J. Appl. Prob. **25**, 184–193 (1988)
Printed in Israel

MODELLING THE FLOW OF COALESCING DATA STREAMS THROUGH A PROCESSOR

V. ANANTHARAM,* *University of California, Berkeley*

Abstract

In a data processing network, two data streams A and B arrive at a node independently at the same Poisson rate λ. Service at exponential rate μ can take place iff there is at least one of each of A and B present. The output is the combined processed data AB. We consider models of this situation with finite buffers, with infinite buffers and with finite buffers for the excess of each input type over the other. We apply the filtering theory for point process functionals of a Markov chain to study whether the output flow is Poisson in equilibrium. The motivation is to examine, if the output is to subsequently be processed by a queueing system, whether it can be treated as an independent Poisson input to that system.

A result of independent interest is that a subset of transitions of a countable-state Markov process does not yield a Poisson process when counted, if the rate matrix of counted transitions is nilpotent, and we prove a generalization of Pakes' lemma for countable-state Markov chains.

FILTERING; MANUFACTURING; OUTPUT PROCESSES; QUASIREVERSIBILITY; STATIONARY DISTRIBUTIONS

1. Introduction

Data arriving into a network of processors are commonly modelled as a Poisson process. The processors in the network are themselves modelled variously according to their service discipline. One of the remarkable discoveries in queueing theory, originating with work of Burke [2], is that for a large class of models of processors, called quasi-reversible, the output flows of the network are also Poisson in equilibrium [4], [5], [6], [9], [10]. This supports the popularity of quasi-reversible nodes as models for processors and facilitates analysis of networks of quasi-reversible nodes.

In a parallel processing system, where portions of a job are worked on by different processors, a given processor may require inputs from more than one data stream before beginning processing. Alternately, to manufacture a commodity it may be necessary to have available different types of raw materials. We call a node modelling such a situation a *manufacturing* node. In this paper we investigate the possibility of coming up with a reasonable model of a manufacturing node with an equilibrium situation where Poisson

Received 9 May 1986; revision received 8 July 1986

* Present address: School of Electrical Engineering, Phillips Hall, Cornell University, Ithaca, NY 14853, USA.

inputs yield a Poisson output. The availability of such a model would greatly extend the range of data processing situations which can be modelled by quasi-reversible nodes.

The $M/M/1/S$ queue is one of the simplest models for a service facility with finite waiting room. Customers arrive in a Poisson stream of rate λ and require independent service times exponentially distributed with rate μ. Any customers arriving when the queue size is S are rejected. It is known that the output of the $M/M/1/S$ queue is not Poisson in equilibrium. In Section 3 we consider the following model for a manufacturing node with two inputs and finite waiting room: customers of types A and B arrive according to independent Poisson streams of rate λ. Service at exponential rate μ can take place only if there is at least one of both A and B present. The output is the combined commodity AB. If an arrival of type A finds N_A customers of type A in queue on arrival, it is rejected. N_B is defined similarly. We call this the model with finite buffers. We prove that the output in equilibrium of this manufacturing node is not Poisson.

If the buffer size S is very large, the $M/M/1/S$ model may be simplified by allowing for arbitrary queue size, giving the familiar $M/M/1$ model. This is a quasi-reversible node, and its output is therefore Poisson in equilibrium. By analogy, if N_A and N_B are very large, one might try to simplify the model with finite buffers by allowing for infinite buffers. It is natural to expect the output of this node to be Poisson in equilibrium. Somewhat unexpectedly, it turns out that the question is meaningless, because, even if only one of the buffers is infinite, there is no equilibrium situation. This is proved in Section 4. The point is that any excess of customers of one type over another tends to remain unchanged, whatever the rate of service. The behaviour of the model is similar to that of a null-recurrent Markov chain. We also prove that *any* countable-state Markovian model of a manufacturing node without blocking of arrivals cannot admit an invariant distribution.

