bits. The errors are therefore observed to be bunched together in groups. The distribution of the lengths of these groups will be discussed in the next section.

The corresponding empirical probability distribution function for error-free 31-bit words between word errors appears in Fig. 8. The fact that the derivative of this curve changes comparatively slowly implies that small groups of errors are themselves bunched together, since the separation of errors would be much greater if they were randomly distributed.

V. ERROR BURSTS AND DROPOUTS

Knowledge of the duration and error density of a burst of errors is important, since it is desirable to avoid combining bits into words in such a way that a substantial fraction of the bits in a single word is likely to be in error. Let us define an error burst of density $1/b$ to be any sequence of bits starting with an error bit and at least b bits long such that every block of b bits within the sequence will contain at least one error bit. In other words an error burst is a sequence which begins with an error bit and does not contain b or more consecutive correct bits. We shall define the length of the burst to be the length of the

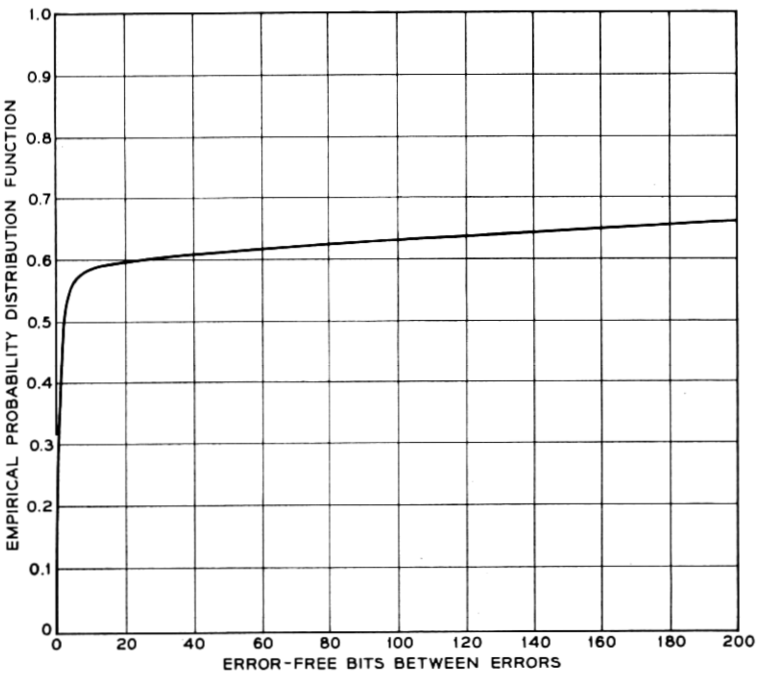

Fig. 7 — Empirical probability distribution function for error-free bits between errors.

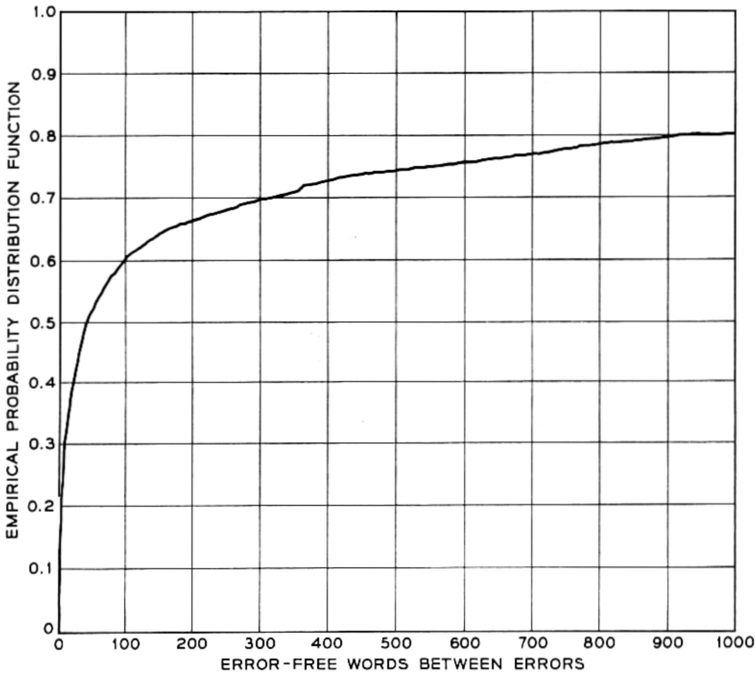

Fig. 8 — Empirical probability distribution function for error-free words between errors.

longest sequence consistent with the above definition. For example, consider the following sequence:

$$0000000000101100000100000000000.$$

Let us assume that the 0's represent bits which were correctly received and the 1's represent bits which were incorrectly received. According to the above definition this sequence contains two bursts of density 1/5. The first burst begins with the eleventh digit in the sequence and is eight bits long. The second burst begins with the twentieth digit of the sequence and is five bits long. The sequence could also be thought of as containing a single burst of density 1/10 beginning with the eleventh digit and 19 bits long.

The empirical probability distribution functions of the lengths of bursts of densities 1/5, 1/10, and 1/31 were calculated and are plotted in Figs. 9–11. It can be seen that most high-density bursts are fairly short. We observe that the bursts become considerably longer for error densities of less than 1/10, which is in agreement with Fig. 7.

