INFORMATION AND CONTROL 13, 134–139 (1968)

Exponential Distributions in Markov Chain Models[1] for Communication Channels

EDWARD L. COHEN[2] AND SHIMSHON BERKOVITS

The MITRE Corporation, Bedford, Massachusetts

A brief history of Markov chain models for communication channels is given. One such model, the most complicated devised to date, is discussed in more detail. For this model, some statistics are presented in terms of sums of exponential functions. As an example, error free runs of two of the numerous trophospheric channels for which we have chosen parameters are studied here.

1. INTRODUCTION

During the last decade, a serious attempt has been made to determine the behavior of error patterns in communication links. This is only natural with the advent of error-correcting codes, some of which can correct a sizeable amount of burst errors without too much loss in information rate; for example, the Gorenstein and Zierler (1961) code or an interleaved Golay (1949) code, both of which have been used at MITRE. However, as the equipment for the implementation of a certain error-correcting code is usually expensive, more thought has been given to deriving models of communication links before going ahead with the implementation.

In this connection, one of the early published studies of telephone links was prepared by Alexander, Gryb, and Nast (1960). From this collection of data, Gilbert (1960) developed a two-state (3 parameter[3]) Markov chain model. A further study by Berkovits, Cohen, and Zierler

[1] This paper reports work performed at the MITRE Corporation for the Air Force Electronic Systems Division, Air Force Systems Command under Contract AF 19(628)-5165.

[2] Present address: Department of Mathematics, University of Ottawa, Ottawa 2, Ont. Canada.

[3] By n parameters we mean the minimum number n of probabilities that determine the Markov process completely. This definition differs from Fritchman's (1967), p. 225.

(1965) resulted in another two-state (4 parameter) Markov chain model. This latter study was concerned with one telephone link and one troposcatter link. Along the same lines, Fritchman (1967) developed a 3 and 4 state model (4 and 6 parameters) for HF radio links. So far, these are the only published studies of Markov type models.

A similar study, introduced by Berger and Mandelbrot (1963), is concerned with the Pareto distribution. For those interested in this kind of model, we suggest the papers of Sussman (1963), Sussman (1965), and Lewis and Cox (1966). We must point out that the Pareto model has been applied only to telephone circuits, as have most of the non-Markov type models.

2. THE MARKOV MODELS

We introduce another Markov chain model, the results of which can be found in Berkovits and Cohen (1967a). The model consists of three states (9 parameters), and it is the most complicated Markov model introduced to date. Unfortunately, as the number of parameters grows, the degree of complexity also grows. However, we can present some statistics that Markov models produce without going into a detailed explanation.

Let us define

$$u(n) = P[x_1 = \cdots = x_n = 0 \mid x_0 = 1] = P[0^n \mid 1]. \tag{1}$$

This is a statistic which is relatively easy to obtain by experiment and which is believed to be particularly useful for estimating the performance of certain types of channels and error correctors. Markov models produce run distribution functions like $u(n)$ which are sums of exponential terms (each state giving one term). Therefore, for our three state model, we can write

$$u(n) = AJ^n + BL^n + CM^n \tag{2}$$

where

$$u(0) = A + B + C = 1. \tag{3}$$

Furthermore, we have

$$u(1) = AJ + BL + CM \tag{4}$$

$$u(2) = AJ^2 + BL^2 + CM^2. \tag{5}$$

Now a statistic which is essential to all channel modeling is the error rate, $P_1 = \dfrac{\text{the number of bit errors}}{\text{the number of total bits}}$.

We use P_1 in the equation

$$1 = \sum_{i=0}^{\infty} P(10^i) = P_1 \sum_{i=0}^{\infty} u(i) = P_1 \left\{ \frac{A}{1-J} + \frac{B}{1-L} + \frac{C}{1-M} \right\}, \quad (6)$$

to yield 4 equations in 6 unknowns.[4] However, if we can estimate J and A, this will leave 4 equations in 4 unknowns. These equations are not linear, but we have available a Fortran program, identified as Program I in Berkovits and Cohen (1967b) which will solve equations (3), (4), (5), and (6) for B, C, L, and M. Also, for each approximation to J and A, $AJ^n + BL^n + CM^n$ can be computed for those values of n for which $u(n)$ is known; thus the best determination of A, B, C, J, L, and M can be made by direct comparison with the data. For $n = 1, 2, 3, 2^2, 2^3, \cdots, 2^{25}$, the calculations of $u(n)$ are done by Program I.