As a final attempt to get a simple model with the 'infinite' features which make a Poisson output plausible, we consider the following. Arrivals and service are as before. An arrival of type A is rejected on arrival if the *excess* of customers of type A over those of type B he sees on arrival is N_A. N_B is defined similarly. We call this the model with finite buffers for the excess. We show that even for this model, the output flow is not Poisson. In Section 5 we set up the Markov process associated with the model and determine a stationary range of the parameters. We prove that a stationary distribution exists if $2\lambda < \mu$. The exact range of parameters permitting an equilibrium distribution is not determined — it appears to depend on N_A and N_B. Calculations to prove that the output flow is not Poisson in equilibrium are carried out in Section 6.

Section 2 deals with some elementary probabilistic preliminaries. Concluding remarks are made in Section 7.

2. Preliminaries

Let $(X_t, t \geqq 0)$ be a stationary Markov process with countable state space X, infinitesimal generator Q, and equilibrium distribution π. Given a subset D of $\{(i, j)$ such that $i, j \in X$ and $i \neq j\}$, define the counting process $(D_t, t \geqq 0)$ by

$$(2.1)\qquad \begin{aligned} D_t &= \sum_{0<s\leq t} 1\{(X_{s-}, X_s)\in D\}, \\ D_0 &= 0. \end{aligned}$$

Define the matrices Q_D and Q^D by

$$(2.2a)\qquad Q_D(i,j) = Q(i,j)1\{(i,j)\notin D\}$$

and

$$(2.2b)\qquad Q^D(i,j) = Q(i,j)1\{(i,j)\in D\}$$

for each $i, j\in X$. Thus Q^D gives the rates with which $(X_t, t\geqq 0)$ makes transitions in D and $(D_t, t\geqq 0)$ counts the transitions in D.

Lemma 2.1. Let P_t denote the column vector with entries $P_t(i) = \Pr(D_t = 0 \mid X_0 = i)$. Then

$$(2.3)\qquad \dot{P}_t = Q_D P_t.$$

Proof.

$$\begin{aligned} P_{t+dt}(i) &= \Pr(D_{t+dt} = 0 \mid X_0 = i) \\ &= \sum_{j\in X} \Pr(D_{t+dt} = D_{dt}, D_{dt} = 0, X_{dt} = j \mid X_0 = i) \\ &= \sum_{j\in X} \Pr(D_{t+dt} = D_{dt} \mid X_{dt} = j, D_{dt} = 0, X_0 = i)\Pr(X_{dt} = j, D_{dt} = 0 \mid X_0 = i). \end{aligned}$$

Using the Markov property, this equals

$$\begin{aligned} &\sum_{j\in X} \Pr(D_{t+dt} = D_{dt} \mid X_{dt} = j)\Pr(X_{dt} = j, D_{dt} = 0 \mid X_0 = i) \\ &\quad = \sum_{j\in X} P_t(j) Q_D(i,j)dt + P_t(i) + o(dt) \end{aligned}$$

from which the claim follows.

Lemma 2.2. Let ξ_t denote the row vector with entries $\xi_t(i) = \Pr(X_t = i \mid D_t = 0)$. Let $h_t = \xi_t Q^D \mathbf{1}$, where $\mathbf{1}$ denotes a column vector of 1's. Then

$$(2.4)\qquad \dot{\xi}_t = \xi_t Q_D + h_t \xi_t.$$

Proof.

$$\begin{aligned} \xi_{t+dt}(i) &= \Pr(X_{t+dt} = i \mid D_{t+dt} = 0) \\ &= \Pr(X_{t+dt} = i \mid D_{t+dt} = D_t, D_t = 0) \\ &= \Pr(X_{t+dt} = i, D_{t+dt} = D_t \mid D_t = 0)\{\Pr(D_{t+dt} = D_t \mid D_t = 0)\}^{-1} \end{aligned}$$

by Bayes' rule.

Therefore

$$\xi_{t+dt}(i)(1+h_t dt+o(dt)) = \sum_{j\in X} \Pr(X_{t+dt}=i, D_{t+dt}=D_t, X_t=j \mid D_t=0)$$

$$= \sum_{j\in X} \xi_t(j)\Pr(X_{t+dt}=i, D_{t+dt}=D_t \mid X_t=j, D_t=0)$$

$$= \xi_t(j)Q_D(i,j)dt + \xi_t(i) + o(dt)$$

from which the claim follows.