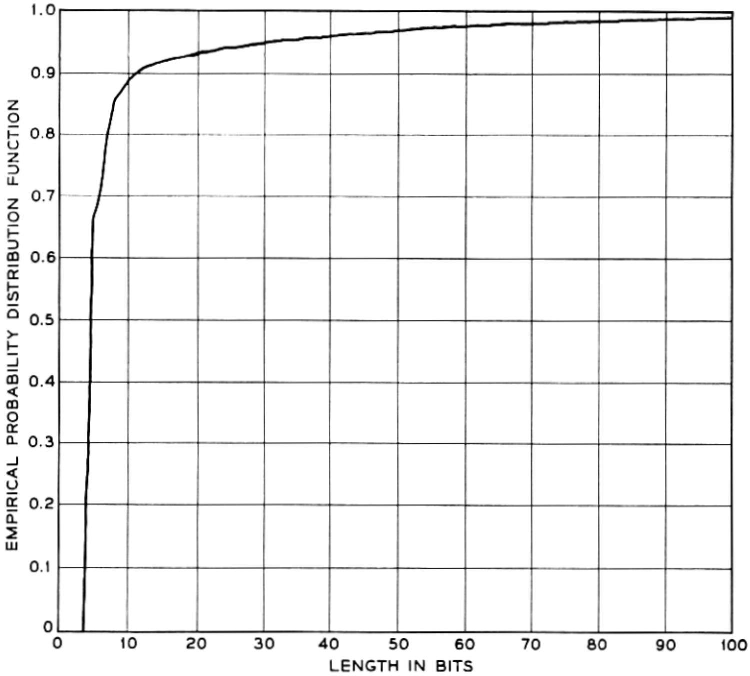

Fig. 9 — Distribution of lengths of bursts of density 1/5.

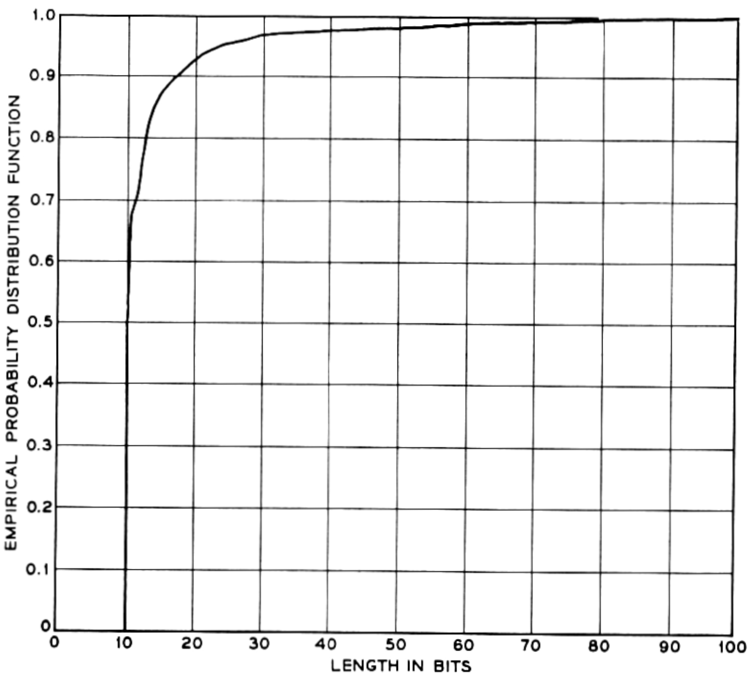

Fig. 10 — Distribution of lengths of bursts of density 1/10.

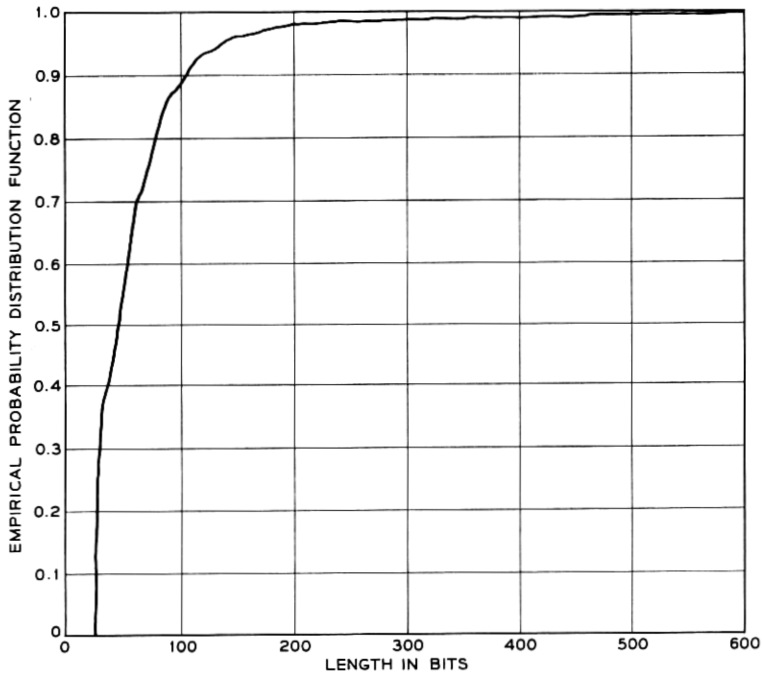

Fig. 11 — Distribution of lengths of bursts of density 1/31.

A dropout is a phenomenon whereby the connection is temporarily interrupted and the line signal drastically attenuated or completely lost for a fraction of a second. In the system tested a dropout caused only 1's to be received regardless of the transmitted message. Any sequence containing at least ten bit errors and in which only 1's were received was deemed to be a dropout. On the basis of this definition, about two per cent of the high-density bursts were found to be dropouts. These were contained in 44 different calls.