To get J and A, we assume (e.g., see the channels below) that $J \gg L \geqq M$, so that when n is large (say $n \geqq 2^{15}$), L^n and M^n are negligible in comparison to J^n. Then $u(n_1) = AJ^{n_1}$, $u(n_2) = AJ^{n_2}$ and hence $u(n_2)/u(n_1) = J^{n_2-n_1}$, which can be solved for J. Next set $u(n_2) = AJ^n$ or $u(n_1) = AJ^{n_1}$ and find approximations for A. If enough data is available, it is advisable to use several values of $n \geqq 2^{15}$ to obtain more than one approximation to A and J.

3. TWO TROPOSPHERIC CHANNELS

We have been successful in selecting channel parameters for a number of tropospheric channels. We present two examples here. The first (Run #55), described by Terzian (1966), took place from East Island, Puerto Rico to Grand Turk Island, Bahamas on 7 October 1965 at 1203. There were 2100 $\times$ 2400 bits sent at 2400 bits per second. The number of bit errors = 24,277; hence $P_1 = .0056669$.

Using the algorithm stated in section 2, with $P_1 = .0056669$, $u(1) = .8$, and $u(2) = .644$, we arrived, after several trials, at what seemed to be a very good approximation. Estimating $J = .9999584$ and $A = .007$,

[4] The first equality in (6) is true, for if the noise digits are scanned in reverse, that is $x_0, x_{-1}, x_{-2}, x_{-3}, \cdots$, a 1 must be encountered sooner or later. Now $P(10^i)$ is the probability that the first 1 occurs at x_{-i}, and these events are mutually exclusive. Therefore, $\sum P(10^i) = 1$.

TABLE I

Distribution of Error-Free Runs of Length n or Greater—Channel 1

	Data	Model
$u(1)$	.7949	.8000
$u(2)$	.6300	.6440
$u(3)$	.5132	.5222
$u(2^2)$	.4271	.4270
$u(2^3)$	.2096	.2122
$u(2^4)$	.1067	.0937
$u(2^5)$	.0568	.0543
$u(2^6)$	.0276	.0276
$u(2^7)$	.0134	.0109
$u(2^8)$	.0078	.0071
$u(2^9)$	.0062	.0069
$u(2^{10})$	.0057	.0067
$u(2^{11})$	.0053	.0064
$u(2^{12})$	.0053	.0059
$u(2^{13})$	.0047	.0050
$u(2^{14})$	.0036	.0035
$u(2^{15})$	.0019	.0018
$u(2^{16})$	.0005	.0005

it was seen that $L = .9476$ and $M = .777317$. From these we obtained Table I, which shows a comparison between the data and the model calculations.

The channel of the second test was a composite communications link mostly made up of the troposcatter type, but including some coaxial cable. It is believed that most of the errors were introduced in the troposcatter section. The total length of the link was well over 6000 circuit miles. The test took place in April 1966. $P_1 = 987/6{,}478{,}336 = .000152$, $u(1) = .664$, and $u(2) = .4656$. Using the same methods as before, it was estimated that $J = .999994$ and $A = .0394$. From this we obtained $L = .97$ and $M = .59872$. Table II gives a comparison between the data for this channel and the model calculations.

4. COVARIANCE AND ERROR CLUSTERS

In the framework of Markov process models, probability distribution functions which are the sums of exponential terms are not uncommon. Being, in that respect, just like the function $u(n)$, they are just as useful in the modeling. We mention two of them here.

