91

21537

Order Number 9121537

Data-analytic and monitoring schemes for a class of discrete point processes

Chandramouli, Yegnanarayanan, Ph.D.

The University of Arizona, 1991

U·M·I
300 N. Zeeb Rd.
Ann Arbor, MI 48106

DATA-ANALYTIC AND MONITORING SCHEMES FOR A CLASS OF DISCRETE POINT PROCESSES

by

Yegnanarayanan Chandramouli

A Dissertation Submitted to the Faculty of the

DEPARTMENT OF SYSTEMS AND INDUSTRIAL ENGINEERING

In Partial Fulfillment of the Requirements
For the Degree of

DOCTOR OF PHILOSOPHY

In the Graduate College

THE UNIVERSITY OF ARIZONA

1 9 9 1

THE UNIVERSITY OF ARIZONA
GRADUATE COLLEGE

As members of the Final Examination Committee, we certify that we have read the dissertation prepared by Yegnanarayanan Chandramouli

entitled Data Analytic and Monitoring Schemes for a Class of Discrete Point Processes

and recommend that it be accepted as fulfilling the dissertation requirement for the Degree of Doctor of Philosophy.

Marcel F. Neuts — 12/18/90
Marcel F. Neuts — Date

Sidney J. Yakowitz — 12/18/90
Sidney J. Yakowitz — Date

Paul J. Sanchez — 12/18/90
Paul J. Sanchez — Date

Robert S. Maier — 12/18/90
Robert S. Maier — Date

Date

Final approval and acceptance of this dissertation is contingent upon the candidate's submission of the final copy of the dissertation to the Graduate College.

I hereby certify that I have read this dissertation prepared under my direction and recommend that it be accepted as fulfilling the dissertation requirement.

Marcel F. Neuts — 12/18/90
Dissertation Director — Date

STATEMENT BY AUTHOR

This dissertation has been submitted in partial fulfillment of requirements for an advanced degree at The University of Arizona and is deposited in the University Library to be made available to borrowers under rules of the Library.

Brief quotations from this dissertation are allowable without special permission, provided that accurate acknowledgment of source is made. Requests for permission for extended quotation from or reproduction of this manuscript in whole or in part may be granted by the head of the major department or the Dean of the Graduate College when in his or her judgement the proposed use of the material is in the interests of scholarship. In all other instances, however, permission must be obtained from the author.

SIGNED: ____________________

ACKNOWLEDGEMENTS

I wish to express my sincere appreciation to my advisor and thesis supervisor, Professor M. F. Neuts, without whose persevering efforts and constant encouragement this work would not have been possible.

I also wish to thank the other members of my dissertation committee, Professors R. S. Maier, P. J. Sanchez and S. J. Yakowitz for several stimulating discussions and helpful suggestions.

I am greatly indebted to my wife Radhika and my parents for their support during this endevour.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

LIST OF TABLES

ABSTRACT

A point process model for the packet stream arising in teletraffic processes is the discrete, non-negative integer-valued, stationary process introduced by Neuts and Pearce. In this thesis, we examine an empirical approach for developing a monitoring scheme for that point process. Monitoring is a procedure of tracking a stochastic process to identify quickly the development of anomalous situations in the evolution of the process and detect their assignable causes. Further, a data-analytic scheme to evaluate the order of a Markov chain that quantifies the local dependence embedded in the point process and Walsh spectral techniques are examined.

CHAPTER 1

A POINT PROCESS MODEL FOR THE PACKET STREAM

1.1 Introduction:

Packet-switched communication networks are being developed for supporting a variety of services such as voice and data communication with a single high speed channel to transfer the packets from one end of the network to another. These different services induce heterogeneous traffic streams, each having specific individual characteristics and stringent performance constraints required for reliable transmission through the network. For instance, the possibility of dropping of packets and the variability in the transmission delay of the voice packets can lead to packet starvation in the destination node which results in the distortion of the reconstructed voice signal. Therefore, the efficient design of teletraffic systems requires as a first step, a detailed study of these different traffic streams for accurately predicting performance of the network. To support the performance analysis, plausible stochastic models for the packet stream are needed which incorporate its physics and capture the qualitative features that may have a bearing on the performance.

The packet stream arriving at a packet network is the pooled stream of the packets generated by several sources which are initiated at different time epochs. Each source has a finite holding time during which the source transmits a message and is comprised of an alternating sequence of active periods and inactive periods. During active periods of the source, packets are generated. The inactive periods are those during which the source is not generating a packet. The complexity of the aggregate packet stream is primarily due to the bursty nature of the packetized process generated by an individual source and the fluctuations in the number of active sources. From the point of view of modelling, it is important to take into account the finer information about the way the source is generating packets during its holding time.

Recently, significant research effort is being devoted to the performance modelling of packet communication systems. Queueing arises very naturally in the study of packet-switched networks: the packets generated by several active sources, arriving at the entry point of the network on the way to their respective destination nodes, are buffered, processed and sent through the transmission link by a statistical multiplexer. The multiplexer is modelled as a single server queue with the packet stream as its input. The time spent in the buffer by an arbitrary packet waiting for transmission and the long run fraction of packets lost are some measures of performance for the network. The most frequently used models to describe the packet stream are the Poisson process, the interrupted Poisson process and the Markov modulated Poisson process. The interrupted Poisson process, which has been used as a model for overflow streams Kaczura [8], is a Poisson process which is alternatingly switched between two arrival rates, of which one is zero, for independent exponential lengths of times. A Markov modulated Poisson process is a doubly stochastic Poisson process whose rate process is governed by a finite state Markov process. The primary appeal in using these models for the packet stream lies in the fact that, analytically tractable queueing models for the analyzing the performance of the network can be obtained.

Sriram and Whitt [23] study the cumulative covariance among the consecutive interarrival times of the aggregate packet arrival process by using the notion of the index of dispersion for intervals to characterize the process. From the general theory of point processes, it is well known that the index of dispersion provides a high degree of information about the correlations present in the process; refer to Cox and Lewis [4]. Further, in Sriram and Whitt [23], they have shown that for low traffic intensities of the queue the Poisson process will indeed serve as an adequate descriptor of the packet stream for the purpose of predicting overall packet losses and delays. For high traffic intensities, they confirm the inadequacy of the Poisson process and to account for the dependencies in the packet stream, they develop a renewal process approximation with a high coefficient of variation for the inter-arrival time distribution using moment matching methods. The performance measures of the resulting queue is obtained by heavy traffic queueing results. The argument for the use of Poisson process for low loads involves the application of the Palm-Khinchin theorem which establishes that the superposition of a large number of independent, identically distributed and *uniformly sparse* renewal processes will be indistinguishable from a Poisson process over short periods of time; for the precise statement and proof of the Palm-Khinchin theorem refer to Karlin and Taylor [9]. Heffes and Lucantoni [6] globally approximate the aggregate arrival process by a two state Markov modulated Poisson process (MMPP). The MMPP provides the means to accommodate with in one's model the fluctuations in the packet stream. The parameters of the MMPP were chosen so as to match several statistical properties of the packet stream, in particular, the index of dispersion curve for the cumulative number of packet arrivals in [0,t] over its entire range. Several system performance measures of the resulting queueing model are evaluated using the matrix analytic methods Neuts [15]. Recently, Ramaswami and Latouche [20] have proposed a model for the packet stream motivated by the need to model the fluctuating rate at which packets are generated within each message. The fluctuations are caused by the alternation of talk spurts and silent periods or by different operational modes within messages. In addition to the modulation of the arrival rate, there is a lower frequency modulation due to the varying number of messages active over time. There has been considerable amount of related work on the performance modelling of communication systems, and for further references on this subject we refer to Ramaswami [19].

Our goal in this chapter is to examine a discrete time point process model developed by Neuts and Pearce [18] for the aggregate packet stream, obtained as the limit process of the superposition of the packetized processes generated by a large number of sources, that takes into account in a general manner the statistical properties in the packet generation of a single source. The behavior of a single source is modelled as an irreducible, finite state, discrete time Markov chain, whose state space is partitioned into two subsets of states. At visits to the active set of states, either a *one* or *zero* is generated depending upon the state visited in that set. In the inactive set of states neither symbol is generated. The output process is a Markovian stream generated by a single channel. A sojourn in the active set of states corresponds to a message. The intervening sojourns in the inactive states represent the silent periods between messages. The transition probability matrix of the Markov chain governing a single source is defined so that the alternating periods spent in the active and inactive sets of states are independent.

1.2 Motivation:

During the active periods, the source is transmitting a message in a stream of packets (ones) and empty slots (zeros). It is clear that the empty slots occurring during a message are quite different from those of the silent periods. Mathematically that distinction is essential in examining the limit process obtained by superposition of point processes. A key condition required in the proof of the Palm-Khinchin theorem is that the *thinness* of the superimposed point processes is assured by *stretching* the time variable proportionately to the number of component processes superimposed. In models for the packetized traffic, while imposing the condition of thinness on the superimposed process, it is meaningful to assume that the silent periods between consecutive messages are long, but the empty slots occurring within a message are not stretchable. By stretching the time variable globally, the silent periods would be long, but so would the empty slots between packets for a given message, so that the active periods would be long too. The principal motivation for studying the limit process to be discussed here is that in analyzing the superposition of a large number of such point processes, it is essential to take the finer physical detail into account. The continuous-time analogue of the limit processes to be discussed here is studied in Neuts [16].

In modelling the source, we shall assume that the durations of the sojourns in the silent and the active sets of states have discrete PH–distributions with irreducible representations (β, S) and (α, T), respectively with $m(0)$ and $m(1)$ phases, where α, β are probability vectors and S, T are substochastic matrices. The means γ and μ of the silent and active periods are given by $\gamma = \beta(I - S)^{-1}\mathbf{e}$ and $\mu = \alpha(I - T)^{-1}\mathbf{e}$.

The column vectors $\mathbf{T}^{\circ}$ and $\mathbf{S}^{\circ}$ are defined by $\mathbf{T}^{\circ} = \mathbf{e} - T\mathbf{e}$ and $\mathbf{S}^{\circ} = \mathbf{e} - S\mathbf{e}$, where each $\mathbf{e}$ is a column vector of appropriate dimension with all its components equal to one. We shall assume the familiarity with the standard notation and basic properties of PH–distributions, as presented in Chapter 2 of Neuts [11].

The transition probabilities governing the i-th time slot in a single source are represented by the generating function matrix $P(\mathbf{z}(i))$ defined by

$$P(\mathbf{z}(i)) = \begin{vmatrix} S & \mathbf{S}^{\circ}\alpha\Delta(\mathbf{z}(i)) \\ \mathbf{T}^{\circ}\beta & T\Delta(\mathbf{z}(i)) \end{vmatrix},$$

where $\Delta(\mathbf{z}(i))$ is a diagonal matrix with the components of the vector $\mathbf{z}(i) = (z_1(i), \cdots, z_{m(1)}(i))$ as its diagonal elements. The role of each $\mathbf{z}(i)$ is in assigning identifying tags to the active states (whether each state is associated with a one or a zero symbol). For notational convenience, we shall write $\Delta(\mathbf{z}(k))$ as Δ_k whenever it is unnecessary to write the argument vector explicitly. The matrix $P = P(\mathbf{e})$ is an irreducible, aperiodic, stochastic matrix. The invariant probability vector of P is partitioned into an $m(0)$–vector $\mathbf{x}(0)$ and an $m(1)$–vector $\mathbf{x}(1)$. The vectors γ° and $\mathbf{x}$ are the stationary probability vectors of the matrices $S + \beta S^{\circ}$ and $T + \alpha T^{\circ}$ respectively and are given by

$$\gamma^{*} = \gamma^{-1}\beta(I - S)^{-1}, \tag{1}$$

$$\pi = \mu^{-1}\alpha(I - T)^{-1}, \tag{2}$$

and the constant

$$a = \frac{\mu}{\gamma + \mu},$$

which represents the steady-state fraction of time spent in the active set states.

Thin Sequences: A stream generated by a Markovian source is defined to be *thin* if the steady state fraction of time spent in the active set of states is small and the silent periods are long. In formalizing this notion, we shall require that for every k, the probability a silent period lasts k units of time $\beta S^{k-1}S^{*}$ is negligibly small. Formally, we consider a sequence of models in which the probability distribution of the length of active periods remains fixed, but that of the silent periods for every positive integer k

$$\beta_{\nu} S_{\nu}^{k-1} S_{\nu}^{*} = o(1) \quad \text{as} \quad \nu \to \infty. \tag{3}$$

It is readily seen that this implies that the means γ_{ν} tend to infinity, and therefore that the fractions of time spent in the active states are small and satisfy

$$a_{\nu} = o(1). \tag{4}$$

1.3 The Superposition Theorem:

Let us now consider the superposition of N stationary independent and identically distributed streams generated by the Markovian sources discussed in the preceding section.