The empirical probability distribution function of the lengths of the observed dropouts is shown in Fig. 12. It should be pointed out that this distribution may be biased, since some of the longer dropouts probably caused the system to lose synchronization and were not included in the distribution. The large jump in the curve in the neighborhood of 145 bits was contributed entirely by four transcontinental calls. One plausible explanation for the occurrence of dropouts of this length is that they were caused by echo suppressors. Since the echo suppressors were not disabled during transmission, a high-energy noise impulse in the reverse channel could momentarily activate an echo suppressor. This would cause a dropout of approximately 145 bits duration.

The empirical probability distribution function of the number of bits between dropouts is shown in Fig. 13. It can be seen that the dropouts exhibit some tendency to be bunched together in time. This apparent bunching suggests fading rather than other possible causes.

The error data exhibited some asymmetry. There were about 15 per cent more $0 \rightarrow 1$ errors than $1 \rightarrow 0$ errors, a result consistent with the effect of dropouts. The fact that $1 \rightarrow 0$ errors were slightly more prevalent than $0 \rightarrow 1$ errors in calls not containing dropouts supports the conclusion that dropouts caused the asymmetry.

A more convenient distribution for some purposes is the distribution of the number of error bits appearing within a block of a given length. Following Elliott's[5] notation we shall define the function $P(m,n)$ to be the probability that exactly m bits will be in error in a block of n bits. The functions $P(m,n)$ for $n = 10, 15, 21, 23, 31, 63, 115$, and 230 are plotted in Fig. 14. These curves demonstrate quite vividly the effect of

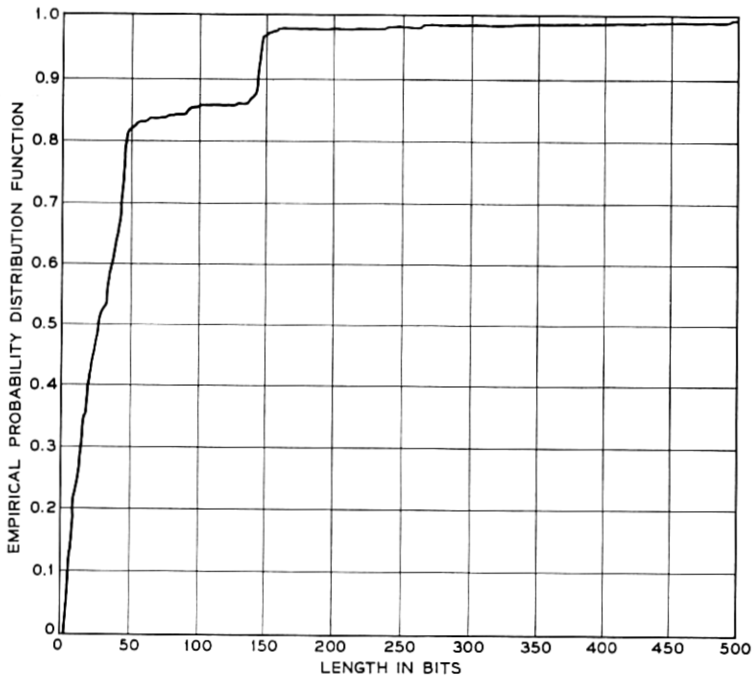

Fig. 12 — Distribution of dropout lengths.

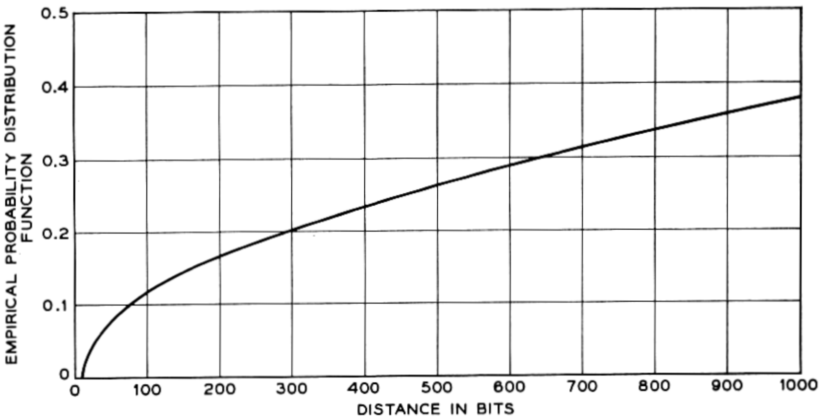

Fig. 13 — Distribution of distances between dropouts.

dropouts. Most of the curves exhibit a local maximum at about $n/2$. This is due mainly to the occurrences of dropouts longer than n. On the assumption that 0's and 1's were transmitted with approximately equal probabilities, dropouts of at least n bits in length would contribute a component in the form of a symmetrical binomial distribution to the $P(m,n)$ function. This is illustrated in Fig. 15, in which the function $P(m,31)$ is plotted with and without dropout components.

Elliott[5] has suggested that a good approximate evaluation of the performance of a code can be made by assuming that all permutations of any given number of error bits in a block are equally likely. Using his methods and the function $P(m,31)$ the estimated number of bits between undetected errors was calculated to be 8.55×10^8. As stated previously, an average of 9.85×10^8 bits between undetected errors was actually observed. This is excellent agreement, although it should be remembered that the number of observed undetected errors was too small to assure good convergence of the observed average to a true average.

