Estimates of Distributions of Random Variables for Certain Computer Communications Traffic Models

by

E. Fuchs and P. E. Jackson

Bell Telephone Laboratories, Incorporated

Holmdel, New Jersey

ABSTRACT

A study of multiaccess computer communications has characterized the distributions underlying an elementary model of the user-computer interactive process. The model used is elementary in the sense that many of the random variables that generally are of interest in computer communications studies can be decomposed into the elements of this model. Data were examined from four operational multiaccess systems, and the model is shown to be robust; that is, each of the variables of the model has the same distribution independent of which of the four systems is being examined. It is shown that the gamma distribution can be used to describe each of the continuous variables of the model, and that the geometric distribution can be used to describe the discrete variables. Approximations to the gamma distribution by the exponential distribution are discussed for the systems studied.

Estimates of Distributions of Random Variables for Certain Computer Communications Traffic Models

by

E. Fuchs and P. E. Jackson
Bell Telephone Laboratories, Incorporated
Holmdel, New Jersey

Introduction

Since time sharing burst on the world some 6 or 7 years ago, many analytical studies have been published of the behavior of such systems. [1,3,7,9,13,14,15,16,20] In general, the completion of an analytical study of a real process requires several steps to be performed: construction of a process model, analysis of the model, estimation of the model parameters, and verification of the results. It is sad to report that in almost all of the published studies, the last two steps are omitted.* It is evident that the basic reasons for these omissions are (1) the difficulties encountered in the collection of necessary data due to the

---

*The pioneering work of Alan Scherr[24] was of course supported by extensive measurements on the M.I.T. Project MAC CTSS System, and his results were verified by simulations. Other investigations which were supported by measurements were undertaken for the JOSS system at RAND Corp.[2,25], the Q-32 Time Sharing System at S.D.C.[6,17,18,26], and additional investigations at Project MAC[10]. Each of these investigations was performed for a specific problem for the system at hand, with no attempt at generalization. However, the results of these studies have been quoted in lieu of measurements by authors of more general studies. An excellent summary and comparisons of these investigations may be found in[23].

complexity of requisite simulations, the potential impairment of the efficiency of real systems by the measurement process, and the problem of avoiding violation of the proprietary constraints of systems applications; (2) the costs in time and dollars for conducting such studies; and (3) the questionable utility of such data in light of the rapid evolution of system capabilities and user characteristics. Nevertheless, as was first pointed out by Sackman[23] in 1967, inferences drawn from such models for the design of systems without empirical determination of parameter values and without testing of the model with the estimated parameters rest on extremely shaky foundations.

Clearly, the third reason is the most difficult to respond to. Many systems are changing so rapidly that a detailed characterization of any one will probably be outdated before it is completed. However, the architecture of computer communication systems has matured to the point that the potential for insight gained from analysis of operational systems for testing models and for forming a basis for research aimed at improvement far outweigh the drawback of obsolescence. Indeed, this situation calls for continued study and review.

If analytical models are to be of value in the design of systems, then the first two problems can be resolved. Efforts have been underway for some time at Bell Telephone Laboratories to model the user-computer interaction process in on-line multi-access computer systems as an aid in the development of new computer communication systems and services. The studies include extensive efforts at the collection of data from representative working systems* to obtain realistic estimates of the parameters of the models. In a previous paper,[12] Jackson and Stubbs reported some of the results of these efforts; specifically, a data stream model of the interaction process was presented, together with estimates of the average values of the basic random variables of the model as obtained from measurements on working systems. In this paper, we report additional results. First, we present the results of goodness of fit tests in which standard probability density functions are fitted to the empirical estimates of the distributions of the random variables of the model. Second, we examine the significance levels of the fits for the various probability density functions and find that analytically tractable probability density functions can be used for the variables with reasonable significance levels. Third, we note a consistency between systems of widely varying types and applications in characterization of key variables and comment on the significance of this consistency.

In a small way, these analyses are analogous to the early studies of Erlang[8] and others 60-70 years ago,† in which representative examples of traffic data were collected for the

---

* In every case these data are obtained on the premises of the computer service provider and with his full permission and cooperation. To ensure the privacy of the four systems under discussion, however, they are not identified by name.