Theorem 1: For the superposition of N independent and stochastically identical Markovian streams which satisfy the thinness condition (4) and $a = N^{-1}\lambda + o(N^{-1})$, for some positive constant λ, the joint probability generating function of the numbers of times all active states are visited at the time points $0,1, \cdots, n$ converges as $N \to \infty$ to

$$\Psi[\mathbf{z}(0), \cdots, \mathbf{z}(n)] = \exp\{-\lambda[1 - C_n] - \lambda\mu^{-1}\sum_{k=1}^{n} [1 - D_{k,n}]\}, \tag{5}$$

where C_n and $D_{k,n}$ are given by

$$C_n = \pi\Delta_0\mathbf{T}^{\circ} + \pi\Delta_0 T\Delta_1\mathbf{T}^{\circ} + \cdots + \pi\Delta_0 T\Delta_1 \cdots T\Delta_{n-1}\mathbf{T}^{\circ} + \pi\Delta_0 T\Delta_1 \cdots T\Delta_n\mathbf{e},$$

$$D_{k,n} = \alpha\Delta_k\mathbf{T}^{\circ} + \alpha\Delta_k T\Delta_{k+1}\mathbf{T}^{\circ} + \cdots + \alpha\Delta_k T\Delta_{k+1} \cdots T\Delta_{n-1}\mathbf{T}^{\circ} + \alpha\Delta_k T\Delta_{k+1} \cdots T\Delta_n\mathbf{e}.$$

This theorem was established in Neuts and Pearce [18] and we shall omit the proof here.

The limit process obtained here by the superposition of indefinitely many thin Markovian streams, is a discrete parameter, non-negative integer-valued stationary stochastic process whose joint distribution is given by formula (5) is a discrete time point process model for the packet stream. This process can be constructively defined as follows:

Theorem 2: The point process model for the packet stream for a given λ and a discrete phase type distribution with representation (α,T) is stochastically equivalent to the following process:

The number of messages initiated at time 0 has a Poisson distribution with mean λ.

Each message in progress at time 0, chooses an initial state among the active states $1,\cdots,m(1)$ according to the stationary vector π and is allowed to evolve, independently of other messages, through a set of transient states of an absorbing Markov chain until it ends when absorption occurs. That is the probability distribution of the duration of each message is a discrete phase type distribution with representation (π,T).

In addition to the messages started at time 0, independent random numbers of new messages are initiated at epochs $k \geq 1$. The number of the new messages have independent Poisson distributions with parameter $\lambda\mu^{-1}$. The duration of each new message has a $PH-$ distribution with parameters (α,T).

Proof: Let us consider a message started at time 0 according to the vector π. The joint probability generating function of the labels of the transient states visited (if any) at the time points $0, 1, \cdots, n$, is given by C_n. The first $n-1$ terms of C_n correspond to the alternatives where the message ends before time n; the n–th term is the contribution of the case where the message is still alive at time n.

Since a Poisson number of independent messages with parameter λ are initiated at time 0, the joint probability generating function of the counting variables of all transient states visited at times $0, 1, \cdots, n$ is given by

$$\Phi_0[z(0), \cdots, z(n)] = \exp\{-\lambda[1 - C_n]\}. \tag{6}$$

By a similar argument, we see that the joint probability generating function of the labels of transient states visited at times $k, \cdots, n$ by a Poisson number of messages with parameter $\lambda\mu^{-1}$ started at time epochs $k \geq 1$ is given by

$$\Phi_k[z(0), \cdots, z(n)] = \exp[-(\lambda/\mu)[1 - D_{k,n}]]. \tag{7}$$

The independence assumptions underlying this construction readily imply that the joint probability distribution of all counts of the labels of the transient states visited in $0, \cdots, n$ is given by

$$\prod_{k=0}^{n} \Phi_k[z(0), \cdots, z(n)] = \Psi[z(0), \cdots, z(n)]. \tag{8}$$

Remark:

For applications of the limit process to describe packet streams, the counts of visits to all the transient states of the absorbing Markov chain governing the duration of a message are usually not of interest. Each state in the transient set of states has a label, either a one or a zero, depending on whether a visit to that state generates a packet or not. We are mostly interested in the joint probability generating function of the numbers of packets that arise at the epochs $0, \cdots, n$. That probability generating function is immediately obtained by special choices of the vectors $\mathbf{z}(k)$, by setting each component of $\mathbf{z}(k)$ equal to z_k or to 1, according to whether the corresponding state generates a packet or not. Therefore, in the calculations in the remainder of this chapter, we shall denote $\Delta(z_k)$ for the diagonal matrix with diagonal elements which are z_k at the indices of states where a packet is generated and 1 where there is not. The derivative Λ of $\Delta(z_k)$, is a matrix which has the diagonal element 1 at the indices of states where packets are generated and 0 at the indices of the other states. The formula

$$\Delta(z) = I - (1-z)\Lambda. \tag{9}$$

is useful in the sequel. The constant $\xi = \pi\Lambda e$ is the expected number of packets generated by a message active at an arbitrary time slot.

Theorem 3: The joint probability generating function $H(z_0,z_n)$ of the numbers N_0 and N_n of packets generated at time 0 and n in the limit process is given by

$$H(z_0,z_n) = \exp\{-\lambda\xi[(1-z_0)+(1-z_n)-\rho(n)(1-z_0)(1-z_n)]\}, \tag{10}$$

where $\rho(n) = \xi^{-1}\pi\Lambda T^n\Lambda e$ is the *correlation coefficient* of N_0 and N_n.

Proof: In the expression for C_n, by setting $\Delta_k = I$ for $0 < k < n$ and using (9) for k=0, n, we obtain

$$C_n = 1 - \pi\Lambda e(1-z_0) - \pi T^n \Lambda e(1-z_n) + (1-z_0)(1-z_n)\pi\Lambda T^n \Lambda e\ .$$

Similarly the formula for $D_{k,n}$ simplifies as

$$D_{k,n} = 1 - (1-z_n)\alpha T^{n-k}\Lambda e,$$

so that

$$\sum_{k=1}^{n}(1-D_{k,n}) = (1-z_n)\mu\pi(I - T^n)\Lambda e\ .$$

Substitution in (5) and simplifications yields (10).

Routine derivations show that $\rho(n)$ is indeed the correlation coefficient of N_0 and N_n. We note that the correlation coefficient $\rho(n)$ of N_0 and N_n does not depend on λ and is always nonnegative. The number of packets generated at an arbitrary time clearly has a Poisson distribution of mean $\lambda\xi$. To conclude, we note that the class of limit processes considered here as a model for the packet stream, consists of a family of stationary, non-negative integer-valued, discrete parameter processes of dependent Poisson random variables.

CHAPTER 2

SIMULATING THE POINT PROCESS

2.1 Introduction:

For the point process model for the packet stream discussed in Chapter 1, there are few analytical results or algorithmic methods to obtain qualitative information about the process. This is primarily due to the non-Markovian dependence present in the process. Given the paucity of analytically tractable methods for the point process, the design of monitors for such models calls for novel methodological approaches. Guided by results on the small class of models that remain algorithmically tractable, an investigation in a data-analytic spirit appears to be most promising. Simulation methodology, often an alternative to analytic and algorithmic methods, provides the necessary tool for an empirical study on stochastic models. The need for careful computer experimentation as a legitimate source of obtaining information about stochastic models is stressed in Neuts [14].

Our objective in this Chapter is to describe a procedure for simulating the discrete, non-negative integer-valued, stationary point process of dependent Poisson random variables studied as a model for the packet stream. Using the output data obtained from the simulation of the process, we examine data-analytic, spectral, and monitoring schemes for the point process in the subsequent chapters. The constructive definition of the process given in Theorem 2 of Chapter 1 is useful from the point of view of simulating the point process.

Towards our goal of simulating the process, as a first step, we describe an algorithm presented in Neuts and Pagano [17], to simulate the number of time epochs and the sequence of states visited until absorption in a discrete time, absorbing Markov chain with finitely many states. Generating a random variate from a discrete probability distribution with finite support is a basic tool needed for the purpose of simulating Markov chains. An efficient and fast algorithm for generating discrete random variates is the alias method proposed by Walker [24] and refined by Kronmal and Peterson [10]. The alias method is based on the following key result:

Theorem: Any discrete distribution with density $p(i)$, $1 \le i \le n$, with finite support (n) can be expressed as a equiprobable mixture of n distributions $q_i(.)$, $1 \le i \le n$, where each distribution $q_i(.)$ supported on two points, in such a way that i is a mass point of $q_i(.)$.

The Alias Algorithm:

The alias algorithm to generate a discrete a random variable X with density $p(i)$, $1 \le i \le n$, in the notation of Kronmal and Peterson [10] is as follows:

a. Generate U uniform on $(0,n)$.

b. Set $I \leftarrow \lceil U \rceil$.

$\lceil U \rceil$ refers the smallest integer greater then U.

Now I is uniform on $\{1, 2, \cdots, n\}$.

c. Set $U \leftarrow I - U$.

U is uniform on (0,1) and independent of I.

d. Set

$$X \leftarrow \begin{cases} I & \text{if } U \le F_i \\ L_i & \text{if } U > F_i . \end{cases}$$

F_i and L_i are constants which depend on the given density $p(.)$. The fractions F_i, $1 \le i \le n$ can be described as the cutoff values since they are compared with U and L_i, $1 \le i \le n$ can be described as the the aliases to which the I is transformed if the comparison fails (i.e. if rejection occurs). Define $F_i = q_i(i)$ and L_i = the mass point of $q_i(.)$ that is not i, where $q_i(.)$ are the elements of the equiprobable mixture.

To implement the alias method, initially we have to calculate two arrays of length n each from the given discrete density $p(.)$. The first array contains the cutoff values F_i and the second array gives the aliases L_i for $i = 1, 2, \cdots, n$. An explicit algorithm for computing the table of cutoff values and the aliases from $p(.)$ is given in Kronmal and Peterson [10]. The generation of the table of cutoff and alias values is performed once and the number of operations to calculate the table is proportional to n. The most important feature of the alias method is that, for generating one random variate from a discrete distribution, one comparison and at most two table lookups are needed. Even though the table-generation part is an overhead, this method is particularly suited for generating large number of random variates from discrete distributions.

2.2 A Procedure to Simulate a Discrete Time Absorbing Markov Chain:

Consider a discrete phase type distribution with representation (α, T) and the associated Markov chain with transition probability matrix P given by

$$P = \begin{vmatrix} T & T^* \\ & \\ 0 & 1 \end{vmatrix} .$$

T is a substochastic matrix of order m and α is a row vector of dimension m. The vector e is a column vector with all its components equal to 1. The vector T° is given by $T^\circ = e - Te$, and $\alpha e + \alpha_{m+1} = 1$. Let T_i be the i–th row of the matrix T and T_i° be the i–the component of the vector T°. The alias method is used to generate the initial state from the initial probability vector (α, α_{m+1}) and the subsequent transitions in the Markov chain P from the rows (T_i, T_i°), $1 \le i \le m$. The simulation algorithm to generate a discrete phase type random variable Y and a sample path $(x_0, x_1, x_2, \cdots, x_Y)$ of the Markov chain P until absorption with initial probability vector (α, α_{m+1}) is as follows:

Step 1:

The table of aliases and the cutoff values for (α, α_{m+1}) and for (T_i, T_i°), $1 \le i \le m$ are computed. Let Y, the number of transitions before absorption occurs in the Markov chain P, be initialized to 0.

Step 2:

a. At time 0, choose the initial state I from (α, α_{m+1})

$x_0 \leftarrow I$

b. If I is the state m+1, then go to d

else

c. At time epochs $n \geq 1$,

Choose the next state J from (T_I, T_I^*)

If J is the state m+1, then go to d

else

$Y \leftarrow Y + 1$

$I \leftarrow J$

$x_n \leftarrow I$

Return to c

d. The Markov chain is absorbed

Return the value of Y

Return $(x_0, x_1, x_2, \cdots, x_Y)$

This algorithm is implemented in FORTRAN and to gain efficiency, our data structures take advantage of the the standard FORTRAN features such as *column major ordering* and *call by reference.* The matrix P is stored in its transposed form of dimension $(m+1) \times (m+1)$, so creating a matrix with stochastic columns. The m+1–st column contains the vector (α, α_{m+1}). Throughout this work, uniform variates on (0, 1) are generated using a portable random number generator devised by Schrage [22]. In our implementation, to avoid serial dependencies in the high-dimensional structure of the successive uniform random variates, we have used 62,089,911 as the multiplier in the linear congruential random number generator as suggested in Lewis and Orav [7]. It should be noted that this multiplier is different from the multiplier 16,807 recommended in Schrage [22]. This procedure to simulate the transitions from the evolution of a absorbing finite Markov chain forms a basic ingredient in our algorithm for simulating the point process.