The function $P(m,n)$ changes radically as n becomes very much larger. Calls were divided into 1-minute and 5-minute time intervals. The cumulative empirical probability distribution functions of the numbers of bit errors and word errors occurring within these time intervals were calculated and are plotted in Figs. 16 and 17. It is interesting to note that the distributions for 5-minute intervals are almost identical to the corresponding distributions for 1-minute intervals except for a scale factor. This suggests that the numbers of errors occurring in suc-

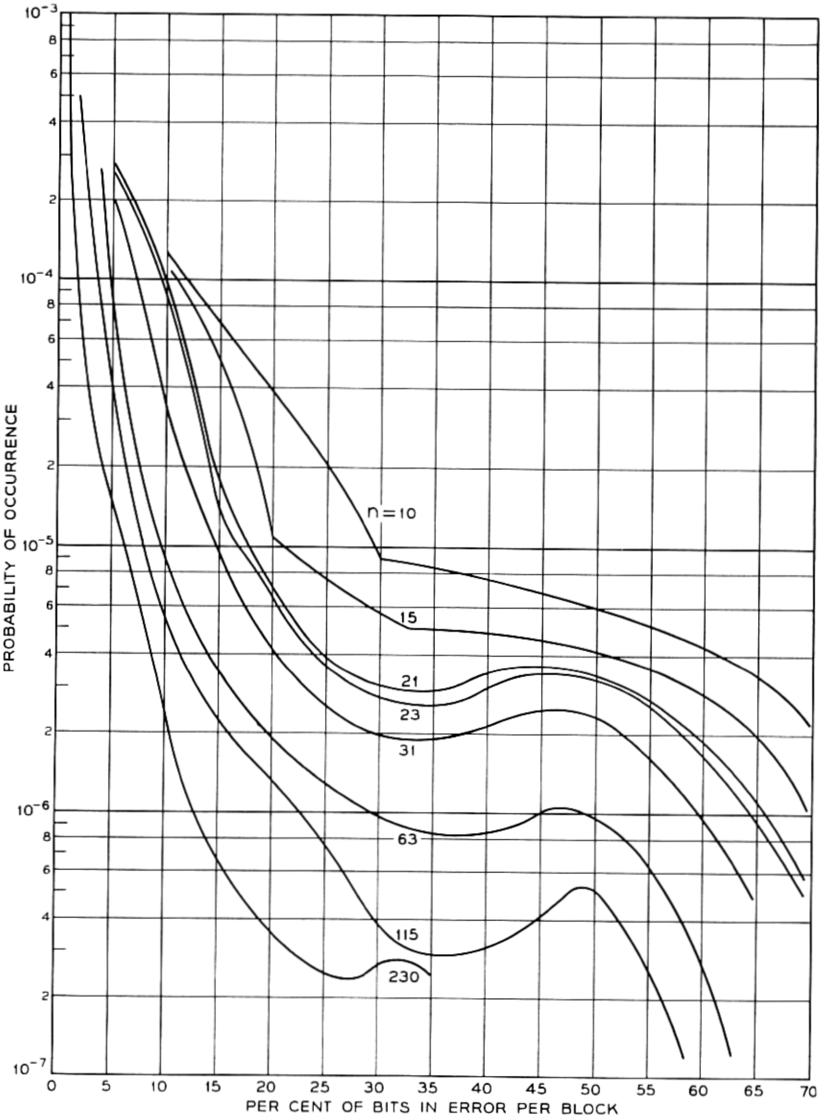

Fig. 14 — $P(m,n)$: the probability that exactly m bits will be in error in a block of n bits.

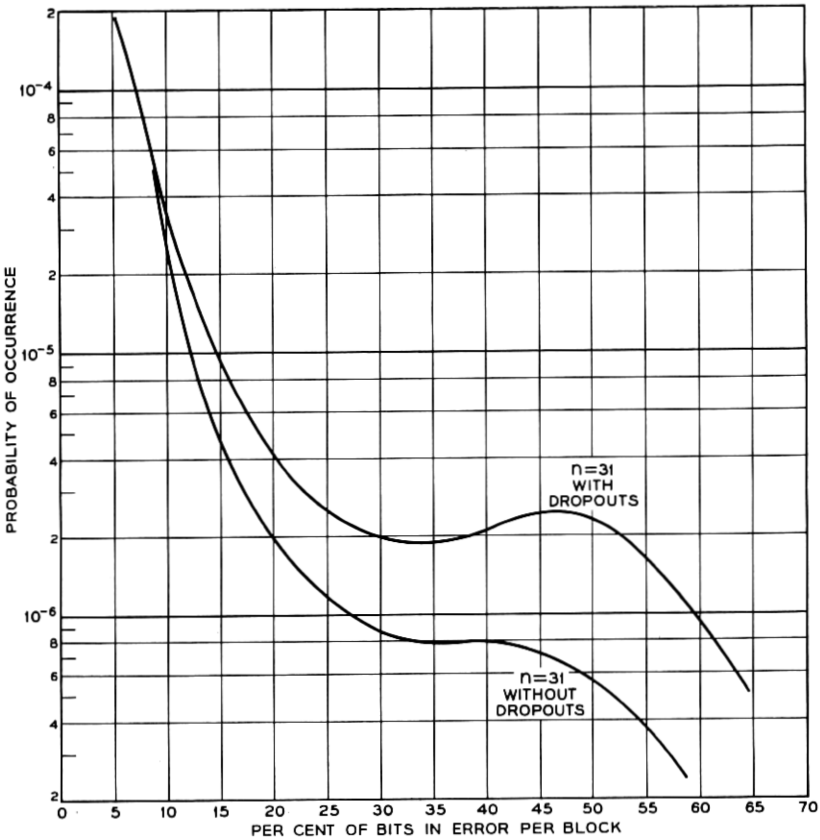

Fig. 15 — $P(m,31)$, with and without dropout components.