Equation (2.4) is the continuous part of the filtering equations for point process functionals of a Markov chain — [1], [7], [8], [10]. Note that if $(D_t, t \geqq 0)$ is Poisson with intensity η, then $h_t \equiv \eta$.

3. The model with finite buffers

The Markov process associated to the model with finite buffers has the finite state space

$$X = \{(a,b) \text{ such that } 0 \leqq a \leqq N_A, 0 \leqq b \leqq N_B\}$$

and the transition matrix Q is easily written down. It is irreducible, therefore admits a unique stationary distribution. In the following theorem we give an easily verifiable necessary condition for the point process counting a subset of transitions of a countable-state Markov process to be Poisson.

Theorem 3.1. Let $(X_t, t \geqq 0)$ be a stationary Markov chain with countable state space X, infinitesimal generator Q, and invariant distribution π. Let D be a subset of transitions and $(D_t, t \geqq 0)$ the associated point process, as in Section 2. If Q^D (Equation (2.2b)) is nilpotent then $(D_t, t \geqq 0)$ is not Poisson.

Proof. Let P_t denote the column vector with entries $P_t(i) = \Pr(D_t = 0 \mid X_0 = i)$. If $(D_t, t \geqq 0)$ is to be Poisson with intensity η, we have

$$\pi P_t = \exp(-\eta t). \tag{3.1}$$

Differentiating (3.1) k times with respect to t and using Lemma 2.1 gives

$$\dot{\pi}(Q_D)^k P_t = (-\eta)^k \exp(-\eta t) \tag{3.2}$$

where we may set $t = 0$ to get

$$\pi(Q_D)^k \mathbf{1} = (-\eta)^k. \tag{3.3}$$

Now supose that, instead of counting every transition the Markov process makes in D, we choose $0 \leqq p \leqq 1$ and make, at each transition, an independent (of the past of $(X_t, t \geqq 0)$ at the instant of choice) decision of counting the transition with probability p. Thus we get a point-process functional $(D_t^p, t \geqq 0)$ of the Markov process, which is a Bernoulli sampled version of $(D_t, t \geqq 0)$. Since a p-Bernoulli sampling of a Poisson process with intensity η yields a Poisson process of intensity ηp, $(D_t^p, t \geqq 0)$ is Poisson with intensity ηp.

Let P_t^p denote the column vector with entries $P_t^p = \Pr(D_t^p = 0 \mid X_0 = i)$. Exactly as in Lemma 2.1 we can show that

$$\dot{P}_t^p = (Q_D + qQ^D)P_t^p \tag{3.4}$$

where $q = 1 - p$. Using (3.4) and arguing as above gives

$$\pi(Q_D + qQ^D)^k\mathbf{1} = (-\eta p)^k \tag{3.5}$$

for all $k \geqq 0$.

Now think of Equation (3.5) as a family of equations parametrized by p and differentiate k times with respect to p. It is easily checked that this gives

$$\pi(Q^D)^k\mathbf{1} = \eta^k \tag{3.6}$$

for all $k \geqq 0$.

From Equation (3.6) we conclude

(a) When X is finite, η is an eigenvalue of Q^D. In fact, if $\mathrm{Min}(x)$ denotes the minimal polynomial of Q^D, so that $\mathrm{Min}(Q^D) = 0$ and $\mathrm{Min}(x)$ is the polynomial of least degree achieving this, we see from Equation (3.6) that $\mathrm{Min}(\eta) = 0$.

(b) If Q^D is nilpotent, i.e. some power of it is 0, the point process associated to D cannot be Poisson.

To apply the above to the output of the model with finite buffers, note that the subset D of transitions which gives the output process has an associated Q^D which is nilpotent. Thus the output process is not Poisson.

Remark. As an application of the above, neither process of accepted arrivals is Poisson, since they are point process functionals corresponding to the sets of transitions

$$A = \{((a, b), (a + 1, b)) \text{ such that } 0 \leqq a \leqq N_A - 1, 0 \leqq b \leqq N_B\}$$

for process of accepted arrivals of type A, and

$$B = \{((a, b), (a, b + 1)) \text{ such that } 0 \leqq a \leqq N_A, 0 \leqq b \leqq N_B - 1\}$$

for type B, and the appropriate matrices are nilpotent. Theorem 3.1 may be applied to make a number of similar statements about the flows of various queues.