2.3 A Procedure to Simulate the Point Process:

The characterization of the stationary version of point process given in Theorem 2 of the previous chapter, is primarily useful from the point of view of simulation. In particular, the initial conditions for the simulation can be selected according to the explicit steady state description of the point process. By simulating the stationary version of the process, we can avoid the considerations due to initialization effects in our simulation runs. The process is governed by the parameters λ, the arrival rate of messages and the duration of each message which is of phase type with representation (α, T) of order m. For simplicity, we have assumed that $\alpha e = 1$. The mean of the phase type distribution with representation (α, T) is given by $\mu_1' = \alpha (I - T)^{-1}$ e. π is the invariant probability vector of the stochastic matrix Q, where $Q = T + T^{\circ}\alpha$. The matrix T is partitioned into two disjoint sets of states A and B. At visits to the set of states A packets are generated and empty slots are generated during sojourns in the set of states B. A realization of the point process is characterized by specifying Y_n, $n \geq 0$, the counts of packets generated which is equivalent to the number of Markov chains that are in the set A of transient states at time epoch n.

In the simulation, we maintain one counter for each of the $m+1$ states of the Markov chain P. Let $N_n(j)$ be defined as the number of Markov chains in state j, $1 \le j \le m+1$ at time $n \ge 0$. Let I_A denote the indicator function of the set A which takes the value 1 if the Markov chain is in the set A and 0 otherwise. A summary of the procedure to simulate the point process is as follows:

Step 1:

a. At time epoch 0, a Poisson random variable M_0 with parameter λ is generated.

b. M_0 independent random variates from the discrete density $(\pi, 0)$ are generated. Based on those M_0 variates, the vector $N_0(j)$, $1 \le j \le m+1$ is initialized.

c. $Y_0 = \sum_{j=1}^{m} I_A \, N_0(j)$.

Step 2:

a. At time epochs $n \ge 1$, each Markov chain that is not absorbed at time $n-1$ is allowed to evolve one transition, where each transition is governed by the transition probability matrix P.

b. According to the states visited by the active Markov chains at time n, the vector $N_{n-1}(j)$, $1 \le j \le m+1$ is updated. The updated vector is denoted as $\mathbf{N}'_{n-1}$.

c. At each time epoch $n \ge 1$, a Poisson random variate M_n with parameter $\lambda(\mu_1')^{-1}$ is generated.

d. M_n independent random variates from the discrete density $(\alpha, 0)$ are generated and stored as a vector $\mathbf{u}$. The vector $\mathbf{u}$ contains the initial states of the M_n Markov chains started at the time epoch n. By adding the vectors $\mathbf{N}'_{n-1}$ and $\mathbf{u}$ the final updated vector $\mathbf{N}_n$ is obtained.

e. For $n \ge 1$, $Y_n = \sum_{j=1}^{m} I_A \, N_n(j)$.

f. The procedure in Step 2 is repeated until the process is simulated for a specified duration of time.

The alias method is used to draw random variates from the discrete densities $(\pi, 0)$ and $(\alpha, 0)$. Updating the vector $N_{n-1}(i)$, $1 \le i \le m+1$ from time epoch $n-1$ to n, is an important module in this algorithm and this can be done efficiently. To that effect the algorithm to simulate the sequence of states visited by a Markov chain until absorption and the alias method to generate a random variate from a discrete distribution with finite support are used. For each $N_{n-1}(i)$, $1 \le i \le m$, the number of Markov chains which are in state i at time $n-1$, a vector containing the next states visited at time n is computed. This can be accomplished, for each i, by drawing $N_{n-1}(i)$ independent random variates from the discrete density (T_i, T_i^*). As a result, a table S_{ij} of dimension $m \times m+1$ is obtained in which the ij–th element of the table denotes that, of the $N_{n-1}(i)$ active Markov chains at time $n-1$, the number of Markov chains in state j at the next time unit n is given by S_{ij}. By summing over each column we can obtain the vector $\mathbf{N}'_{n-1}$, and the i–th component $N'_{n-1}(i)$, $1 \le i \le m+1$ represents the number of Markov chains in state i at time n.

The algorithm for generating a Poisson random variable is based on the following relation between the Poisson and the exponential distributions with parameters λ and λ^{-1} respectively. Let $Z_1, Z_2, \cdots$ be a sequence of independent and identically distributed random variables and let $X = \max\{\, i : \sum_{j=1}^{i} Z_j \le 1 \,\}$. Then the distribution of the Z_i's is exponential (λ^{-1}) if and only if the distribution of X is Poisson (λ).

Using the algorithm described in this chapter, we simulate the stationary version of point process for N time units where N is chosen to be a large number. In our experiments, to ensure that we have sufficient data for the empirical studies to be described in the subsequent chapters, we have chosen N to be 200,000. The output of the simulation contains a record $(Y_0, Y_1, Y_2, \ldots, Y_N)$ of $N+1$ observations from the point process.

CHAPTER 3

AN APPROACH FOR QUANTIFYING LOCAL DEPENDENCE

3.1 Introduction:

One of the point process models for the packet stream is the discrete, non-negative integer-valued, stationary process of dependent Poisson random variables. A commonly used measure to quantify the dependence present among random variables is the correlation coefficient. For the point process considered here, the correlation coefficient $\rho(n)$ of N_0 and N_n, the number of packets generated at an arbitrary time and n time units later is derived in Chapter 1 and is given by

$$\rho(n) = \xi^{-1} \pi \Lambda T^n \Lambda e.$$

In a different approach to characterize the local dependence in the process, a procedure to determine the order of a Markov chain that captures the some of the local dependence present in the point process, is examined. A sequence of random variables $\{X_\nu, \nu \geq 1\}$ is a k–th order Markov chain if for integers n, k and $k < n$

$$P(X_n \mid X_1, X_2, \cdots, X_{n-1}) = P(X_n \mid X_{n-k}, X_{n-k+1}, \cdots, X_{n-1}). \tag{11}$$

Mathematically, the k–th order Markov chain is a vector generalization of the simple Markov chain in which the present state X_n depends on the vector consisting of k past states $X_{n-k}, \cdots, X_{n-1}$. The case $k = 1$ corresponds to the ordinary Markov chain. The order of a Markov chain is a useful notion in assessing dependence, in particular, in quantifying how the successive random variables are temporally related. Further, the order of a Markov chain contains information on how much of the recent history of the process has an impact on the present, which can be described as the memory retained in the evolution of a stochastic process. A classical reference on the theory of statistical inference of Markov processes is Billingsley [2]. A procedure to evaluate the order of a Markov chain from biological signals recorded from the human heart or brain is outlined in Destexhe [5].

The data obtained from the simulation of the point process can be regarded as the observations from the evolution of a discrete, non-negative integer-valued, stationary time series. In this chapter, we describe a data-analytic procedure, guided by standard statistical techniques, to determine the order of a Markov chain embedded in the data. Towards the goal of fitting a Markov chain, it is desirable to work with a smoothing of the data which induces the concept of states of a stochastic process, yet retains information regarding the dependence and persistence present in the process. A smoothing of the data is obtained as follows: a. Instead of recording the value of the process at each time epoch, we observe the value of the process; decide whether that value can be classified uniquely as one of l distinct levels according to some criterion; record the symbol corresponding to that level.

The resulting sequence of symbols can be described as the states visited by a simplified stochastic process. The criterion for fixing the levels should reflect some physical meaning associated with the states. One approach to fixing the levels is based on the result proved in Chapter 1, that the marginal distribution of each random variable in the point process is a Poisson distribution with parameter $\lambda\xi$. Selected percentiles of that Poisson distribution, serve as the cutoff values for fixing the l levels. Let $Y_1, Y_2, \cdots, Y_N$ denote data obtained from the simulation of the point process and $p_1, p_2, \cdots, p_{l-1}$ denote the selected percentile points of the Poisson distribution with parameter $\lambda\xi$. X_n for $n \geq 0$ is defined as follows:

$$X_n = \begin{cases} 0 & \text{if } 0 \leq Y_n < p_1 \\ 1 & \text{if } p_1 \leq Y_n < p_2 \\ \cdot & \cdot \\ l-1 & \text{if } Y_n \geq p_{l-1}. \end{cases}$$

The sequence of symbols $X_1, X_2, \cdots, X_N$ so obtained contains information about the thresholds encountered sequentially and the persistence of various patterns among the symbols. Using the sequence of symbols, we empirically estimate the statistical properties of patterns among the symbols.

3.2 A Procedure to Determine the Order of a Markov Chain:

A word of length n is the contiguous occurrence of n symbols $X_1 X_2 \cdots X_n$ where each symbol is chosen from the set $\{0, 1, \cdots, l-1\}$. The probability of observing the word $X_1 X_2 \cdots X_n$ among the sequence of symbols is $P(X_1 X_2 \cdots X_n) = p_i$ for $i = 1, 2, \cdots, l^n$. The number of words of length n that can be formed by selecting n symbols with replacement from a set of l symbols is l^n. The subscript i refers to the i-th word among the l^n words of length n which are sorted in lexicographic order.

Let $P^{(k)}(X_1 X_2 \cdots X_n)$, $1 \le k < n$, denote the probability of observing the word $X_1 X_2 \cdots X_n$ by assuming that the word is generated by a k-th order Markov chain and is given by $p_i^{(k)}$ for $i = 1, 2, \cdots, l^n$. By using the Markov property, it can be verified that

$$P^{(k)}(X_1 X_2 \cdots X_n) = P(X_1 X_2 \cdots X_k)\, P(X_{k+1} \mid X_1 X_2 \cdots X_k)$$
$$P(X_{k+2} \mid X_2 \cdots X_{k+1}) \cdots P(X_n \mid X_{n-k} \cdots X_{n-1}) \quad (12)$$

The case $k = 0$ corresponds to the word $X_1 X_2 \cdots X_n$ generated by independent trials and the probability of that word is given by

$$P(X_1 X_2 \cdots X_n) = P(X_1)\, P(X_2) \cdots P(X_n).$$

The probabilities p_i and $p_i^{(k)}$ for $1 \le i \le l^n$, $1 \le k < n$ are empirically estimated by the frequencies f_i and $f_i^{(k)}$. The frequency f_i for each i can be evaluated from the sequence of symbols by using the formula

$$f_i = \frac{\text{number of words of type } i}{\text{total number of words of length } n}. \tag{13}$$

The frequencies $f_i^{(k)}$, for each k, $1 \le k < n$, are calculated by using the appropriate frequencies and conditional frequencies as estimates of the probabilities in relation (12). In order to calculate $f_i^{(k)}$ for a fixed k by using (12), it is readily seen that the frequencies of all words of length k and $k+1$ are needed. Hence, a basic module required in our scheme is an efficient algorithm of determining the counts of words of length m, $1 \le m \le n$.

For a fixed value of m, the counts of each of the l^m words of length m can be obtained by a search procedure which incorporates a moving window scheme. To keep track of the counts of l^m words among the sequence of symbols, a counter for each word is maintained. Each word, a string of m symbols selected from the set $\{0, 1, \cdots, l-1\}$, can be regarded as a number to the base l. Associated with each word is an identifier obtained by converting that string of m symbols to a decimal number. As initialization, we choose a window of length m which consists of the first m symbols $X_1, X_2, \cdots, X_m$ among the sequence of symbols. The identifier a for the word $X_1 X_2 \cdots X_m$ is given by

$$a = X_1\, l^{m-1} + X_2\, l^{m-2} + \cdots + X_m\, l^0.$$

The counter corresponding to this word referred to by the identifier a is incremented by 1. The next word is selected by moving the initial window and the next window contains the symbols $X_2, X_3, \cdots, X_m, X_{m+1}$. The identifier b for the next word $X_2 X_3 \cdots X_m X_{m+1}$ can be updated from the identifier of the previous word and b is given by

$$b = l\, [\, a - (X_1\, l^{m-1})\,] + X_{m+1}.$$

We increment the counter for the word associated with the identifier b by 1. This procedure is repeated until all $N - m + 1$ words of length m among the sequence of symbols are accounted. Once the counts of all the words are determined, their frequencies can be obtained by using formula (13). The principal advantages of this scheme are that, during the search for various words, the sequence of symbols is scanned through only once and only the words appearing within the sequence of symbols are searched and counted. Further, for each window chosen, the identifier for that word is efficiently updated.