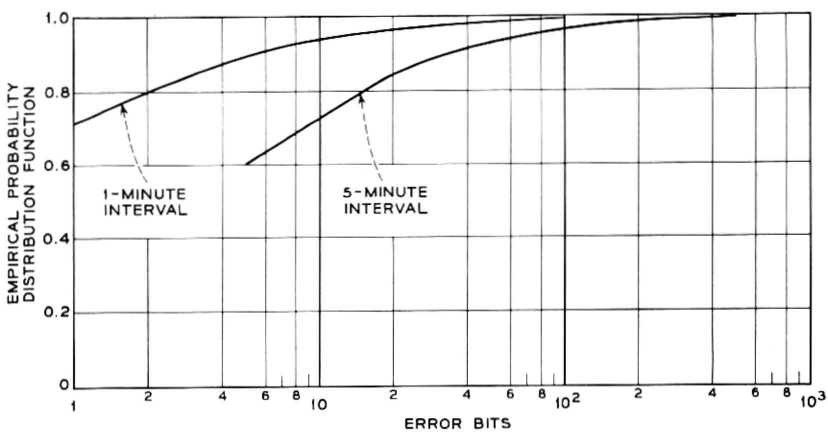

Fig. 16 — Distribution of error bits per one- and five-minute time interval.

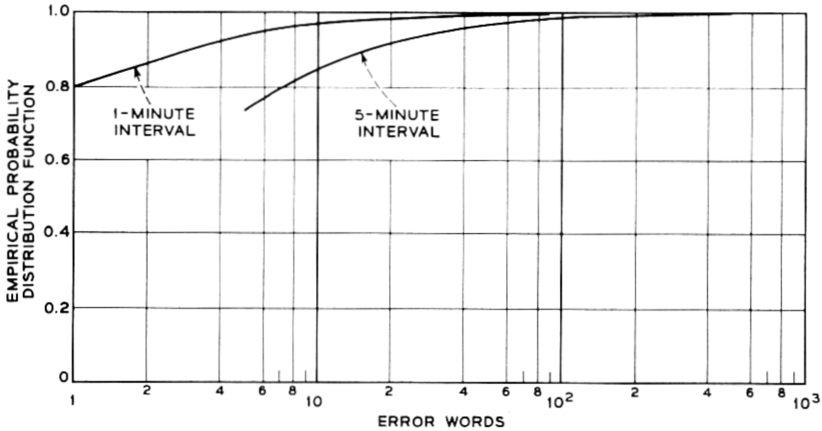

Fig. 17 — Distribution of error words per one- and five-minute time interval.

cessive intervals as long as a minute are essentially independent. Also, there is no noticeable effect of dropouts.

VI. SUMMARY

The test demonstrated that it is possible to provide data transmission over the switched telephone network with extremely low undetected error rates by means of a coding technique of moderate complexity.

The statistical properties of the error data appear to be similar to those observed in other tests. Distributions of bit error rates without regard to coding showed a strong similarity to the results of Alexander, Gryb, and Nast, despite the fact that different modems operating at different speeds were used in the two tests. The digital errors were strongly correlated with each other, and the error rates were highly nonstationary. Bit errors were observed to occur in groups of two or three and generally had a density of at least one error bit per ten good bits. These small groups of errors were themselves bunched together. Dropouts occurred frequently in certain calls, but it was difficult to determine their cause.

VII. ACKNOWLEDGMENTS

The authors wish to acknowledge the efforts of a number of people who have contributed to the results given here. A. B. Covey of the American Telephone and Telegraph Company was instrumental in making arrangements for the field tests. L. P. McRae, G. S. Robinson, and W. J. Wolf assisted in the testing program, and V. Koll assisted in the reduction of the test data. The authors also wish to thank A. B. Brown, W. R. Cowell, E. O. Elliott, M. D. Fagen, F. E. Froehlich, and R. Morris for their many helpful suggestions and technical advice. Personnel of the various operating companies involved were most cooperative and helpful.

APPENDIX A

The Error Detecting System and Performance Measuring Apparatus

A preliminary requirement on the code used for the error control experiment was that it be capable of detecting approximately 99.9 per cent of all transmission errors occurring in data transmitted by means of a DATA-PHONE data set 201A over switched, long distance telephone connections. Ease and economy of implementation were other factors affecting the selection of a code. Computer studies of a Bose-Chaudhuri[2] (31,21) code indicated that this code, subsequently used in the experimental system, had the desired error detecting ability. The above notation indicates that data were transmitted in blocks 31 bits long consisting of 21 information and 10 check bits.

The code is cyclic with a minimum distance of five* and is therefore capable of detecting any four or fewer bit errors in a 31-bit block. Furthermore, all single-error bursts† of length 10 bits or fewer, 511/512 of all 11-bit error bursts and 1023/1024 of all error bursts 12 to 31 bits in length are detected. The generator polynomial of this code is $h(X) = X^{10} + X^7 + X^6 + X + 1$, which is equivalent to saying that the code is the null space of the matrix:

$$H = \begin{bmatrix} 1000000000100110101001000011111 \\ 0100000000110101111101100010000 \\ 0010000000011010111110110001000 \\ 0001000000001101011111011000100 \\ 0000100000000110101111101100010 \\ 0000010000000011010111110110001 \\ 0000001000100111000010111000111 \\ 0000000100110101001000011111100 \\ 0000000010011010100100001111110 \\ 0000000001001101010010000111111 \end{bmatrix}.$$

---

* I.e., every code word (block) differs in at least five places from every other code word.

† The length of a "burst" in this context is the number of bits between and including the first and last bits in error in a 31-bit block.