4. The model with infinite buffers

In this section we study the model of Section 3 with either one or both of N_A and N_B infinite. For the sake of definiteness assume that N_A is infinite. Also assume $1 \leqq N_B \leqq \infty$ to avoid triviality. We show below that the Markov process associated with this model does not admit a stationary distribution.

The state space associated with the model is $X = \{(a, b)$ such that $0 \leqq a \leqq \infty, 0 \leqq b \leqq N_B\}$, and the infinitesimal generator Q is easily written down. Suppose the process admits the invariant distribution π. Define $D_n \subseteq X$, $n \geqq 0$, by

$$D_n = \{(a, b) \in X \text{ such that } a - b = n\},$$

so that D_n is the subset of states where the excess of customers of type A over those of type B is n. From the balance of probability flow in equilibrium, one easily sees that

$$(4.1) \qquad \sum_{(a,b)\in D_n} \pi(a,b)\lambda \leqq \sum_{(a,b)\in D_{n+1}} \pi(a,b)\lambda$$

for each $n \geqq 0$.

Since $\sum_{(a,b)\in X} \pi(a,b) = 1$, this can only hold if

$$(4.2) \qquad \pi(a,b) = 0$$

for each $(a,b) \in \bigcup_{n\geqq 0} D_n$.

It is well known that if an irreducible Markov process admits an invariant distribution, the distribution is strictly positive on the state space. Our process is clearly irreducible, so that the above contradiction proves it cannot admit an invariant distribution. Note that the size of μ/λ was completely irrelevant to the above argument.

More generally, we can show that any countable-state Markovian model of a manufacturing node without blocking of inputs cannot admit an invariant distribution. Let $(X_t, t \geqq 0)$ be Markov, with countable state space X. The total number of components of type A is given by a function $g_A : X \to \mathbf{Z}_+$. Similarly, we have $g_B : X \to \mathbf{Z}_+$. Arrivals and departures correspond to subsets of transitions $A, B, D \subseteq \{(i,j)$ such that $i, j \in X$ and $i \neq j\}$, where

$$(i,j)\in A \quad \text{iff } g_A(j) = g_A(i) + 1,$$

$$(i,j)\in B \quad \text{iff } g_B(j) = g_B(i) + 1,$$

$$(i,j)\in D \quad \text{iff } g_A(j) = g_A(i) - 1 \text{ and } g_B(j) = g_B(i) - 1.$$

The arrival processes

$$A_t = \sum_{0<s\leqq t} 1\{(X_{s-}, X_s)\in A\}$$

and

$$B_t = \sum_{0<s\leqq t} 1\{(X_{s-}, X_s)\in B\}$$

are independent Poisson processes of rate λ.

It suffices to observe that, under the above assumptions, $(g_A(X_t) - g_B(X_t),\ t \geqq 0)$ executes a simple symmetric random walk on the integers.

5. The model with finite buffers for the excess

The state space X of the model with finite buffers for the excess is

$$(5.1) \qquad X = \{(a,b) \text{ such that } a, b \geqq 0,\ a - b \leqq N_A \text{ and } b - a \leqq N_B\}.$$

The off-diagonal entries of the infinitesimal generator are the rates of the various transitions and are given by

(5.2a) $$Q((a,b),(a+1,b)) = \lambda 1\{a-b \neq N_A\},$$

(5.2b) $$Q((a,b),(a,b+1)) = \lambda 1\{b-a \neq N_B\},$$

(5.2c) $$Q((a,b),(a-1,b-1)) = \mu 1\{a>0\}1\{b>0\}.$$

In this section we prove that the process admits a stationary distribution when $2\lambda < \mu$. To do this we consider the imbedded discrete-time Markov chain (x_n), which has the same state space X, and has transition matrix Γ with

(5.3) $$\Gamma(i,j) = \frac{Q(i,j)}{Q(i)},$$

where

(5.4) $$Q(i) = \sum_{j \in X, j \neq i} Q(i,j),$$

and prove that (x_n) is a positive recurrent chain. Since the original Markov process is stable and conservative, the existence of an invariant distribution for the process is equivalent to the existence of an invariant distribution for the imbedded chain. We need the following generalization of Pakes' lemma, whose ordinary version, applicable to chains with state space the non-negative integers, is proved in, e.g., [9].