In order to determine value of k, the order of a Markov chain, that quantifies the dependence within the various words, we compare the two discrete distributions f_i and $f_i^{(k)}$, $i = 1, 2, \cdots, l^n$, for increasing values of k. The statistic we have used is analogous to the χ^2 statistic, and is given by

$$\chi^2 = \sum_{i=1}^{l^n} \frac{[f_i - f_i^{(k)}]^2}{f_i}.$$

In our experiments, while computing this statistic, if the value of f_i is smaller than 0.001 for some i, then the adjacent cells of i are lumped until this condition is satisfied. This condition ensures that exceptional values of the statistic are not caused by the contribution of the terms corresponding to rare events. A similar rule of thumb is also widely adopted in the classical goodness of fit statistical testing situations primarily to obtain better asymptotic distributional properties of the test statistic. Another statistic analogous to the χ^2 statistic is

$$\delta = \frac{1}{l^n} \sum_{i=1}^{l^n} \frac{|f_i - f_i^{(k)}|}{f_i},$$

and δ is a measure of the average relative error between the distributions. The test statistic is plotted versus the order of the Markov chain. Empirical evidence of convergence of the test statistic to 0 for values of $k \geq k_0$, is our criterion to decide on a value of the order of the Markov chain.

3.3 Numerical Results:

In our computer experiments, we have chosen for simplicity to work with three levels low, medium and high. To fix the three levels, the 20–th and the 80–th percentile points of the Poisson distribution with parameter $\lambda\xi$ were used as the cutoff values for deciding low and high levels respectively. In order to ensure that there are enough data to obtain good estimates of the frequencies f_i and $f_i^{(k)}$, the point process is simulated for long durations of time typically up to 200,000 time points. Caution must be exercised in choosing the value of n, the length of the word, since the number of all possible words to consider grows geometrically as n increases. Based on computational experience, we have chosen n to be 9 which is large enough for the purpose of illustrating the methodology discussed here.

After extensive experimentation with various parameters for generating the point process, in most cases we have considered, the results of our study reveal that the order of a Markov chain that can be deduced from the statistic is always greater than 1 and at most 4. The consistency of this result is further empirically verified, by replicating the simulation experiment several times for different random number seeds for a given set of parameters of the process. This result indicates that, for the point process, locally within a word, the memory of any event extends several steps backward in time. As a numerical example, the χ^2 statistic versus the order of a Markov chain is displayed in Table 1. The significant frequencies f_i are sorted and the corresponding frequencies $f_i^{(k)}$ for $k = 1, \cdots, 4$ are displayed in Table 1. A close inspection of the frequencies in Table 1 supports our observation deduced from the test statistic. In Table 1, it is remarkable to note that for each fixed i the frequency $f_i^{(k)}$ uniformly tends to f_i as k approaches 4. Firstly, of the 3^n all possible words of length n, only a small fraction of words appear with significant frequencies among the sequence of symbols. This suggests that some words are encountered more frequently than some others. Further, among the words that have significant frequencies, a systematic pattern is often exhibited by the symbols that constitute those words. Such a pattern indicates that there is high persistence of symbols within the words. The assignable causes for that behavior can be traced to the dependence present in the process and further evidenced the preference of the chance mechanism governing the word generation towards words of systematic word configurations. Secondly, the frequencies of words describing the up-crossings and down-crossings of levels are nearly equal. This result is an empirical verification of a conservation law which states that in the stationary version of the process, the frequencies of up-crossings and down-crossings of the levels are equal. Such a result is also useful in confirming the accuracy of our numerical results. In addition to the conservation law, those systematic level crossings can be attributed to the asymptotic behavior of the point process. It is our conjecture that under suitable shifting and rescaling conditions, asymptotically the point process converges in distribution to a standard Gaussian stochastic process, a process whose level crossing properties are well understood. While these comments are related to a specific numerical example, the qualitative conclusions we have drawn are consistent with many other examples for which we have performed numerical computations.

In conclusion, we first note that the dependence information obtained from the order of a Markov chain is complementary to the correlation coefficient. In this approach, the joint distributional properties of each word of finite length n is empirically determined from data and a Markov chain of some order less than n provides an approximation to the distribution of that same word. However, it is important to stress that the order of a Markov chain is a reduction of local information only within a word of finite length n. Our study does not imply that such a reduction to a lower order Markov chain holds for a sequence composed of more than n symbols. These questions require an investigation to determine a higher order Markov chain that globally approximates a realization of the process. We have not yet undertaken such a study and that would require entirely different methods. The empirical study to evaluate the order of a Markov chain, as discussed here, provides interesting insights on some aspects of the local dependence and structural properties of the point process.

CHAPTER 4

A MONITORING SCHEME FOR THE POINT PROCESS

4.1 Introduction:

One of the fundamental challenges facing the performance evaluation of packet-based information transport, besides the modelling aspect, is determining approaches to efficient congestion and flow-control strategies. Towards that objective, we examine a procedure to detect the situations in the evolution of the model for the packetized traffic that are statistically incompatible with the proposed model. These situations require control measures which may depend upon how those situations had developed over time. Monitoring is a novel approach towards detecting the occurrence of such anomalous situations. Our goal in this chapter is to develop a monitoring scheme for the discrete, non-negative integer-valued, stationary point process proposed in Neuts and Pearce [18] as a model for the packet stream.

Monitoring can be defined as the procedure of tracking a random process so as to guarantee early identification of anomalous situations, which could be caused by undesirable changes in parameters of the process. Once such a change is detected, control actions can be implemented depending upon how that situation arose and how it is likely to evolve in the future. The occurrence of such anomalous situations can be due to rare random fluctuations, which remain statistically compatible with the posited model for the process, or could be indicative of systematic changes in the parameters of the process. While, in some cases, short-lived interventions in response to random fluctuations may be necessary, the most important purpose of monitoring is the prompt identification of truly anomalous situations that are no longer adequately described by the posited model. The importance of monitoring stochastic phenomena is stressed in Neuts [13], while a detailed study of monitoring a $M/G/1$ queue with group arrivals is described in Neuts [12].

As a first step towards monitoring the point process, a threshold K, the monitoring level, is fixed. It is often desirable that the criterion for fixing the threshold is chosen so that the monitor is infrequently actuated as long as the process operates within certain specifications. Further, since recognizing overload conditions and detecting their assignable causes quickly are the purposes of monitoring, the monitoring level should be selected based on the local information about the process instead of long run average behavior. Once such a level is chosen, the monitor is not activated as long as the process stays within that threshold. If the process exceeds the level K the monitor is activated and then the monitor keeps track of both the amount and the duration of exceedance above the threshold.

In addition to fixing a monitoring level, for the excursions above the threshold, we define the profile curves for the point process. The notion of profile curves was introduced by Neuts [12] for single-server queues, as a means to identify the excessive excursions in the queue length above a fixed threshold. Profile curves can be described as the upper stochastic envelope which captures the path functions that have exceeded the level K and are considered to be statistically compatible with the process. These curves would serve in classifying the excessive excursions as soon as possible and in formalizing monitoring schemes for stochastic processes.

Therefore, as long as the process stays between the level K and the profile curve monitoring continues. If the process returns back to the level K, before a specific time has elapsed, then that excursion is attributed to a random fluctuation in the process and the monitor is turned off. Alternatively, should the process exceed the profile curve or fail to return back to the level K soon enough, then the process is declared out of control and an appropriate control action is initiated. In the remainder of this chapter, we discuss the design of a monitoring level and develop a procedure to determine the profile curves for the Neuts-Pearce point process. Further, we derive some performance measures for the monitoring scheme.

Towards the goal of developing a monitoring scheme for the point process few analytical results can be derived. Given that the point process is indeed a good model for the packet stream and motivated by the paucity of analytically tractable results for the point process, an investigation of a monitoring scheme in a data-analytic spirit appears to be most promising. Guided by the algorithmically tractable methodology developed for monitoring a class of single-server queues in Neuts [12], a monitoring scheme for the point process, based on empirical methods is examined.

The data from evolution of the point process are obtained by simulating the process using the algorithm described in Chapter 2. Let $Y_0, Y_1, Y_2, \cdots, Y_N$ denote the data corresponding to a realization of the point process where Y_i, $0 \leq i \leq N$, is the number of packets generated at time epoch i. The criterion for fixing the monitoring level K is based on the result proved in Chapter 1, that the marginal distribution of the number of packets generated at an arbitrary time is Poisson with parameter $\lambda\xi$. The monitoring level K is chosen so that K is the first index for which

$$\sum_{i=0}^{K} e^{-\lambda\xi} \frac{(\lambda\xi)^i}{i!} \geq \gamma, \tag{13}$$

is satisfied. Typical values of γ we have used to fix the threshold K are $\gamma = 0.85$ or $\gamma = 0.90$.

4.2 Some Preliminaries:

After fixing the monitoring level K, we collect various exploratory statistics regarding the excursions above and sojourns below the threshold. From the data $\{Y_n, n \geq 0\}$ summary statistics $\{(I_n, a_n, b_n), n \geq 0\}$ are recorded. I_n is the indicator function which corresponds to the time epochs at which the monitor is initiated and defined as follows: At time epoch n, $n \geq 0$,

$$I_n = \begin{cases} 1 & \text{if } Y_n \geq K \\ 0 & \text{if } Y_n < K. \end{cases}$$

The variables a_n and b_n correspond to the durations of excursions above and sojourns below the threshold respectively and are defined as follows: If $I_n = 1$, then for $j \geq 1$

$$a_n = \begin{cases} j & \text{if } Y_{n-j+1} \geq K, \cdots, Y_n \geq K \\ 0 & \text{otherwise.} \end{cases}$$

If $I_n = 0$, then for $j \geq 1$

$$b_n = \begin{cases} j & \text{if } Y_{n-j+1} < K, \cdots, Y_n < K \\ 0 & \text{otherwise.} \end{cases}$$

Using such summary statistics collected from the data, we can empirically determine the profile curves for the point process. In defining the profile curves, we need to consider the joint conditional frequency of the number of packets generated and of the duration of the excursion, given that the process has exceeded the threshold K. Informally, the profile curves are then the loci of given percentiles of the conditional frequencies of the excursions above the threshold. As a first step, we have to obtain an estimate of the joint conditional probability density from the data, to describe both the amount and the duration of excursions above the level K on the set of lattice points (i, j), $i \geq K$, $j \geq 1$, where i denotes the number of packets generated during an excursion above the level K that has lasted j units of time.

Let $\psi(i, j)$, $i \geq K$, $j \geq 1$, denote the number of times we observe from the data, i packets and for j consecutive units of time the process has been above the level K. $\psi(i, j)$ is given by

$$\psi(i, j) = \sum_{n=0}^{N} \phi_{ij}\,(Y_n I_n, a_n), \tag{14}$$

where ϕ_{ij} is the indicator function defined as

$$\phi_{ij}\,(Y_n I_n, a_n) = \begin{cases} 1 & \text{if } Y_n I_n = i,\ a_n = j \\ 0 & \text{otherwise.} \end{cases}$$

The conditional frequency $f(i, j)$, $i \geq K$, $j \geq 1$, that during an excursion lasting for j units of time above the level K, the counts of packets generated is i, given that the process exceeds the level K at time $j = 1$, can be determined by

$$f(i, j) = \frac{\psi(i, j)}{\sum_{i \geq K} \psi(i, j)}. \tag{15}$$

For each j, $f(i,j)$ is an empirical density function with support on the lattice points $i \geq K$. From the summary statistics collected from the data, we extract additional information in order to determine various frequencies related to monitoring events. Let κ_n, $n \geq 1$ be defined as

$$\kappa_n = \begin{cases} b_{n-1} & \text{if } I_{n-1} = 0, I_n = 1 \\ 0 & \text{otherwise.} \end{cases}$$

Using this statistic, we can obtain the empirical distribution of the sojourn below the threshold K. Let $\beta(j)$, $j \geq 1$ denote the number of times we observe from the data, a sojourn below the threshold K lasting for j time units and $\beta(j)$ can be obtained using

$$\beta(j) = \sum_{n=1}^{N} \phi_j(\kappa_n),$$

where ϕ_j is the indicator function defined as

$$\kappa_n = \begin{cases} b_{n-1} & \text{if } I_{n-1} = 0, I_n = 1 \\ 0 & \text{otherwise.} \end{cases}$$

The frequency $\tau(j)$,$j \geq 1$ that a sojourn below the threshold K lasts for j, time units is given by

$$\tau(j) = \frac{\beta(j)}{\sum_{j \geq 1} \beta(j)}$$

Let $\Phi(j)$ for $j \geq 1$ denote the frequency that an excursion above the threshold lasts for j units of time. It can be verified that the $\Phi(j)$ is given by

$$\Phi(j) = \frac{\sum_{i \geq K} \psi(i, j)}{\sum_{i \geq K} \psi(i, 1) + \sum_{n \geq 1} (\kappa_n - 1)^+}, \tag{16}$$

where $(\kappa_n - 1)^+$ denotes the non-negative part of $\kappa_n - 1$, defined as $\max(0, \kappa_n - 1)$. $\Phi(1)$ is an important performance index for the monitoring level designed and can be interpreted as the frequency with which the monitor is initiated.