Examination of row 10 of the H matrix (in its canonic form) reveals that the first check bit transmitted is the modulo 2 sum of information bits d_1, d_2, d_3, d_4, d_5, d_6, d_{11}, d_{14}, d_{16}, d_{18}, and d_{19}. Information bits are numbered in the order of their transmission—i.e., d_1 is the first information bit in the block, d_2 the second, etc.

The encoder was implemented by the feedback shift register shown in Fig. 18,[6] which operates as follows. With the feedback path closed (i.e., S in position 1), 21 information bits were shifted from the data source into the encoder and simultaneously transmitted. After the twenty-first bit had been encoded and transmitted, the feedback path was disabled by setting S to position 2, and the contents of the shift register (i.e., the 10 check bits) were transmitted. Thus, each block consisted of 21 information bits transmitted as a group in their original order followed by the 10 associated check bits.

Decoding was accomplished by the same circuit. The decoder was synchronized with respect to the received data and thus was able to distinguish between information and check bits. To decode a 31-bit

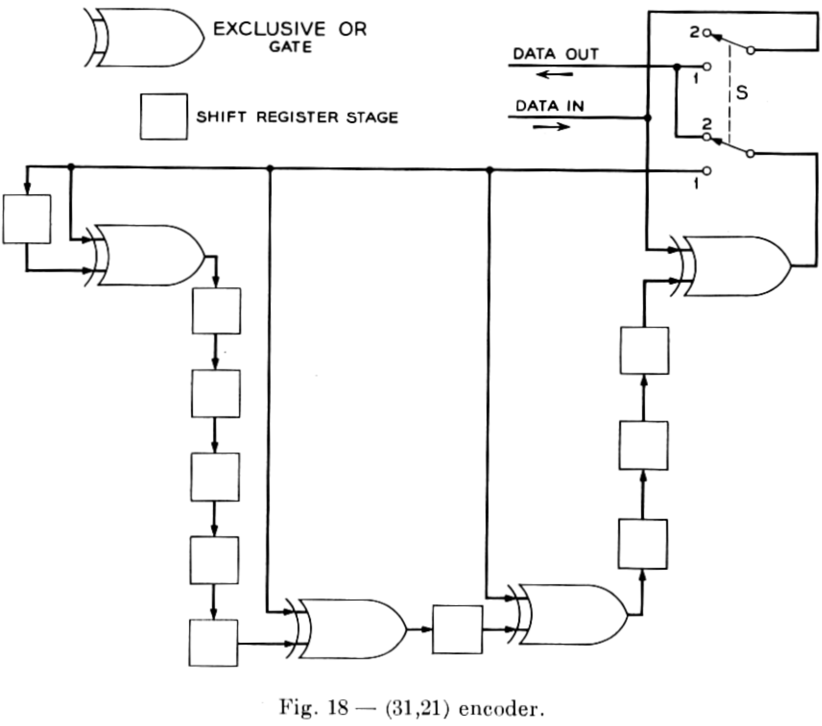

Fig. 18 — (31,21) encoder.

block the information bits were shifted from the demodulator into the encoder, which generated 10 check bits. These were compared to the 10 check bits received following the 21 information bits. Any difference between the two sets of check bits indicated the occurrence of a transmission error.

A block diagram of the error detection system is shown in Fig. 19. Both the source of data and the reference source are provided by a pseudo-random sequence generator. The output of this device is a repetitive 511-bit sequence containing every 9-digit binary sequence except the all-0 sequence. Since these data are meaningless so far as information content is concerned, a continuous chain of timing pulses shifts both the encoder and sequence generator. The output of the source is disregarded while check bits are shifted from the encoder. Since 31 and 511 are relatively prime, all possible 21-bit sections of the 511-bit sequence are transmitted as data. In practice the data source must stop after delivering 21 bits, while the 10 associated check bits are transmitted. For purposes of the test this would be impractical, since 511 and 21 have a common factor of 7, and therefore only 73 of the possible 511 21-bit

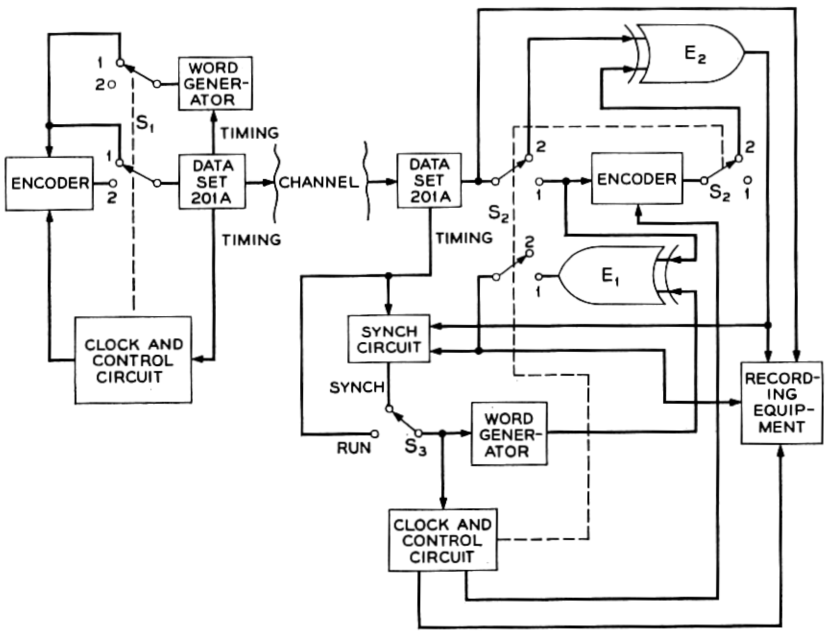

Fig. 19 — The test system.

sequences would be encoded. Timing is provided by the data set 201A, which generates 2000 timing pulses per second. The function of the clock and control circuit is to operate the encoding circuits at the transmitting and receiving terminals in the manner described earlier. The clock is driven by timing pulses from the data set and produces a periodic output signal which is "on" for 21 bits and "off" for 10. This signal is used by the control circuit to operate S (see Fig. 18), S1, and S2.*

At the transmitting terminal S and S1 were set to position 1 while information bits were shifted into the encoder and data set, then switched to position 2 while check bits were shifted from the encoder. Switches S and S1 then were reset to position 1 and the process was repeated for the next block of information.