Lemma 5.1. Let $(x_n, n \geqq 0)$ be an irreducible Markov chain, with state space X. Let there be given a function $g: X \to Z_+$, such that, for each $N>0$, $\{x \in X$ such that $g(x) \leqq N\}$ is finite. For $i \in X$, define

(5.5) $$\gamma(i) = E(g(x_{n+1}) - g(x_n) \mid x_n = i).$$

Suppose

(a) $|\gamma(i)| < \infty$, for all $i \in X$.

(b) There is $N>0$ and $-a<0$ such that, for each $i \in X$ with $g(i) \geqq N$, we have $\gamma(i) \leqq -a$.

Then $(x_n, n \geqq 0)$ is a positive recurrent chain. In particular, it admits an invariant distribution.

Remark. We take as known that for an irreducible Markov chain of period d, and $i, j \in X$

$$\lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} \Pr(x_n = j \mid x_0 = i) = \pi(j)$$

exists. Also, $\pi(j) > 0$ if and only if the chain is positive recurrent, admitting invariant distribution $\pi(\cdot)$ (and equals 0 otherwise).

Proof. Let $b = \max\{\gamma(i)$ such that $g(i) < N\}$. Let $E_i\{\cdot\}$ denote $E(\cdot \mid x_0 = i)$. We have

$$-g(i) \leqq E_i\{g(x_{n+1}) - g(x_0)\} = E_i\{\gamma(x_0) + \cdots + \gamma(x_n)\}$$

$$\leqq (a+b) \sum_{m=0}^{n} P_i(g(x_m) < N) - (n+1)a$$

where now P_i denotes probabilities starting with $x_0 = i$.

Divide by n and let $n \to \infty$. We get

$$(a+b) \sum_{\{i \text{ such that } g(i) < N\}} \pi(i) \geqq a > 0,$$

where we have used the fact that $\{i \text{ such that } g(i) < N\}$ is finite. This proves that $\pi(\cdot)$ cannot be identically zero, so that the chain is positive recurrent.

To apply the above, let $X_n \subseteq X$ be given by $X_n = \{(a, b) \in X \text{ such that } \min\{a, b\} = n\}$. Define the gauge $g: X \to Z_+$ by: $g(x) = n$ if and only if $x \in X_n$. If $2\lambda < \mu$ then $g(\cdot)$ satisfies the conditions of the lemma. Also, our chain is irreducible and periodic with period 3.

6. Intensity calculations

Suppose $(D_t, t \geqq 0)$ is Poisson with intensity η. We first observe that

$$\eta = \mu \sum_{(a,b) \in X} \pi(a, b) 1\{a > 0\} 1\{b > 0\}. \tag{6.1}$$

But

$$\lambda \sum_{i \in X_n} \pi(i) > \mu \sum_{i \in X_{n+1}} \pi(i). \tag{6.2}$$

Summing over all n gives

$$\lambda > \eta. \tag{6.3}$$

This will be used below.

Let ξ_t denote the row vector with entries $\xi_t(i) = \Pr(X_t = i \mid D_t = 0)$. From Lemma 2.2,

$$\begin{aligned} \xi_t(a,b) = {} & \lambda \xi_t(a-1,b) 1\{a > 0\} 1\{a - 1 - b < N_A\} \\ & + \lambda \xi_t(a, b-1) 1\{b > 0\} 1\{b - 1 - a < N_B\} \\ & - \xi_t(a, b)(\lambda 1\{a - b < N_A\} + \lambda 1\{b - a < N_B\} + \mu 1\{a > 0\} 1\{b > 0\}) \\ & + \xi_t(a, b)\eta. \end{aligned} \tag{6.4}$$

Sum over $\{(a, b) \text{ such that } a, b > 0\}$ to get

$$
\begin{aligned}
0=\frac{h_t}{\mu}=\lambda \sum_{0<b\leq N_B} \xi_t(0,b)+\lambda \sum \xi_t(a,b)1\{a>0\}1\{b>0\}1\{a-b<N_A\} \\
+\lambda \sum_{0<a\leq N_A} \xi_t(a,0)+\lambda \sum \xi_t(a,b)1\{a>0\}1\{b>0\}1\{b-a<N_B\} \\
-\sum \xi_t(a,b)1\{a>0\}1\{b>0\}(\lambda 1\{a-b<N_A\}+\lambda 1\{b-a<N_B\} \\
+\mu 1\{a>0\}1\{b>0\})+\frac{\eta^2}{\mu}.
\end{aligned} \tag{6.5}
$$