4.3 Profile Curves:

With these preliminaries, we can formally define the profile curves for the point process when monitored at the level K. For a given value of α, $0 < \alpha < 1$, the α level profile curve corresponding to the threshold K is the locus of the points (i, j), $j \geq 1$, where i is the smallest integer for which

$$\sum_{r=K}^{i} f(r, j) \geq \alpha.$$

For given values of α and K, the profile curve is a sequence of indices and each index $i^*(j)$ in that sequence is the (approximate) 100α-th percentile of the conditional density of the excess above K, given that for j successive units of time the process has been above the level K. The choice of α is determined by how stringently we want to monitor the point process. The smaller the value of α, more frequently the process will exceed the profile curve. However, if moderate excursions above the level K can be tolerated, values of α such as 0.9 or 0.95 can be chosen. As soon as the empirical density function $f(i,j)$ for each j is determined, each index $i^*(j)$ for $j \geq 1$ is specified by the α quantile of the corresponding empirical distribution and can be obtained by using an efficient binary search.

In practice, it is advisable to specify an upper bound J on the values of j up to which the profile curve has to determined. The reason lies in the fact that monitor should also report on the occurrence of excursions above K which are of unusually long duration, yet do not involve an upward crossing of the profile curve. The value of J can be selected so that the frequency of an excursion lasting longer than J time units is less than ε. In our computer experiments, we have chosen the value of ε to be 0.0001. The profile curve can now be expressed as a finite sequence of indices $\{i^*(j), 1 \leq j \leq J\}$ for given values of α and K. We shall refer to the set of indices (i,j) with

$$1 \leq j \leq J, \qquad K \leq i \leq i^*(j),$$

as the monitoring region where i is the number of packets and j is the number of time units for which the process is in excess of K. As long as the process remains within this region, the process is monitored. An excursion is classified as abnormal or a random fluctuation depending how that excursion exits the monitoring region.

Remark:

In addition to the rule to fix the value of J, it is important to assess if there are sufficient data available to estimate the empirical distributions for all $j \leq J$. Since, our procedure to develop a monitoring scheme is based on the data from a realization of the point process, there may not be enough data on excursions above the level K that last for very long durations of time. It is stressed that, by computing the profile curves beyond a reasonable value of J, the information we obtain is regarding the long range behavior of the profile curves, which is only of marginal interest from our point of view of monitoring the point process. However, an asymptotic result on the conditional distribution of the excess during an excursion that has lasted for j consecutive time units as $j \to \infty$, is proved in Neuts [12] for a class of Markov processes arising in single-server queues. The implication of this result is that the profile curve for that class has a horizontal asymptote. This result would serve as a guiding example in our current investigation for non-Markovian processes.

4.4 Monitoring Procedure:

In this section, we discuss the following specific monitoring procedure for the point process using the the monitoring level K and the profile curves to identify the situations which may require control measures. Let us consider two further indices $J_1 < J_2 < J$, chosen according to a similar rule as J.

If the process crosses over the threshold K, the monitor is initiated. If at any time j with $1 \leq j \leq J_1$, the process stays above K and exceeds $i^*(j)$, the monitor signals a very rapid increase. Such a situation may call for an immediate flow control action. If the rapid increase occurs infrequently, it is probably due to a short-lived trend in the process.

Consider an excursion above K which has not crossed the profile curve up to time J_1, but exceeds $i^*(j)$ for some j with $J_1 < j \leq J_2$. Such an excursion is typical of an overload condition, i.e., during an excursion of relatively short duration, unusually high number of packets are generated which is indicative of the short term, local dependence present in the process.

Once an excursion lasts longer than J_2 time units without having exceeded the profile curve, we may be faced with a process whose parameters are permanently above design criteria. If the process exceeds the profile curve between times J_2 and J, the monitor will declare that a steady, long-lived upward trend of the process is in progress.

Finally, the excursions which stay between the threshold K and the profile curve for J time units at least, may be indicative of a very slow upward drift in the process. The antecedents of such excursions may be attributed to the persistence and the dependent nature of the point process.

4.5 Performance Measures:

In this section, we examine various performance measures in reference to the specific monitoring procedure discussed in the previous section and sketch their derivation. Such a study will be useful in assessing the performance of the profile curves in detecting specific parameter changes in the point process. In order to determine the performance measures to be discussed, we need to collect statistics regarding the excursions that exceed the level K but also remain below $i^*(1)$ at the time of up-crossing. Conceptually, the evaluation of these performance measures involves elementary probabilistic arguments. The procedure we have adopted to collect such statistics is analogous to the scheme described in section 2. If $I_n = 1$, then for $n \geq 0$ and $j \geq 1$

$$c_n = \begin{cases} j & \text{if } K \leq Y_{n-j+1} \leq i^*(1), \cdots, K \leq Y_n \\ 0 & \text{otherwise.} \end{cases}$$

Let $\psi'(i, j)$ denote the number of times we observe from the data, i packets and for $j-1$ successive units of time the process remains between the level K and the profile curve and at the j–th time unit of the excursion the process is above the level K. $\psi'(i, j)$ can be determined by using

$$\psi'(i, j) = \sum_{n=0}^{N} \phi'_{ij}(Y_n I_n, c_n),$$

where ϕ'_{ij} is defined as

$$\phi'_{ij}\,(Y_n I_n, c_n) = \begin{cases} 1 & \text{if } Y_n I_n = i,\ c_n = j \\ 0 & \text{otherwise.} \end{cases}$$

Some performance measures of practical interest are the following:

1. $\xi_1(j)$, $2 \le j \le J$, the frequency that an excursion stays between the level K and the profile curve up to the j–th time unit. It can be verified that $\xi_1(j)$ is given by

$$\xi_1(j) = \frac{\sum_{i=K}^{i^*(1)} \psi'(i, j)}{\sum_{i \ge K} \psi(i, 1) + \sum_{n \ge 1} (\kappa_n - 1)^+} \tag{17}$$

and can be interpreted as the fraction of paths that require a monitoring decision at time epoch $j + 1$. From the point of view of monitoring the process, it is desirable that this quantity decreases as j increases.

2. $\xi_2(j)$, $2 \le j \le J$, is the conditional frequency that an excursion remains between the level K and the profile curve up to the j–th time epoch, given that the process has crossed the level K at time $j = 1$. Since, the frequency that the process exceeds the level K is $\Phi(1)$, we obtain

$$\xi_2(j) = \frac{\xi_1(j)}{\Phi(1)}. \tag{18}$$

The speed at which this quantity drops off as time elapses is indicative of the sensitivity of the monitor to excessive fluctuations.

The following frequencies are conditioned on the event that the process exceeds the monitoring level K at time $j = 1$ and assumes one of the values $K, \cdots, i^*(1)$. We shall refer to this event as A.

3. $\xi_3(j)$, $2 \le j \le J$, is the conditional frequency that an excursion remains in the monitoring region until the j–th time unit, conditional on the event A. $\xi_3(j)$ is given by

$$\xi_3(j) = \frac{\sum_{i=K}^{i^*(j)} \psi(i, j)}{\sum_{i=K}^{i^*(j-1)} \psi(i, j-1)}. \tag{19}$$

4. $\xi_4(j)$, $2 \le j \le J$, is the conditional frequency that an excursion exceeds $i^*(j)$ for the first time at the j–th time unit, conditional on the event A. By using a similar argument, we obtain

$$\xi_4(j) = \frac{\sum_{i > i^*(j)} \psi(i, j)}{\sum_{i=K}^{i^*(j-1)} \psi(i, j-1)} \tag{20}$$

5. $\xi_5(j)$, $2 \le j \le J$, is the conditional frequency that an excursion returns below the level K for the first time at the j–th time unit, conditional on the event A. It is readily seen that for each j

$$\xi_3(j) + \xi_4(j) + \xi_5(j) = 1, \tag{21}$$

which determines the value of $\xi_5(j)$, once we compute $\xi_3(j)$ and $\xi_4(j)$ using the relations (19) and (20). These performance measures are useful in assessing the frequency with which various events related to monitoring will occur purely due to chance, in a process which operates with specific parameter values.

In addition to their use in formalizing a monitoring procedure, the profile curves may also serve in the comparison of different models or of various approximations to the same point process. For the posited model of the point process, we may, for instance, change one of the parameters of the process to obtain a model with different parameters. The changed model is also monitored at the same threshold K and with respect to the profile curves defined for the posited model. A first measure of identifying the change in the parameter is the frequency with which the monitor is initiated. Further, for the changed model, the change in the parameter should be reflected by more frequent systematic excursions above monitoring level. A procedure analogous to the scheme described in this section, can adopted to evaluate the performance measures for the changed model. As a second indicator, we can compare the measures of performance obtained for the posited and the changed model. A comparison of these frequencies yields refined information on the sensitivity of the monitor scheme, to a change in the parameter of the process. Such an investigation is analogous to that of examining the power of statistical tests against various alternate hypotheses.

4.6 Numerical Examples:

In this section, we shall present numerical examples to illustrate the monitoring methodology developed in this chapter and discuss the interpretations of these examples. We have selected four simple examples to demonstrate the feasibility of the procedure developed and to draw qualitative inferences on the performance of the proposed monitoring scheme. The principal conclusion to be drawn from these examples is that the monitoring scheme identifies moderate changes in the parameters of the point process. In particular, the performance measures do reflect on the sensitivity of the monitor to excessive fluctuations when compared to the posited model. While the discussion of the following numerical examples pertain to a specific model, it is important to note that the conclusions we have drawn are based on extensive experimentation with various parameters for the process and replication of the experiments under several random number seeds.

The posited model of the point process has two parameters λ, the arrival rate of messages and the holding time of each message which is of phase type with representation (α, T). Besides these two parameters, the indices of the states within the matrix T that generate packets have to specified. These parameters of the posited model are listed in Table 2. The data from the evolution of the point process with those parameters are obtained, by simulating the process for 500,000 time units. The process is monitored at the level $K = 27$ chosen according to the 75–th percentile of the Poisson distribution with mean $\lambda\xi = 24.0$. ξ can be computed from the parameters of the point process and its value is 0.80.

Table 3 lists the following quantities for the posited model: the frequencies relating to excursions above the threshold, the 80 and 95 percentile profile curves, performance measures for the monitor, the empirical distributions of the duration of an excursion above the level and a sojourn below the level. A scatter plot of a data set containing 10,000 observations regarding the successive durations of excursions above threshold and sojourns below the threshold for the posited model is displayed in Figure 1. The frequency with which the monitor is initiated is 0.0909, and this indicates that the monitor is activated infrequently which is a desirable property of the monitor. It is important to note that the frequency that an excursion above the threshold lasting beyond the 25–th time unit is smaller than 0.0043. From the relatively narrow spacing between 80–th and 95–th percentile profile curves, we can infer that the process does not exhibit large fluctuations in the path functions that have exceeded the threshold. This is further confirmed by the performance measures. The conditional frequency $\xi_4(j)$ of exit from the monitoring region by crossing the profile curve is small compared to the other frequencies. Besides the monitoring aspect, this study reveals some of the behavioral characteristics of the point process. For instance, we can infer from the performance measures and the scatter plot that most of the path functions that exceed the level either return below the level quickly or else tend to exhibit long-lived and slowly varying fluctuations within the monitoring region.

We have considered three changes in the parameters of the posited model yielding three different changed models. As the second example, the mean duration of each message is increased while the other parameter is kept unchanged. Similarly, only the arrival rate of messages is increased in the third example. Within the duration of each message, only the distribution governing packet generation is changed in the fourth example. The parameters of the changed models are displyed in Table 4. The processes with the changed parameters are first simulated for 500,000 time units and then monitored with respect to the same threshold $K = 27$ and the 80–th percentile profile curve defined for the posited model. Table 5, Table 6 and Table 7 list the performance measures, empirical distributions for durations above and sojourns below the threshold for the second and third examples respectively. Substantially, similar conclusions can be drawn from Table 5, Table 6 and

Table 7. Firstly, $\xi_2(j)$, the conditional fraction of paths having exceeded the level K and awaiting a monitoring decision, drops off at a faster rate compared to the posited model. This implies that the monitor is sensitive to the change in the parameter. Secondly, the conditional frequency $\xi_4(j)$, is significantly higher for the changed model. This conveys that in the changed model, the path functions that remain in the monitoring region, will exceed the profile curve more frequently. Overall, the performance measures indicate that the excursions in the changed model either cross the profile curve soon or return back within the level K after remaining for a moderate duration of time in the monitoring region. A visual comparison of the scatter plots in Figure 1 and Figure 2 also demonstrates the effect of a change in the parameter. More frequent excursions that last for longer durations of time above the threshold in the changed model are shown in the scatter plot.