At the receiver S and S2 were set to position 1 while information bits were received and encoded. During this time the received information was examined for transmission errors by E1, which produced an output whenever a received information bit differed from the output of the synchronized reference sequence generator. (The method of synchronization will be described in the next paragraph.) After the 21 information bits had been received and encoded, S and S2 were switched to position 2 while the check bits were received. Each of the 10 check bits in a block was compared with the output of the local encoder by E2. An output signal from E2 indicating the occurrence of detected errors was produced whenever a received check bit differed from the corresponding locally generated check bit. The outputs of E1 and E2, the received data, and timing information from the clock and control circuit were delivered to the recording equipment.

The clock and sequence generator at the receiving terminal were synchronized with respect to the demodulated data by means of the synch circuit. This circuit was activated manually by switching S3 to "synch." With S3 set to "synch" the phase of the sequence generator and clock was automatically shifted by one bit with respect to the received data in response to each output pulse from E1. When the outputs from E1 and E2 were observed to remain constant for one second or more, S3 was switched manually to "run" and the recording equipment started. Error data could not be recorded with S3 in the "synch" position.

The performance measuring apparatus was located at the receiving terminal as indicated in Fig. 19. Signals from E1, E2 and the control circuit were combined to indicate the occurrence and type of block errors.

* Switches S, S1, and S2 are implemented with solid-state circuits.

If, for a given 31-bit block, an error indication was received from E2 then transmission errors occurred and were detected. Undetected errors occurred in a block when errors were indicated in the information section by E1 but not in the check section by E2.

The received data and the error information derived from the outputs of E1 and E2 were recorded on dual-channel magnetic tape. The output of the demodulator was transformed into a series of positive and negative pulses and recorded on one channel. Following each 31-bit block one, two or three framing pulses were recorded on the second channel to indicate that the preceding block contained no errors, detected errors or undetected errors respectively. The inputs to both channels of the tape for each of the three conditions are shown in Fig. 20. Data were recorded on channel 1 and block framing on channel 2. In each case only one framing pulse follows block $K - 1$, which is assumed to be error free.

Cumulative error data for each call were recorded on five electronic event counters. The two types of bit errors ($0 \rightarrow 1$ and $1 \rightarrow 0$) occurring in user information were derived from the output of E1 and recorded separately on two counters. A third counter was incremented whenever a block was received containing any errors in the information section. The fourth and fifth counters recorded detected and undetected block errors respectively. The counters were photographed automatically at 20-second intervals during each call. A clock was included at the camera's

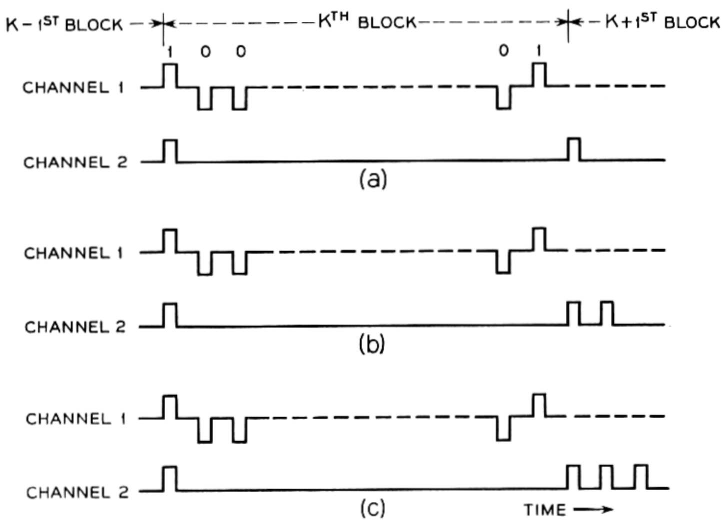

Fig. 20 — Block framing; three cases.

field of view and the film dated so that photographic data could be correlated with log book records.

APPENDIX B

Field Test Procedure

The purpose of this section is to describe the field test procedure in some detail. To review briefly, the receiving error control terminal remained fixed at Murray Hill or Holmdel and the transmitting terminal was carried to the locations listed in Table III. One day was spent at each location, during which time data were transmitted over a number of dialed connections established from the receiving terminal and used for approximately 30 minutes each.

Before initiating a series of calls between two locations, the line loss from the transmitter to the central office was measured with a 12B transmission measuring set. The data set's transmission level then was adjusted so that the signal strength at the central office was approximately −8 dbm and the data set then placed in the on-hook automatic answer mode. The location, telephone number, local loop loss, transmitting level, and type of central office serving the transmitting terminal were recorded in the test log. The encoding equipment was started and ran continuously throughout the series of calls (i.e., data were generated and encoded continuously and transmitted automatically whenever a connection was established from the receiving terminal).