† Molina[21] reports that G. T. Blood of the AT&TCo in 1898 found a close agreement between the terms of a binomial expansion and the results of the observations on the distribution of busy telephone calls. This is the earliest reference that we have been able to find to empirical studies aimed at verification of assumptions employed in telephone traffic modeling.

purpose of characterizing local and toll telephone system behavior. The Poisson arrival rate process and exponential interarrival time distribution were results of some of the earliest of these studies. It is interesting to note that the validity of these characterizations has been retained throughout the years despite the many technological changes in telephone systems and sociological changes in telephone usage.

To provide a framework for presentation of the results, we first give an overview of the study methods and review the data stream model presented in Reference [12]. We then discuss the techniques employed to characterize the variables. Finally, we present the results of the study.

Methods and Models

The modus operandi for this study is an in-depth analysis of selected multiaccess computer communication systems. These systems were selected on the basis that they are representative of the advanced state of the art, that the providers of the particular system are knowledgeable in communications, that the systems are fully operational with the initial break-in period accomplished, and that the computer service providers are willing to participate in the study. More detail on the selection procedure is given in Reference [12].

The data which are utilized in the results reported here are the detailed relationships of the flow of message characters to and from users and computers during on-line transactions. The model describes the communications process in terms of random variables which give intercharacter times and the sizes of clusters of characters as they are transmitted through the communication interface so the raw data could be collected at the computer ports of active multiaccess computer systems. The model did not require nor did we collect data from internal computer processes such as the length of various internal queues.*

Figure 1 illustrates the data stream model. A "call" (or a connect-disconnect time period) is represented as the summation of a sequence of time periods during which the user sends characters without receiving, interleaved with time periods during which he receives characters without sending. (This implies half-duplex operation. Simple modifications to the model would allow the accommodation of full-duplex operation.) The periods during which the user is sending characters to the computer are defined as user burst segments. The periods during which he is receiving characters sent from the computer are computer burst segments. A burst segment, by definition, begins at the end of the last character of the previous computer burst segment. Similarly, a computer burst segment begins at the end

---

* It is apparent that a model which portrays the interplay of the internal computer processes, such as memory management and processor time scheduling algorithms, with the communication processes would be more satisfactory for joint optimization of computer and communication performance. However, acquisition of data describing the former processes was not within the scope of this study.

of the last character sent by the user. The first burst segment of a call begins when the call is established, and the last burst segment ends when the call is terminated as measured at the computer interface.

Within a given burst segment, there are periods of line activity and of line inactivity. The first inactive period of a user burst segment is defined as think time. That is, think time is the time that elapses from the end of the last previous computer character until the beginning of the first user character in that burst segment. In most cases, think time is employed by the user to finish reading the previous computer output and to think about what to do next. The corresponding inactive period in a computer burst segment is called idle time. In some systems idle time represents time during which the user waits for the return of "line feed", after sending "carriage return"; in other systems, idle time represents the time during which the user's program is being processed or is in queue. The remaining inactive periods within a burst segment are called intercharacter times and interburst times. A prerequisite for their definition is the definition of a burst.

Two consecutive characters are defined as belonging to the same burst if the period of inactivity between the characters is less than one-half character width. Thus, each burst is the longest string of consecutive characters where the period of inactivity between any two consecutive characters is less than one-half character width. All of the characters in a burst must, of course, be transmitted from the same party (user or computer). For example, every character of an unbroken string of characters sent at line rate is in the same burst.

For characters within the same user burst, an inactive time between two consecutive characters is called a user intercharacter time. The corresponding variable for computer bursts is computer intercharacter time. For bursts within the same user (computer) burst segment, the inactive time between two consecutive bursts is called a user (computer) interburst time. Five final variables of the data stream model are: number of user bursts per burst segment, number of computer bursts per burst segment, number of characters per user burst, number of characters per computer burst, and temporal character width (time from start to end of one character).