4.7 Conclusions:

In this approach to monitoring, once the process exceeds the threshold, that path function is tracked until it returns back below the level. The profile curves are defined with respect to the conditional probability induced on the space of all path functions of the process that have exceeded the threshold. Hence a crucial step in our present study is, determining an estimate of the joint conditional density of the amount and duration of the excursion above the threshold. To that end, we have defined an empirical conditional distribution for each time unit of the excursion.

This would tantamount to assuming implicitly that the successive excursions above the threshold are independent. Under such a condition, the collection of the sample paths of the process that have exceeded the threshold would be composed of independent replicates of excursions, which would in turn justify the procedure we have adopted here to determine the empirical conditional density. In contrast, we observe from the scatter plot shown in Figure 1, that long sojourns below the threshold often trigger short excursions above the threshold and vice versa. However, it is our conjecture that mixing properties hold for the point process which would suggest asymptotic independence. Besides this result, we draw attention to the fact that for the classical queueing models, the profile curves can be determined in an algorithmic approach. The monitoring approach developed for well-understood models such as single-server queues provides the general guidelines into an exciting, but unwieldy area of monitoring non-Markovian processes which calls for a variety of mathematical methods of analysis.

In this chapter, we have developed a monitoring scheme for the Neuts-Pearce point process based on data obtained from the simulation of the process. Our approach is general and can be applied to any discrete, non-negative integer-valued, stationary process, given a long record of observations from the evolution of that process. The distributional property regarding the marginals is convenient, but not a requirement. Since, an empirical estimate of the marginal density can be obtained from the data in the absence of a specific marginal distribution. Finally, it is important to stress that our approach to monitoring the process falls within a general framework of empirical, model free methods for the statistical analysis of a class of stochastic processes.

CHAPTER 5

WALSH SPECTRAL METHODS

5.1 Introduction:

One of the most frequently used descriptive techniques to summarize the information embedded in data recorded from a time series is classical Fourier spectral analysis. The essence of Fourier spectral analysis lies in the representation of a set of data in terms of sinusoidal functions. For a clear exposition regarding the Fourier spectral analysis of data, we refer to Bloomfield [3]. The Fourier spectrum thus obtained encodes global information about the fluctuations and reveals the hidden periodicities present in the time series. The simple physical interpretation of the frequency notion and the computational simplifications in view of the FFT (fast Fourier transform) algorithm are some of the appealing features of Fourier analysis. As an alternative to the classical Fourier spectral methods, in this chapter, we examine Walsh spectral techniques to analyze data from a time series. The principal motivation for such an investigation is derived from the advantages gained in the computational simplicity of the Walsh functions, and the novel interpretation of the Walsh spectrum.

Walsh functions play a role analogous to the trigonometric functions used in the Fourier analysis. A synthesis of a complex waveform can be performed using Walsh functions which is equivalent to a decomposition of data into components that are expressed in terms of Walsh functions. However, it should be noted that in some cases Walsh functions are more suitable than their Fourier counterparts for approximating non-smooth waveforms. The Walsh functions form a complete, orthogonal sequence of functions. They take values from the set $\{+1, -1\}$, a property ideally suited for numerical computations. Analogous to the FFT algorithm, there is a fast Walsh transform. Due to the binary-valued nature of the Walsh functions, the fast Walsh transform can be evaluated on most modern digital computers more quickly than the FFT. For a detailed discussion of basic mathematical properties, applications and computational features of Walsh functions, we refer to Beauchamp [1] and the references therein. Our interest in studying Walsh spectral techniques is inspired by their applications in the design of factorial experiments and in the output analysis of simulation experiments; refer to Sanchez and Sanchez [21].

The Walsh function is defined over a fixed interval $[0, T]$. It is written as $WAL(n, t)$, $0 \le t \le T$. The ordering index n, is referred as the sequency of the Walsh function. The Sequency is defined to be proportional to the average number of zero crossings of the function within the interval $[0, t]$ and is a generalization of the frequency notion. If time is discretized by dividing the interval $[0, T]$ into N

equal subintervals, the discrete Walsh function is written as $WAL(n, i)$, $0 \le i \le N-1$. In the i–th subinterval $WAL(n, i)$ has same the value as $WAL(n, t)$, and i is the smallest integer greater than Nt/T. The following indentities regarding orthogonality, symmetry and multiplication of discrete Walsh functions hold:

$$\sum_{i=0}^{N-1} WAL(m, i)\, WAL(n, i) = \begin{cases} N & \text{if } n = m \\ 0 & \text{if } n \neq m. \end{cases}$$

$$WAL(n, i) = WAL(i, n),$$

$$WAL(m, i)\, WAL(n, i) = WAL(m \oplus n, i),$$

where $\oplus$ denotes bitwise modulo-2 addition for the binary representations of m and n. Let $\{x_i, 0 \le i \le N-1\}$ denote observations recorded from a time series. Without loss of generality, N can be selected so that $N = 2^p$ for some integer p, $p \ge 1$. The discrete Walsh transform of a sequence can be written for $k = 0, 1, \cdots, N-1$ as

$$X_k = \frac{1}{N}\sum_{i=0}^{N-1} x_i WAL(k, i).$$

The original time series x_i for $i = 0, 1, \cdots, N-1$ can be recovered as

$$x_i = \sum_{n=0}^{N-1} X_n WAL(n, i).$$

Our goal in chapter, is to determine the Walsh spectrum for the data obtained from the simulation of the point process studied as a model for the packet stream in Chapter 1. In conjunction with the general objectives of this thesis, Walsh spectral analysis provides a novel tool for exploring the statistical information present in the data recorded from the point process. As with Fourier spectral analysis, there are several equivalent approaches to determine the Walsh spectrum. In the first approach, the Walsh spectrum is expressed in terms of the dyadic autocovariance function. We assume that the data are generated by a covariance stationary process with mean 0. The dyadic autocovariance can be obtained from the data using the relation

$$\gamma_\tau = \frac{1}{N}\sum_{i=0}^{N-1} x_i \, x_{i \oplus \tau},$$

where τ denotes the time lag. The Walsh spectral density is then explicitly given by the Walsh transform of the dyadic autocovariance, a result referred to in the literature as the logical version of the Wiener-Khinchin theorem. The second approach, is to determine the Walsh periodogram which is an estimate of the Walsh spectrum. In our computer experiments, we have used the periodogram approach to determine the Walsh spectrum. Let $\{y_i, 0 \le i \le N-1\}$ denote the data recorded from the simulation of the point process studied as a model for the packet stream, where N is the duration for which the process in simulated. The Walsh transform of the data is given by

$$Y_k = \frac{1}{N} \sum_{i=0}^{N-1} y_i WAL(k, i),$$

for $k = 0, 1, \cdots, N-1$, and can be computed efficiently using the fast Walsh transform algorithm. The Walsh periodogram is supported on $(N/2)+1$ spectral points and can be computed from the Walsh transform coefficients as

$$P(0) = Y_0^2,$$

$$P(N/2) = Y_{N-1}^2,$$

and for $1 \leq k \leq (N/2) - 1$

$$P(k) = Y_{(2*(k-1))+1}^2 + Y_{(2*(k-1))+2}^2.$$

5.2 Numerical Results:

In our computer experiments, the point process is simulated for 262,144 time units using the algorithm described in Chapter 2. It is necessary to have a large data set, since our goal is to assess the fluctuations in the point process using the Walsh spectrum. The data recorded from the simulation of the process are centered around the sample mean of the data. Using the fast Walsh transform algorithm, we compute the Walsh transform for this data set. The periodogram can be computed from the resulting Walsh transform coefficients. We replicate this experiment by simulating the process for several random number seeds and determine the periodogram for each replication. The sample average of the periodograms obtained from all the replications is computed. Such an average periodogram is our estimate of the Walsh spectrum. A different procedure is to determine the average of the periodograms obtained from splitting the entire data set into smaller batches. In our present study, in view of the highly dependent nature of the point process, we have adopted to estimate the periodogram by replications.

As a sample of our experiments, a plot of the Walsh spectrum is displayed in Figure 3. An inspection of figure 3 suggests that the Walsh spectrum has dominant low sequencies and a few significant high sequencies. The notion of sequency is related to the average number of zero crossings of the process. In this study, a zero crossing refers to an up or down crossing of the sample mean of the process. The low sequencies indicate that whenever the process stays above the sample mean, then over long periods of time it is likely that the process would remain above the sample mean. This suggests that there is high persistence in the process. Such a behavior may be attributed to the correlation structure of the process, in view of the relation between spectrum and the correlation. The serial correlation coefficient between N_0 and N_n, the number of packets generated at an arbitrary time and n time units later is derived in Chapter 1 and is given by

$$\rho(n) = \xi^{-1}\pi\Lambda T^n \Lambda e.$$

It can be verified that $\rho(n)$ for all n is non-negative. In contrast, the high sequencies reflect on the short range dependence in the process.

It is important to note that a flat Walsh spectrum would indicate that the data are generated by independent random variables or dependent random variables with zero dyadic correlation. In particular, for the application on hand, the non-flat Walsh spectrum as shown in Figure 3 serves as an evidence to the dependence such as the dyadic correlation in the data from the point process. Further, the sequencies obtained from the Walsh spectrum convey the information regarding the fluctuations around the sample mean of the data. In conclusion, we observe that analogous to the Fourier spectrum, Walsh spectrum provides a tool for summarizing information in data.

CHAPTER 6

NUMERICAL RESULTS

Table 1

Numerical Results from the Study of Determining the Order of a Markov Chain

An empirical study of fitting an r-th order MC for the data obtained from the simulation of Neuts-Pearce process for quantifying local dependence in the process

We choose of a word of length 9 and estimate the frequency of its occurrence first and then estimate the frequency of that same word if it were generated by an rth order MC and then compare both the frequencies using a test analogous to the chisquare test

words	frequency	order 0	order 1	order 2	order 3	order 4
111111111	0.17672	0.00583	0.14784	0.17373	0.18198	0.17510
000000000	0.10390	0.00001	0.04909	0.07384	0.08463	0.09699
222222222	0.02898	0.00000	0.02408	0.02732	0.02946	0.03477
111111110	0.01564	0.00287	0.01674	0.01616	0.01512	0.01599
011111111	0.01561	0.00287	0.01674	0.01616	0.01511	0.01598
211111111	0.01143	0.00163	0.01022	0.01084	0.01155	0.01122
111111100	0.01141	0.00141	0.01594	0.01130	0.00965	0.01058
111111112	0.01139	0.00163	0.01022	0.01084	0.01155	0.01122
221111111	0.01137	0.00045	0.00955	0.00922	0.00954	0.00970
222111111	0.01135	0.00013	0.00893	0.00858	0.00835	0.00945
111111000	0.00997	0.00069	0.01518	0.01115	0.00818	0.00815
111111122	0.00993	0.00045	0.00955	0.00922	0.00880	0.00809
100000000	0.00859	0.00002	0.01187	0.01266	0.01229	0.01109
000000001	0.00858	0.00002	0.01187	0.01266	0.01228	0.01109
001111111	0.00858	0.00141	0.01594	0.01130	0.01027	0.01057
222221111	0.00854	0.00001	0.00780	0.00742	0.00729	0.00738

101111111	0.00850	0.00287	0.00385	0.00737	0.00706	0.00790
222222221	0.00715	0.00000	0.00637	0.00662	0.00668	0.00649
122222222	0.00715	0.00000	0.00638	0.00662	0.00668	0.00649
222211111	0.00714	0.00004	0.00835	0.00798	0.00780	0.00758
222222111	0.00713	0.00000	0.00729	0.00690	0.00682	0.00711
222222211	0.00713	0.00000	0.00682	0.00642	0.00626	0.00651
111111011	0.00711	0.00287	0.00385	0.00605	0.00646	0.00609
110111111	0.00709	0.00287	0.00385	0.00605	0.00646	0.00663
111111222	0.00709	0.00013	0.00893	0.00858	0.00770	0.00623
110000000	0.00575	0.00004	0.01247	0.01055	0.00920	0.00730
000000011	0.00574	0.00004	0.01247	0.01055	0.00979	0.00873
111110000	0.00573	0.00034	0.01445	0.01100	0.00819	0.00666
000111111	0.00573	0.00069	0.01518	0.01115	0.00870	0.00915
111111101	0.00570	0.00287	0.00385	0.00737	0.00768	0.00789
111101111	0.00569	0.00287	0.00385	0.00605	0.00577	0.00632
111110111	0.00569	0.00287	0.00385	0.00605	0.00577	0.00575
111111211	0.00569	0.00163	0.00253	0.00298	0.00421	0.00467
212222222	0.00567	0.00000	0.00047	0.00180	0.00241	0.00295
000010000	0.00566	0.00002	0.00141	0.00304	0.00300	0.00309
000100000	0.00565	0.00002	0.00141	0.00304	0.00300	0.00393
110001111	0.00564	0.00069	0.00349	0.00155	0.00095	0.00211