TABLE II

Distribution of Error-Free Runs of Length n or Greater—Channel 2

	Data	Model
$u(1)$	.6636	.6640
$u(2)$	.4650	.4613
$u(3)$	.3191	.3385
$u(2^2)$	.2442	.2636
$u(2^3)$	.1449	.1574
$u(2^4)$	.0922	.1215
$u(2^5)$	.0780	.0897
$u(2^6)$	.0669	.0584
$u(2^7)$	.0578	.0421
$u(2^8)$	.0466	.0394
$u(2^9)$	.0426	.0393
$u(2^{10})$	.0415	.0392
$u(2^{11})$	.0395	.0390
$u(2^{12})$	.0395	.0385
$u(2^{13})$	.0385	.0375
$u(2^{14})$	.0324	.0357
$u(2^{15})$	.0314	.0324
$u(2^{16})$	.0284	.0266
$u(2^{17})$	.0203	.0180
$u(2^{18})$	.0071	.0082
$u(2^{19})$	.0010	.0017

The first statistic is the covariance function

$$\begin{aligned} r(n) &= P[x_n = 1 \mid x_0 = 1], n = 1, 2, 3, \cdots \\ r(0) &= 1. \end{aligned} \tag{7}$$

It can be expressed in terms of the Markov chain parameters, but also can be obtained as sums of exponential terms

$$\begin{aligned} r(n) &= A_1 J_1^n + B_1 L_1^n + C_1 M_1^n \\ r(0) &= A_1 + B_1 + C_1 = 1. \end{aligned}$$

The other statistic is the relative probability of error clusters of length n or greater. This is given by

$$\begin{aligned} e(n) &= P[x_2 = \cdots = x_n = 1 \mid x_0 = 0 \text{ and } x_1 = 1] \\ &= P[1^{n-1} \mid 01], n = 2, 3, \cdots \\ e(1) &= 1. \end{aligned} \tag{8}$$

Again we have

$$e(n) = A_2J_2^n + B_2L_2^n + C_2M_2^n$$

$$e(1) = A_2J_2 + B_2L_2 + C_2M_2 = 1$$

For information about $e(n)$ in terms of Markov parameters, one can read Berkovits and Cohen (1967c). Of course, if two or more functions such as $u(n)$, $r(n)$ or $e(n)$ are used simultaneously to achieve a better fit, or if, having a satisfactory fit to some of these statistics, one seeks information about others, the A's, B's, and C's, the J's, L's, and M's must be expressed in terms of the underlying Markov chain parameters.

RECEIVED: April 1967; REVISED June 19, 1968.

REFERENCES

ALEXANDER, A. A., GRYB, R. M., AND NAST, D. W. (1960), Capabilities of the telephone network for data transmission. *BSTJ* **39**, 431–476.

BERGER, J. M., AND MANDELBROT, B. (1963), A new model for error clustering in telephone circuits. *IBM J. Res. Develop.* **7**, 224–236.

BERKOVITS, S. AND COHEN, E. L. (1967a), A 3-state model for digital error distributions. The MITRE Corp., Bedford, Mass., **ESD-TR-67-73.**

BERKOVITS, S. AND COHEN, E. L. (1967b), Three programs for evaluating the parameters of a communications channel model. The MITRE Corp., Bedford, Mass., **ESD-TR-67-74.**

BERKOVITS, S. AND COHEN, E. L. (1967c), Error clusters in a Markov model for communication channels. The MITRE Corp., Bedford, Mass., **ESD-TR-67-355.**

BERKOVITS, S., COHEN, E. L. AND ZIERLER, N. (1965), A model for digital error distributions. First IEEE Annual Commun. Conv., 103–111.

FRITCHMAN, B. D. (1967), A binary channel characterization using partitioned Markov chains. *IEEE Trans. Inform. Theory* **IT-13**, 221–227.

GILBERT, E. N. (1960), Capacity of a burst-noise channel. **BSTJ 39**, 1253–1266.

GOLAY, M. J. E. (1949), Notes on digital coding. *Proc. IRE* **37**, 657.

GORENSTEIN, D. AND ZIERLER, N. (1961), A class of error-correcting codes in p^m symbols. *J. Soc. Indust. Appl. Math.* **9**, 207–214.

LEWIS, P. A. W. AND COX, D. R. (1966), A statistical analysis of telephone circuit error data. *IEEE Trans. Comm. Tech.* **COM-14**, 382–389.

SUSSMAN, S. M. (1963), Analysis of the Pareto model for error statistics on telephone circuits. *IEEE Trans. Comm. Sys.* **CS-11**, 213–221.

SUSSMAN, S. M. (1965), Communication channel attributes as related to error control. First IEEE Annual Commun. Conv., 5–13.

TERZIAN, J. (1966), Field test of the MITRE 950 coder/decoder. The MITRE Corp., Bedford, Mass., **ESD-TR-66-101.**