After the transmitting terminal had been readied for the series of calls as described in the preceding paragraph, the receiving terminal was attached to a foreign exchange line. A connection was dialed to the transmitting terminal, which answered the call automatically and started transmitting encoded data immediately. Then the receiving error detection system was synchronized with respect to the demodulated data as described in Appendix A and the recording equipment started. Thirty minutes later the recording equipment was stopped and the call terminated from the receiving end. The data recorded on event counters, the times (local) at which the call was started and terminated, and a description of any unusual transmission or operating conditions were entered into the test log. The recording equipment was reset and the receiving terminal attached to a different foreign exchange line for the next call.

The error control equipment, data sets, and error recording apparatus were checked periodically throughout the test. This was done to insure that the data collected during the experiment would not be affected by marginally operating test equipment.

APPENDIX C

Summary of Incomplete Calls

During the test, 59 calls could not be completed for reasons which fall into three general categories. These are:

(1) long dropouts or fades resulting in loss of synchronism between the transmitting and receiving data sets,

(2) the inability to achieve initial synchronization of the terminals within a reasonable length of time, and

(3) lost connections.

These conditions will be more fully described in the following paragraphs. In addition to these 59 calls, 6 calls were interrupted due to human errors and two calls were lost as a result of local power failures.

Loss of synchronism between the transmitting and receiving data sets during otherwise normal communication caused the interruption of 30 calls. This condition usually was caused by long dropouts, particularly on long-haul connections. Dropouts lasting more than approximately 100 milliseconds generally caused the transmitting and receiving data sets to lose synchronism, since timing in the demodulator was derived from the line signal. Within 20 milliseconds of a loss of line signal the timing reverted to the natural resonant frequency of the high-Q circuit in the demodulator's bit synch recovery circuit. The natural frequency of this high-Q circuit was within one cycle of the transmitter frequency. Thus, if the modulator and demodulator remained decoupled long enough, synchronism between the two was lost. This situation was detected easily but resulted in some loss of data, as the terminals had to be resynchronized before data transmission could be resumed. Intense channel noise was observed to have approximately the same effect as dropouts.

Twenty-one dialed connections were sufficiently noisy that the test apparatus could not be synchronized within a reasonable length of time. To recapitulate, initial synchronization was obtained by automatically shifting the phase of the receiving end clock and sequence generator by one bit whenever the synchronization circuit was enabled and a received bit differed from the corresponding locally generated information bit. When the system was synchronized, but the synchronization circuit was not yet disabled, any transmission errors occurring in the information section of a block caused the synchronization procedure to be repeated. Therefore, when the channel was unusually noisy the receiving terminal would not remain synchronized long enough for the synchronization circuit to be disabled manually. If synchronization could not be established within 2 or 3 minutes after the call was placed the connection was dropped and the transmitting terminal called a second time using the same foreign exchange line. The semiautomatic synchronization procedure could have been fully automated. Had this been done, synchronization might have been achieved over a few of the connections for which the semiautomatic method described in Appendix A was unsuccessful. However, since all 21 of these calls were exceptionally noisy it appears doubtful that data set synchronism could be maintained for a full 30-minute period.

Eight connections were lost entirely during data transmission and dial tone was returned to both terminals. This situation was easily detected by the test apparatus. On at least three of these occasions the lost connections appeared to be associated with telephone maintenance operations.

The conditions described above are transmission impairments which cannot be integrated directly into the error rate data. These are, however, situations with which the data communicator must contend and are included here to provide an estimate of their frequency of occurrence. These data are of importance in the design of error control systems which must recognize such transmission impairments and allow for some type of remedial action to be taken, such as manual intervention or automatic resynchronization.

REFERENCES

1. Baker, P. A., Phase-Modulation Data Sets for Serial Transmission at 2000 and 2400 Bits per Second, AIEE Trans., Part 1, Comm. and Elect., **61,** July, 1962, pp. 166–171.
2. Bose, R. C., and Ray-Chaudhuri, D. K., On a Class of Error-Correcting Binary Group Codes, Information and Control, **3,** 1960, pp. 68–79.
3. Alexander, A. A., Gryb, R. M., and Nast, D. W., Capabilities of the Telephone Network for Data Transmission, B.S.T.J., **39,** May, 1960, pp. 431–476.
4. Townsend, R. L., Error Characteristics of a Data Communication System, IEEE Trans. Comm. and Elect., March, 1963, pp. 69–72.
5. Elliott, E. O., Estimates of Error Rates for Codes on Burst Noise Channels, B.S.T.J., **42,** Sept., 1963, pp. 1977–1997.
6. Peterson, W. W., *Error Correcting Codes*, The MIT Press and John Wiley & Sons, 1961.
7. Cowell, W. R., The Use of Group Codes in Error Detection and Message Retransmission, IRE Transactions on Information Theory, **IT-7,** July, 1961, pp. 168–171.
8. Cowell, W. R., and Burton, H. O., Computer Simulation of Use of Group

Codes with Retransmission on a Gilbert Burst Channel, AIEE Trans., Part 1, Comm. and Elect., **58,** Jan., 1962, pp. 577–585.
9. Fontaine, A. R., and Gallager, R. G., Error Statistics and Coding for Binary Transmission over Telephone Circuits, Lincoln Laboratory Report 25G-0023, Oct. 30, 1960.
10. Gilbert, E. N., Capacity of a Burst-Noise Channel, B.S.T.J., **39,** Sept., 1960, pp. 1253–1266.
11. Mertz, P., Error Burst Chains in Data Transmission, Rand Corporation Memorandum RM-3024-PR, Feb., 1962.
12. Morris, R., Further Analysis of Errors Reported in "Capabilities of the Telephone Network for Data Transmission," B.S.T.J., **41,** July, 1962, pp. 1399–1414.