Collected Data and Analysis

During the study, data have been collected for a large number of transactions for each of several multiaccess computer systems. Data from four of the systems are discussed in this paper. These systems are labeled A, B, C and D. Systems A and B have the same computer equipment and basically the same mix of computer applications (scientific/engineering programming and problem solving); although the average loads supported by the two systems during the study periods were quite different. System C has computer equipment different from each of the others and its mix of user applications is oriented toward business problem solving. System D also has computer equipment different from each of the others, and its applications are data collection and data dissemination in an inquiry/response method of operation. All four systems serve low-speed, half-duplex, teletypewriter-like terminals. Table I summarizes the salient characteristics of these systems.

TABLE I

	Systems			
	A	B	C	D
Computer Type	Brand X	Brand X	Brand Y	Brand Z
Transmission Speed (Characters/Second)	10	10	15	10
Primary Application	Scientific	Scientific	Business	Inquiry/ Response
Load*	Moderate	Heavy	Moderate	Light/ Moderate

The random variables of Figure 1 are of two types. Some are discrete, such as the number of characters per burst. Others are continuous, such as think time. Modeling techniques most commonly used in computer communications studies include queueing processes, renewal processes, birth-death models, Markov processes, and to a limited extent, flow models. Most key random parameters of models used in computer-communication studies are either inter-event times such as times between arrivals at a server or burst length counts such as the number of arrivals in a batch arrival process. In solving these types of models, only a very few random functions are tractable, and in some cases allowable. In the category of desirable functional forms fall the Poisson,

---

* The term load denotes the relative occupancy of the processor due to on-line demands and background batch work (if any); nothing is implied directly as to the load on the communication channel.

geometric and binomial distributions for discrete processes, and the gamma distribution family for continuous processes. Hence, we are extremely interested in the extent to which the key parameters of such models can be described by these few desirable distributions.

Data collected from the communication lines at the computer ports of the four operating systems described above were used to seek desirable distributions to describe each of the random variables of the data stream model. These data were laundered to remove ambiguities and were then partitioned into sets describing each of the variables. For each set of data for each variable for each system, goodness-of-fit tests were performed to ascertain which standard probability functions could be used to describe the variables.*

The set of distributions used for goodness-of-fit tests included the normal, Cauchy, Laplace, chi-square, exponential, hyperexponential, gamma, and lognormal distributions for continuous variables and the geometric (with and without mass at the origin), uniform, Poisson, compound Poisson and binomial distributions for discrete variables. For each variable, a compound goodness-of-fit test was performed where the parameters of the hypothesized distributions (those being tested) were adjusted so that

---

* As the existing tests for goodness-of-fit were not satisfactory for our purposes because of their low power or excessive computation time, a new test was devised. This test is briefly outlined in the Appendix.

the mean and variance for a two-parameter distribution were the same as the sample mean and sample variance. For a single-parameter distribution the mean of the distribution was equated to the sample mean.

Results of the goodness-of-fit tests are shown in Table II. From the table, we see that the geometric distribution can be used to describe every discrete process but one (the single exception is an impulse function which is a degenerate form of the geometric distribution). Similarly, each of the continuous random variables of the model can be described by the gamma distribution, and the think times, idle times and interburst times can be described additionally by the lognormal distribution. These results are significant for two reasons.

<u>First</u>, the data stream model, which is elementary in the sense that many of the variables that generally are of interest in computer communication studies can be decomposed into the elements of the model, is shown to be robust; that is, each of the variables of the model is described by the same distribution independent of the computer system being examined.* These results were obtained in spite of the fact that three

---

* Although the truth of the statement for the "number of characters per user burst" is artificial, it is made because even for that case the same distributional form can be used in practice with no operational difficulty by chosing appropriate parameters for the distribution.

TABLE II

RESULTS OF GOODNESS OF FIT TESTS

ACCEPTABLE* DISTRIBUTIONS†

Random Variable	Systems A	B	C	D
No. of Burst Segments per Call	G	G	G	G,CP
Think Time	L,Γ	L,Γ	L,Γ	L,Γ
User Interburst Time	L,Γ	L,Γ	L,Γ	L,Γ
Computer Interburst Time	L,Γ	L,Γ	L,Γ	L,Γ
No. Bursts/User Burst Segment	G	G	G	G
No. Bursts/Computer Burst Segment	G	G	G	G
No. Characters/ User Burst	G	G	I	G,CP
No. Characters/ Computer Burst	G	G	G	G
User Intercharacter Time	Γ	Γ	N/A	Γ
Computer Intercharacter Time	Γ	Γ	Γ	Γ

* Acceptable at the five percent level.