Order of a Markov Chain	χ^2 statistic
0	5.63120
1	0.54395
2	0.30262
3	0.19470
4	0.11801

Table 2

Parameters of the Posited Model in Example 1

arrival rate of messages λ 30.00

parameters of the holding time of each of message

order of the PH-distribution 4

initial probability vector α

0.2500 0.2500 0.2500 0.2500

T-matrix of the PH-distribution

0.5000	0.0000	0.0000	0.0000
0.0000	0.8750	0.0000	0.0000
0.0000	0.0000	0.8750	0.0000
0.0000	0.0000	0.0000	0.9688

vector T^0 of the PH-distribution

0.5000 0.1250 0.1250 0.0312

mean of the PH-distribution 12.50

steady state vector of the matrix $T + T^0\alpha$

0.0400 0.1600 0.1600 0.6400

indices of states that generate packets 3, 4

fraction of time generating packets ξ 0.80

Table 3

Profile Curves for the Posited Model in Example 1

The monitor is initiated when the number of packets generated exceeds the level 27 and monitoring is continued till the process returns back to the level 27

total time the process is simulated 500000

time units	frequency that an excursion lasts for j units of time	80 level profile curve	95 level profile curve
1	0.09094	28	29
2	0.06251	30	31
3	0.05257	30	33
4	0.04262	31	33
5	0.03693	31	36
6	0.03125	32	35
7	0.02983	33	35
8	0.02983	32	35
9	0.02700	32	36
10	0.02557	35	36
11	0.02274	33	36
12	0.02132	32	35
13	0.01848	33	35
14	0.01706	32	34
15	0.01423	31	33
16	0.01281	32	33
17	0.01280	31	31
18	0.00854	31	32
19	0.00854	31	32
20	0.00854	32	33
21	0.00711	31	31
22	0.00711	31	32
23	0.00711	30	32
24	0.00711	32	33
25	0.00427	34	34

performance measures for the posited model

j	$\xi_1(j)$	$\xi_2(j)$	$\xi_3(j)$	$\xi_4(j)$	$\xi_5(j)$
2	0.06026	0.64093	0.43909	0.09738	0.46353
3	0.02793	0.29709	0.15808	0.15784	0.68408
4	0.01911	0.20323	0.07709	0.15364	0.76927
5	0.01470	0.15634	0.30016	0.09975	0.60009
6	0.00882	0.09382	0.16660	0.16622	0.66718
7	0.00588	0.06259	0.00058	0.24971	0.74971
8	0.00441	0.04693	0.00000	0.33308	0.66692
9	0.00294	0.03130	0.49827	0.00116	0.50058
10	0.00147	0.01567	0.00000	0.00000	1.00000

selected percentiles of the empirical distribution of the duration of an excursion above the threshold

0.10000	2
0.25000	2
0.50000	3
0.75000	11
0.90000	17
0.95000	24
0.99000	37

selected percentiles of the empirical distribution of the duration of a sojourn below the threshold

0.10000	2
0.25000	2
0.50000	3
0.75000	13
0.90000	37
0.95000	58
0.99000	67

Table 4

Parameters of Changed Models

Parameters of the Changed Model in Example 2

arrival rate of messages λ 30.0

parameters of the holding time of each message

order of the PH-distribution 4

initial probability vector α

0.2500 0.2500 0.2500 0.2500

T-matrix of the PH-distribution

0.5000	0.0000	0.0000	0.0000
0.0000	0.7500	0.0000	0.0000
0.0000	0.0000	0.9000	0.0000
0.0000	0.0000	0.0000	0.9750

vector T^0 of the PH-distribution

0.5000 0.2500 0.1000 0.0250

mean of the PH-distribution 14.00

steady state vector of the matrix $T + T^0\alpha$

0.03571 0.07143 0.17857 0.71429

indices of states that generate packets 3, 4

Parameters of the Changed Model in Example 3

arrival rate of messages λ 35.0

parameters of the holding time of each of message

order of the PH-distribution 4

initial probability vector α

0.2500 0.2500 0.2500 0.2500

T-matrix of the PH-distribution

0.5000	0.0000	0.0000	0.0000
0.0000	0.8750	0.0000	0.0000
0.0000	0.0000	0.8750	0.0000
0.0000	0.0000	0.0000	0.9688

vector T^0 of the PH-distribution

0.5000 0.1250 0.1250 0.0312

mean of the PH-distribution 12.50

steady state vector of the matrix $T + T^0\alpha$

0.0400 0.1600 0.1600 0.6400

indices of states that generate packets 3, 4

Parameters of the Changed Model in Example 4

order of the PH distribution 5

initial probability vector α

0.2000	0.2000	0.2000	0.2000	0.2000

T-matrix of the PH-distribution

0.5000	0.0000	0.0000	0.0000	0.0000
0.0000	0.8750	0.0000	0.0000	0.0000
0.0000	0.0000	0.8000	0.0750	0.0000
0.0000	0.0000	0.0188	0.9000	0.0500
0.0000	0.0000	0.0250	0.0250	0.9000

vector T^0 of the PH-distribution

0.5000	0.1250	0.1250	0.0312	0.0500

mean of the PH-distribution= 12.76288

steady state vector of the matrix $Q = T + T^0\alpha$

0.03134	0.12536	0.15347	0.35541	0.33441

indices of states that generate packets 3, 4, 5

Table 5

Performance Measures for the Changed Model in Example 2

j	$\xi_1(j)$	$\xi_2(j)$	$\xi_3(j)$	$\xi_4(j)$	$\xi_5(j)$
2	0.07045	0.66028	0.31447	0.02858	0.65694
3	0.04629	0.43377	0.21771	0.06526	0.71703
4	0.03319	0.31102	0.15170	0.03045	0.81784
5	0.02714	0.25437	0.14812	0.03696	0.81492
6	0.02212	0.20729	0.18193	0.09105	0.72703
7	0.01608	0.15071	0.18738	0.00000	0.81262
8	0.01307	0.12247	0.15382	0.15411	0.69207
9	0.00904	0.08475	0.11134	0.11134	0.77732
10	0.00703	0.06588	0.71353	0.00000	0.28647

selected percentiles of the empirical distribution of the durations of excursion above the threshold

0.10000	2
0.50000	4
0.75000	9
0.90000	22
0.95000	25
0.99000	48

selected percentiles of the empirical distribution of the durations of sojourns below the threshold

0.10000	2
0.25000	2
0.50000	4
0.75000	8
0.90000	28
0.95000	38
0.99000	72

Table 6

Performance Measures for the Changed Model in Example 3

j	$\xi_1(j)$	$\xi_2(j)$	$\xi_3(j)$	$\xi_4(j)$	$\xi_5(j)$
2	0.09659	0.61228	0.33364	0.00053	0.66583
3	0.06431	0.40767	0.25005	0.29909	0.45086
4	0.02899	0.18380	0.33179	0.00066	0.66755
5	0.01936	0.12270	0.00099	0.16722	0.83179
6	0.01610	0.10206	0.39849	0.00000	0.60151
7	0.00968	0.06139	0.00132	0.00000	0.99868
8	0.00967	0.06131	0.00132	0.00132	0.99736
9	0.00965	0.06115	0.00132	0.33289	0.66578
10	0.00642	0.04071	0.00000	0.00000	1.00000

selected percentiles of the empirical distribution of the durations of excursion above the threshold

0.10000	2
0.25000	2
0.50000	5
0.75000	15
0.90000	26
0.95000	67
0.99000	126

selected percentiles of the empirical distribution of the durations of sojourns below the threshold

0.10000	2
0.25000	2
0.50000	2
0.75000	9
0.90000	20
0.95000	25
0.99000	29

Table 7

Performance Measures for the Changed Model in Example 4

j	$\xi_1(j)$	$\xi_2(j)$	$\xi_3(j)$	$\xi_4(j)$	$\xi_5(j)$
2	0.03406	0.44103	0.27312	0.18104	0.54583
3	0.01859	0.24073	0.49905	0.00152	0.49943
4	0.00928	0.12023	0.33321	0.00038	0.66641
5	0.00619	0.08012	0.49829	0.49771	0.00400
6	0.00002	0.00032	0.14286	0.00000	0.85714
7	0.00002	0.00027	0.33333	0.00000	0.66667
8	0.00001	0.00018	0.00000	0.00000	1.00000
9	0.00001	0.00018	0.25000	0.00000	0.75000
10	0.00001	0.00014	0.33333	0.00000	0.66667

selected percentiles of the empirical distribution of the durations of excursion above the threshold

0.10000	2
0.25000	2
0.50000	3
0.75000	9
0.90000	32
0.95000	33
0.99000	59

selected percentiles of the empirical distribution of the durations of sojourns below the threshold