Simplifying,

$$
0=\lambda \sum_{0<b\leq N_B} \xi_t(0,b)+\lambda \sum_{0<a\leq N_A} \xi_t(a,0)-\eta+\frac{\eta^2}{\mu}. \tag{6.6}
$$

We see from the above equation that $\Sigma_{0<b\leq N_B} \xi_t(0,b)+\Sigma_{0<a\leq N_A} \xi_t(a,0)$ is constant. Now write

$$
\sum_{0<b\leq N_B} \dot{\xi}_t(0,b)=\lambda \sum_{0\leq b<N_B} \xi_t(0,b)-2\lambda \sum_{0<b<N_B} \xi_t(0,b)-\lambda \xi_t(0,N_B)+\eta \sum_{0<b\leq N_B} \xi_t(0,b),
$$

which can be simplified to give

$$
\sum_{0<b\leq N_B} \dot{\xi}_t(0,b)=\lambda \xi_t(0,0)+(\eta-\lambda) \sum_{0<b\leq N_B} \xi_t(0,b). \tag{6.7a}
$$

Similarly we get

$$
\sum_{0<a\leq N_A} \dot{\xi}_t(a,0)=\lambda \xi_t(0,0)+(\eta-\lambda) \sum_{0<a\leq N_A} \xi_t(a,0). \tag{6.7b}
$$

Summing Equations (6.7a) and (6.7b) or taking them individually as appropriate, we see that $\xi_t(0,0)$ is constant. (Note that at least one of N_A and N_B must exceed 0 to avoid triviality.)

However

$$
\dot{\xi}_t(0,0)=(\eta-2\lambda)\xi_t(0,0) \tag{6.8a}
$$

if $N_A>0$ and $N_B>0$, and

$$
\dot{\xi}_t(0,0)=(\eta-\lambda)\xi_t(0,0) \tag{6.8b}
$$

if one of N_A and N_B is 0.

In either case, since $\eta<\lambda$, the above implies that $\xi_t(0,0)$ cannot be constant. This contradiction establishes that the output flow of the manufacturing system cannot be Poisson in equilibrium.

7. Concluding remarks

In the preceding we have considered several attempts at defining a manufacturing node which admits a stationary situation yielding a Poisson output for Poisson inputs. Our results are negative, in that several natural definitions are shown to not possess the property we seek. Perhaps there is no such model (?). There has been some recent work on measuring the distance of a process from a Poisson process in terms of the intensity [3]. An interesting direction to go from here would be to quantify the deviation of the output process of the above models from a Poisson process.

References

[1] BRÉMAUD, P. (1981) *Point Processes and Queues*. Springer-Verlag, Berlin.

[2] BURKE, P. J. (1956) The output of a queueing system. *Operat. Res.* **4**, 699–704.

[3] BROWN, T. C. AND POLLETT, P. K. (1982) Some distributional approximations in Markovian queueing networks. *Adv. Appl. Prob.* **14**, 654–671.

[5] KELLY, F. (1975) Networks of queues with customers of different types. *J. Appl. Prob.* **12**, 542–554.

[6] MELAMED, B. (1979) On Poisson transition processes in discrete state Markovian systems with applications to queueing theory. *Adv. Appl. Prob.* **11**, 218–239.

[7] WALRAND, J. (1982) On the equivalence of flows in networks of queues. *J. Appl. Prob.* **19**, 195–203.

[8] WALRAND, J. (1981) Filtering formulas and the $M/M/1$ queue in a quasireversible network. *Stochastics* **6**, 1–22.

[9] WALRAND, J. (1985) *Lectures on Queueing Theory* (in preparation).

[10] WALRAND, J. AND VARAIYA, P. (1981) Flows in queueing networks: A martingale approach. *Math. Operat. Res.* **6**, 387–404.