† Γ - gamma distribution, L - lognormal distribution G - geometric distribution, CP - compound Poisson distribution, I - Constant at X = 1.

different computer types and operating systems were investigated. In addition, the computer loads and programming applications were different. Thus, in modeling data communication systems, we can apply the analytical results from a long-holding time system to other long-holding-time systems merely by changing the parameters of the distributions. Jackson and Stubbs[12] have examined the mean values of the model variables for the first three systems and make the following observations:

1. Delays introduced by the computer (primarily idle time and computer interburst delay) can be a large component of holding time and are affected by the number of simultaneous users on the system, probably by the computer scheduling algorithm, and by the characteristics of the communications control unit.
2. The average number of characters sent by the computer to the user is an order of magnitude greater than the number of characters sent by the user to the computer.
3. Delays introduced by the user are a significant contributor to average holding time and are remarkably close in absolute values for the four systems studied.

These three observations are examples of information that may be employed by system designers in investigations into improved communications for multiaccess computers. In modeling (probabilistically) the behavior of present and proposed systems to determine their sensitivity to particular elements of the data stream model, the parameters of the distributions need only be changed and not the distributions themselves. These data are equally valuable for investigations into computer operating systems. For example, one might investigate changes in computer scheduling algorithms as reflected in changes in idle time and interburst delay parameters, changes in transmission speed from computer to user and the converse, and changes in terminal characteristics which may influence (hopefully reduce) user delays. Indeed, recently there have been reported many investigations into the performance of scheduling algorithms as measured by response time[1,3,7,9,13,14,15,16,20]. Almost without exception, these investigations hypothesize arrival rates of requests for CPU time without the support of measurements. Since such arrivals can be approximated from the variables of the data stream model, the above observations as to the efficacy of the results reported in this paper are demonstrated.

<u>Second</u>, the particular distributions obtained in Table II are tractable and are useful in further analytical studies. For example, the geometric distribution was obtained for the discrete distributions and the gamma family for the continuous distributions.

Table III shows the coefficients of variation, V, for the continuous variables for the four computer systems investigated.* Since the exponential distribution belongs to the gamma distribution family and is the special case where V=1, for certain applications it may be possible to use the exponential distribution to

* For one system, the user terminal had an automatic response at the end of a computer burst segment rather than a true user "think time" response. For this system, the estimated value of V for the think time distribution was 0.72, close to that for the hyperexponential distribution (V=0.71).

describe the arrival and delay processes. To illustrate the similarity between the exponential distribution and the gamma distribution with $1.0 \leq V \leq 1.8$, Table IV is included.

TABLE III

COEFFICIENT OF VARIATION FOR GAMMA DISTRIBUTIONS

	Systems			
	A	B	C	D
Think Time	1.56	1.64	0.72	1.61
Idle Time	1.09	1.54	1.59	1.45
User Interburst Time	1.39	1.59	1.49	1.54
Computer Interburst Time	1.56	1.61	1.59	1.64
User Intercharacter Time	1.67	1.54	1.67	1.59
Computer Intercharacter Time	1.67	1.67	1.59	1.56

TABLE IV

DIFFERENCE BETWEEN CUMULATIVE DISTRIBUTIONS OF GAMMA AND EXPONENTIAL VARIATES

	Error in Percent* for Independent Variable at			
Gamma Coefficient of Variation	One-Half Mean Value	Mean Value	Twice Mean Value	Maximum Error in Tail
1.0	0.0	0.0	0.0	0.0
1.2	14.5	3.6	2.0	2.4
1.4	23.1	6.22	3.8	5.0
1.6	28.2	7.3	6.4	8.6
1.8	30.6	6.7	10.2	13.7

* With respect to gamma distribution.