0.10000	2
0.25000	3
0.50000	6
0.75000	13
0.90000	33
0.95000	60
0.99000	74

Figure 1: A Scatter Plot for the Posited Model

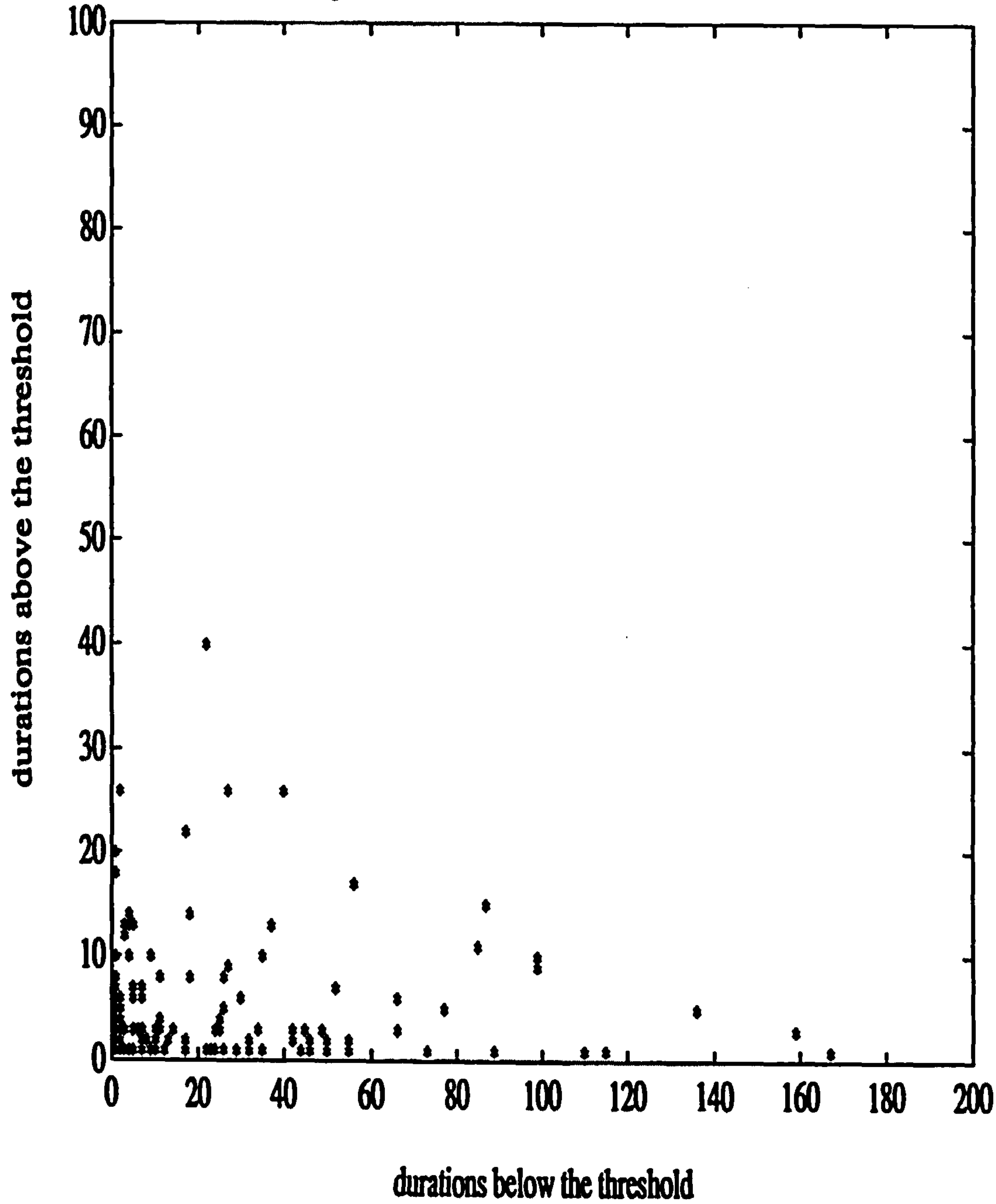


Figure 2: A Scatter Plot for the Changed Model

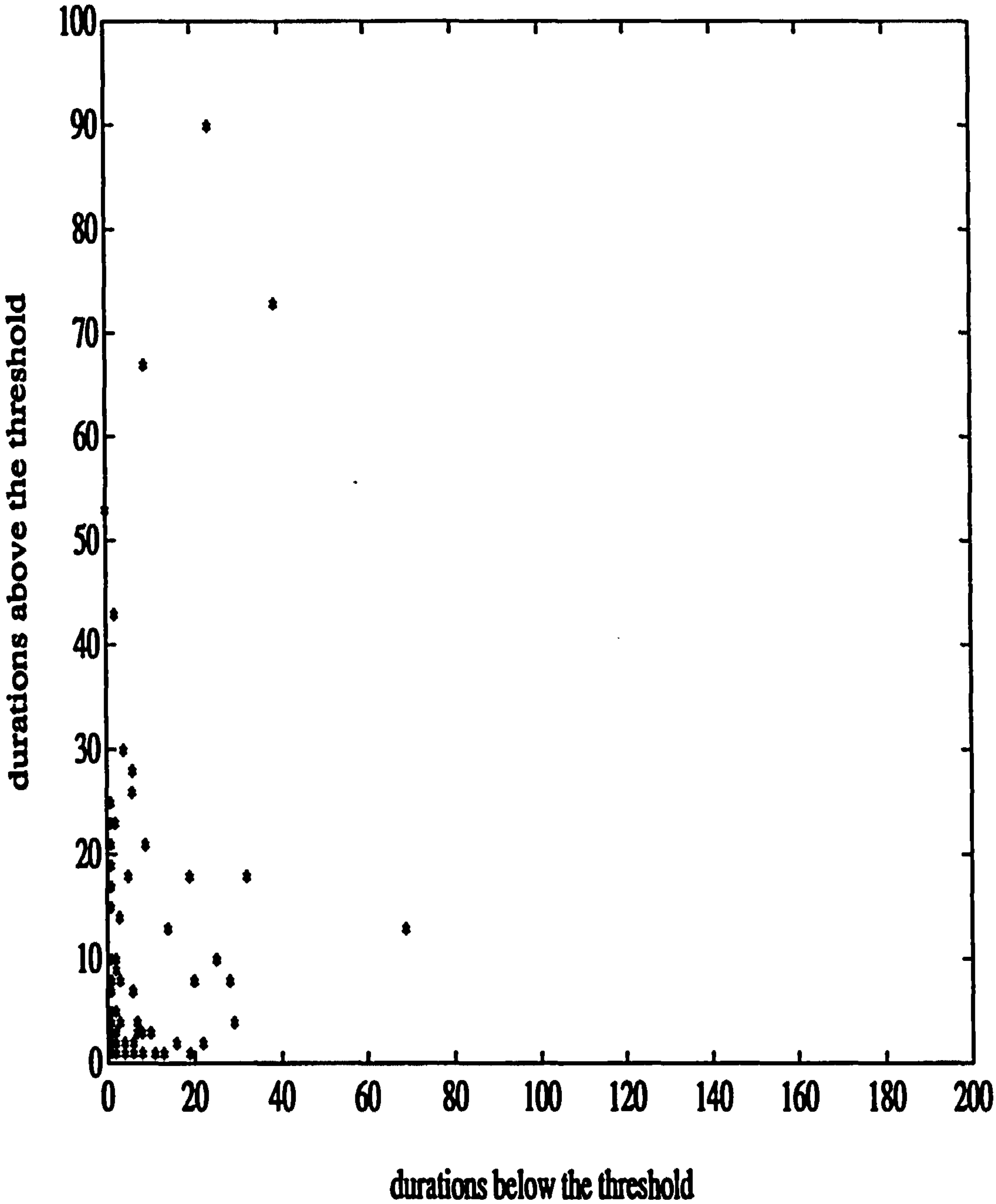
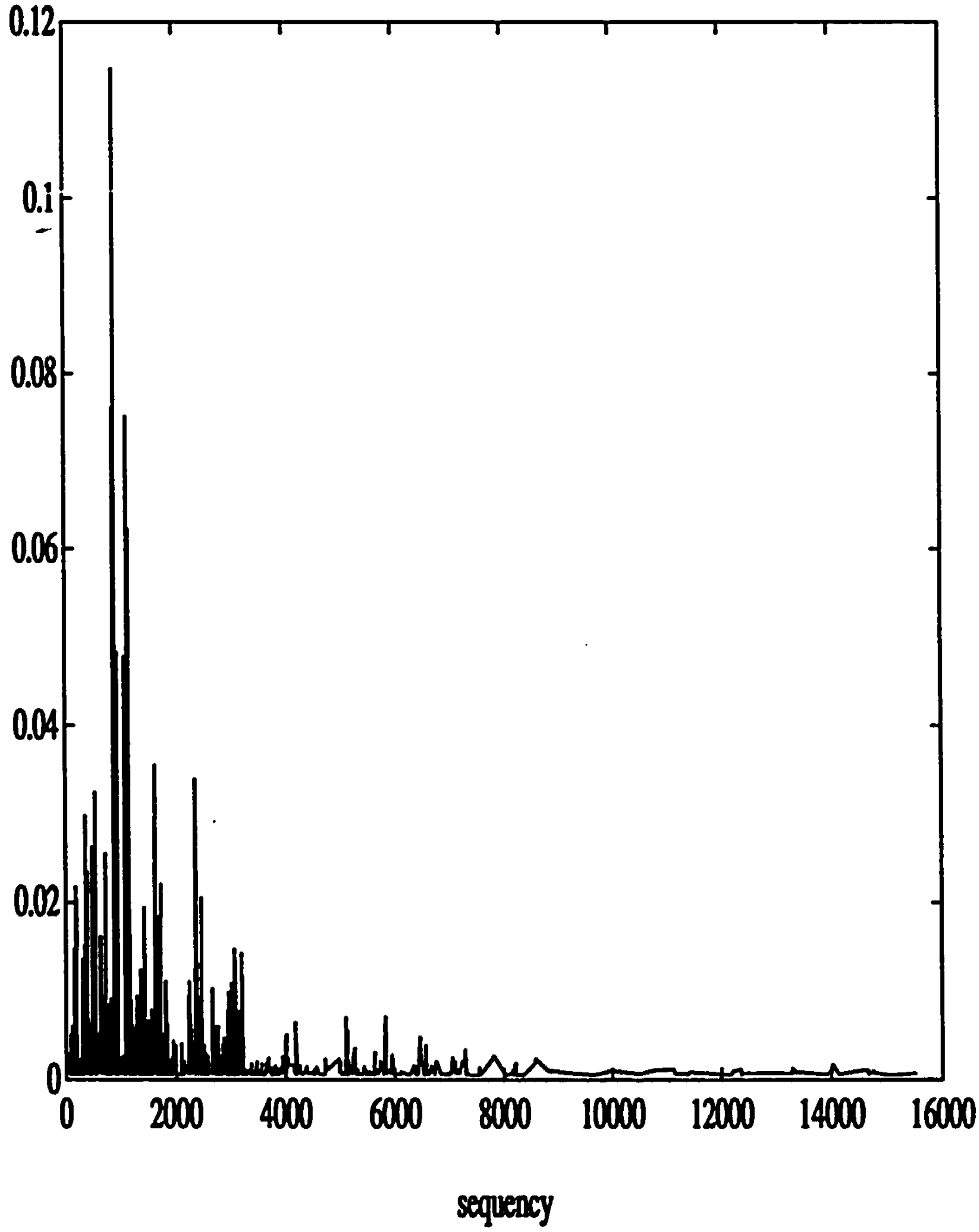


Figure 3: Walsh Spectrum



REFERENCES

[1] Beauchamp, K. G. (1984) *Applications of Walsh and Related Functions With an Introduction to Sequency Theory.* New York: Academic Press.

[2] Billingsley, P. (1965) *Statistical Inference for Markov Processes.* Chicago: University of Chicago Press.

[3] Bloomfield, P. (1976) *Fourier Analysis of Time Series: An Introduction.* New York: John Wiley.

[4] Cox, D. R. and Lewis, P. A. W. (1966) *The Statistical Analysis of Series of Events.* London: Methuen.

[5] Destexhe, A. (1990) Markov Chains from Time Series. *Physics Letters A,* forthcoming.

[6] Heffes, H. and Lucantoni, D. M. (1986) A Markov Modulated Characterization of Packetized Voice and Data Traffic and related Statistical Multiplexer Performance. *IEEE J. on Selected Areas in Communication, Special Issue on Network Performance Evaluation,* SAC-4, 6, 856-868.

[7] Lewis, P. A. W. and Orav, E. J. (1989) *Simulation Methodology for Statisticians, Operations Analysts and Engineers, Vol 1.* Pacific grove: Wadsworth & Brooks/Cole.

[8] Kaczura, A. (1973) An Interrupted Poisson Process as an Overflow Process. *Bell Syst. Tech. J.,* vol. 52, no. 6, 437-438.

[9] Karlin, S. and Taylor, H. M. (1975) *A First Course in Stochastic Processes.* New York: Academic Press.

[10] Kronmal, R. A. and Peterson, Jr., A. V., (1979) On the Alias Method for Generating Random Variables from a Discrete Distribution. *American Statistician* , Vol. 33, no. 4, 214-218.

[11] Neuts, M. F. (1981) *Matrix-Geometric Solutions in Stochastic Models: An Algorithmic Approach.* Baltimore: The Johns Hopkins University Press.

[12] Neuts, M. F. (1988) Profile Curves for the $M/G/1$ Queue with Group Arrivals. *Stochastic Models,* Vol. 4, no. 2, 277-298.

[13] Neuts, M. F. (1988) Monitoring Queueing Systems. *Computer Performance and Reliability*, G. Iazeolla, P. J. Courtois and O. J. Boxma (Eds.), Amsterdam: North-Holland, 59-69.

[14] Neuts, M. F. (1988) Computer Experiments in Applied Probability. *J. Appl. Prob.*, Vol 25A, 31-43.

[15] Neuts, M. F. (1989) *Structured Stochastic Matrices of M/G/1 Type and Their Applications*. New York: Marcel Dekker.

[16] Neuts, M. F. (1990) On the Packet Stream Generated by a Random Flow of Messages of Random Durations. *Stochastic Models*, Vol. 6, no. 3, 445-470.

[17] Neuts, M. F. and Pagano, M. E. (1981) Generating Random variates from a Distribution of Phase Type. *Proceedings of 1981 Winter Simulation Conference*, T. I. Oren, C. M. Delfosse and C. M. Shub (Eds.), 381-387.

[18] Neuts, M. F. and Pearce C. E. M., (1990) The Superposition of Independent Discrete Markovian packet Streams. *Adv. Appl. Prob.* forthcoming.

[19] Ramaswami, V. (1988) Traffic performance modeling for packet communication, whence where and whither ? *Keynote Address*, Third Australian Teletraffic Research Seminar, Melbourne, Australia.

[20] Ramaswami, V. and Latouche, G. (1988) Modeling Packet Arrivals from Asynchronous Input Lines. *Proceedings of 12-th International Teletraffic Conference, Torino, Italy*, 721-727.

[21] Sanchez, P. and Sanchez, S. (1988) The Design and Analysis of 2^k Factorial Experiments Using Walsh Spectral Methods. *Working paper*, 88-017, Department of Systems and Industrial Engineering, University of Arizona, Tucson, AZ.

[22] Schrage, L. (1979) A More Portable Fortran Random Number Generator. *ACM Transactions on Mathematical Software*, 5, 132-138.

[23] Sriram, K. and Whitt, W. (1986) Characterizing Superposition Arrival Processes in Packet Multiplexers for Voice and Data. *IEEE J. on Selected Areas in Communication, Special Issue on Network Performance Evaluation*, SAC-4, 6, 833-846.

[24] Walker, A. M. (1977) An Efficient Method for Generating Discrete Random Variables With General Distributions. *ACM Transactions on Mathematical Software*, 3, 253-256.