The error listed in each column is the difference between the cumulative distributions at the point given. The last column lists the largest value of this error in the upper tail region. From the table, we can see that the approximation becomes less accurate as V becomes larger and is less accurate for smaller values of the independent variable than for larger ones. We further observe that even close to the origin, the class of gamma distributions defined by $V \geq 1$ has the same general shape as the exponential distribution. For much analytical work, the behavior of the distribution function in the neighborhood of the mean and in the upper tail are of the most interest. For these types of problems, if errors of the magnitudes shown in Table IV are allowable, (or alternatively if the relative coefficients of variation shown in Table III are tolerable; note that the coefficients of variations of Table III may be interpreted as relative to the exponential distribution, which has a coefficient of variation of unity) then the exponential distribution may be used in place of the gamma distribution.

Thus, assuming independent interarrivals, one can use the Poisson process to describe any of the arrival processes and have the large body of queueing theory at one's disposal to analyze the communication process of time-shared computer systems. Even for those distributions where the exponential interarrival approximation is not useful, the gamma distribution is tractable for some types of analyses.

Conclusions

In analyzing computer communication systems for time-sharing applications, the results of this work have shown that a variety of techniques can be applied to model the processes. Since the input traffic process has been characterized in terms that are usually tractable for analytical models, realistic results may be obtained using standard analytical techniques. In some models where the estimated distributional forms are not amenable to analysis, appropriate approximation techniques are available.

This work has shown that the communication process between a multiaccess computer and a user at a teletypewriter-like terminal can be represented by an elementary model from which more complex models may be constructed. Further, by using real data from operational multiaccess systems, we have shown that the model is robust and that the distributions obtained for each of the variables are tractable. In certain cases, the character arrival process can be approximated by a Poisson process. Thus, in modeling the communication process of long-holding-time multiaccess computer systems only the parameters of the distributions for the variables change for various computer types, applications and system loading.

These observations can be combined with the observations of Jackson and Stubbs[12] on computer-introduced delays, user-introduced delays and the relative amounts of information flow in each direction on the communication line to give a comprehensive picture of the communication process. For example, these analyses support analytical and simulation studies at Bell Telephone Laboratories which seek solutions to computer access data communications problems, cf. Chu's, Meltzer's and Pilc's studies on asynchronous time division multiplexing.[4,5,19,22]

Studies of multiaccess computer communications are continuing. Data are being collected from systems with different terminal types, system configurations, average holding times, and user applications. Analyses of data for these new systems will expand our understanding of the computer-communication processes involved as we have a broader base from which to draw conclusions and make comparisons.

Acknowledgments

Many people have contributed their efforts to various parts of this study. Data acquisition was accomplished with the considerable help of the American Telephone and Telegraph Company and the Bell System Operating Companies. Contributions to the model and the analyses and many helpful criticisms were made by Messrs. R. J. Roddy and C. D. Stubbs, of Bell Telephone Laboratories, and by Mr. R. J. Price, formerly of Bell Telephone Laboratories.

We would also like to thank Mr. E. Wolman of Bell Telephone Laboratories for carefully reading a draft of this paper.

Our special thanks are extended to the companies whose computer systems are being studied. Without their full permission and very helpful cooperation these analyses would not be feasible.

APPENDIX

A Fourier Series Test of Goodness of Fit

The examination of the data for goodness of fit posed considerable problems. The objective of this part of the study was to determine the suitability of analytically tractable probability density functions (p.d.f.'s) to describe the significant random processes.

The classical goodness-of-fit tests which were first applied in this study were the chi-square test, the Kolmogorov-Smirnov test, and the Cramer-Von Mises test. These tests suffer from the following deficiencies. The power of the chi-square test is very poor, and the realized significance level is sensitive to the number and placement of class intervals. The Kolmogorov-Smirnov test and the Cramer-Von Mises test require ordering the data, which requires considerable computer time even with the most efficient algorithms for the quantities of data involved in this study.* Further, since we are interested in the suitability at an acceptable significance level of analytically tractable p.d.f.'s, the tradeoff between significance level, power, and number of sample points is of concern to us. In this regard, the only available

---

* The advantage of the test used (described in the following paragraphs) over the Kolmogorov-Smirnov test or the Cramer-Von Mises test, in units of computer time, is roughly 25-50 to one.

expression for the power of the Kolmogorov-Smirnov test is a lower bound, and no expression for the power of the Cramer-Von Mises test is known at this time. An additional complication is what we need to perform a compound goodness-of-fit test where we estimate the parameters of the distribution from the data as well as testing a given distribution.

These difficulties led to the development of a new test by Jackson, called the Fourier Series Test of Goodness-of-Fit,[11] which has the following advantages:

1. No decisions about number and spacing of class intervals are required as with the construction of histograms required for the chi-square test;
2. The set of observations need not be ordered;
3. Estimators are easily updated for additions to the data base by recursive relationships which require only minimal operations on the new data;
4. The power of the test is comparable to the power of the best of the classical tests for reasonable alternatives;
5. The computation time for the Fourier test is less than the computation time for the classical tests; and

6. Using the limiting distributions, the power of the Fourier test can be computed analytically for both the simple and compound hypothesis tests, while it cannot be computed for most of the classical tests.

Briefly, the technique proceeds as follows: The probability density function is estimated from the data by a finite linear combination of sine and cosine functions harmonic over the region of support of the function - a truncated Fourier series - where the coefficients of the series are estimated from the data, and the number of terms of the series are determined by a minimization technique. Then, for each prespecified standard distribution, the hypothesis that the estimated Fourier series function is not significantly different from the Fourier series expansion of the p.d.f. of the standard distribution is tested. The test statistic used is a function of the squared differences between the coefficients of the Fourier series expansion of the estimated distribution and the coefficients of a Fourier series expansion of the hypothesized standard distribution.

REFERENCES

1. Bowden, E. K., Jr. (1966). "Priority Assignment in a Network of Computers", IEEE Computer Group Conference Digest, IEEE Catalog Number 69C30-C, June, 1969.

2. Bryan, G. E. (1965). "JOSS: 20,000 Hours at the Console, A Statistical Summary", RAND Corporation.

3. Chang, W. (1966). "A Queueing Model for a Simple Case of Time Sharing", IBM Systems Journal, Vol. 5, No. 2.

4. Chu, W. (1968). "A Study of the Technique of Asynchronous Time Division Multiplexing for Time Sharing Computer Communications", Proceedings of the 2nd Hawaii International Conference on System Sciences, January, 1969.

5. Chu, W. (1969). "An Analysis of Buffer Behavior for Batch Poisson Arrivals and Single Server with Constant Output Rate", unpublished work.

6. Coffman, E. G., and Wood, R. C. (1965). "Interarrival Statistics for TSS", System Development Corporation SP-2161.

7. Coffman, E. G., and Kleinrock, 1. (1968). "Computer Scheduling Measures and Their Countermeasures", 1968 Spring Joint Computer Conference, AFIPS Proceedings, Vol. 32, Washington, D.C.: Thompson, 1968.

8. Erlang, A. K. (1925). "Calcul des Probabilités et Conversations Telephoniques", Revue Generale, de l'Electricité, Vol. XVIII, August, 1925; previously published in a Danish periodical "Nyt Tidsskrift for Matematick".

9. Gaver, D. P., Jr., and Lewis, P.A.W. (1969). "Dynamic Buffer Storage Models", IEEE Computer Group Conference Digest, IEEE Catalog No. 69C30-C, June, 1969.

10. Hastings, T. (1965). "Operating Statistics of the MAC Time-Sharing System", Massachusetts Institute of Technology, Project MAC, Memorandum MAC-M-280.

11. Jackson, P. E. (1969). "A Fourier Series Test of Goodness of Fit", submitted for publication in Journal of the American Statistical Association.

12. Jackson, P. E., and Stubbs, C. D. (1969). "A Study of Multi-access Computer Communications", AFIPS-Conference Proceedings, Volume 34, p 491.

13. Kleinrock, L. (1964). "Analysis of a Time-Shared Processor", Naval Research and Logistics Quarterly, March, 1964.

14. Kleinrock, L. (1967). "Time-Shared Systems - A Theoretical Treatment", Journal of the Association for Computing Machinery, March, 1967.

15. Kleinrock, L. (1968). "Certain Analytic Results for Time-Shared Processors", IFIP Congress 1968, Hardware Vol. 1, Booklet D.

16. Krishnamoorthi, B., and Wood, R. C. (1966). "Time-Shared Computer Service with Both Interarrival and Service Times Exponential", Journal of the Association for Computing Machinery, July, 1966.

17. McIsaac, P. V. (1965). "Time-Sharing Job Descriptions for Simulation", System Development Corp., TM-2713.

18. McIsaac, P. V. (1966). "Job Description and Scheduling in the SDC Q-32 Time-Sharing System", System Development Corp., TM-2996.

19. Meltzer, J. I. (1968). "Simulation of an Asynchronous Time Division Multiplexor for Data Sources", unpublished work.

20. Miller, L. W. and Schrage, L. E. (1965). "The Queue M/G/1 with the Shortest Remaining Processing Time Discipline", RAND Corp. Report P3263, November, 1965.

21. Molina, E. C. (1922). "The Theory of Probabilities Applied to Telephone Trunking Problems", The Bell System Technical Journal, Volume I, 1922.

22. Pilc, R. J. (1968). "A Derivation of Buffer Occupancy Statistics in an Asynchronous Time Division Multiplexor Used with Bursty Sources", unpublished work.

23. Sackman, H. (1967). "Experimental Investigation of User Performance in Time Sharing Computing Systems: Retrospect, Prospect and the Public Interest", System Development Corp., AD 654 624, May, 1967.

24. Scherr, A. L. (1965). "JOSS: Experience with an Experimental Computing Service For Users at Remote Typewriter Consoles", P-3149, The RAND Corp.

26. Totschek, R. A. (1965). "An Empirical Investigation into the Behavior of the SDC Time-Sharing System", System Development Corp., SP-2191/000/00.

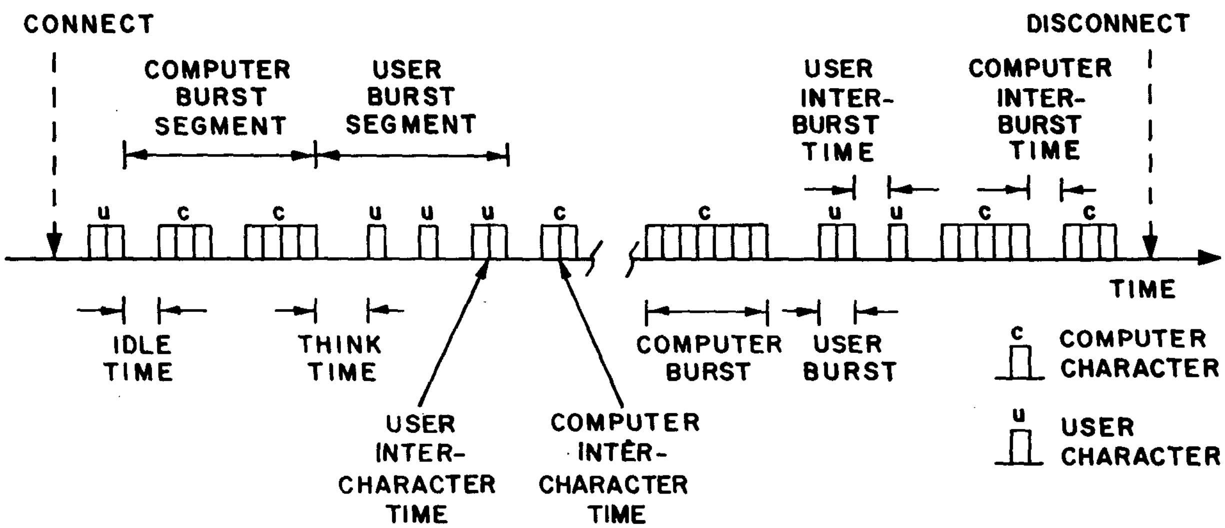

FIGURE I THE DATA STREAM MODEL