Analytic models for performance evaluation of single-buffered banyan networks under nonuniform traffic

T.-H. Lee

Indexing terms: Networks, Modelling

Abstract: The performance of single-buffered banyan networks under certain nonuniform traffic patterns had been studied by Garg and Huang. However, the models used are over simplified and the results obtained may deviate from exact values significantly. Alternative models to achieve more accurate performance estimates are presented. In our models, the destinations of blocked packets residing in the buffers of nodes at stage 1 (and perhaps stage 2, depending on the traffic matrix) are memorised. Compared with those adopted by Garg and Huang, our models are only slightly more complicated. By viewing banyan networks as queueing systems, we apply Little's formula to compute the average packet delays.

1 Introduction

Because of the nice properties such as self-routing, potential VLSI implementation, and easiness of fault diagnosis, multistage banyan networks have attracted increasing interest from researchers and engineers in the areas of multiprocessor systems and telecommunications. Goke and Lipovski defined in Reference 1, a general class of banyan networks. Here we are interested in only the regular SW banyans with spread and fan-out of two. It was proved [4] that the flip network, the omega network, the indirect binary n-cube network, the baseline network, and the regular SW banyan network with spread and fan-out of two are all isomorphic. Therefore, we will not distinguish these terms in this paper.

Patel [2] derived a simple recursive formula to compute the normalised throughput of unbuffered multistage banyan networks under the uniform traffic model. By uniform traffic model it is meant that inlets (sources) generate packets independently with identical rates and each outlet (destination) is equally likely to be the destination of any packet. It was shown that banyan networks are more cost-effective than single-stage crossbar networks for processor-memory interconnection in large multiprocessor systems.

Dias and Jump [5] studied the effect of adding buffers to input links of each switching element of a banyan network. They conclude that, for most applications, the number of buffers between stages should be limited to one or two. Jenq [6] proposed a similar analytic model to estimate the normalised throughput and the average packet delay for single-buffered banyan networks (see Fig. 1). The analyses were also based on the uniform traffic assumption. This assumption, which simplifies the analysis, may not be true for real world systems. Garg and Huang [7] modified the model used by Jenq to study the performance of single-buffered banyan networks under certain nonuniform traffic patterns. Similar work was done by Kim and Leon-Garcia [8]. In general, a higher degree of nonuniformity of traffic flow results in a worse performance.

Unfortunately, the models used by Garg and Huang to analyse the performance of single-buffered banyan networks under nonuniform traffic result in significant deviations from exact values. For example, consider a nonuniform traffic matrix of form I studied in their paper. The ratio of the normalised high-traffic throughput to the normalised low-traffic throughput should be m_1/m_2. The reason is that the banyan network can be viewed as a queueing system and the normalised high-traffic and low-traffic throughputs are, respectively, equal to the effective arrival rates of high-traffic and low-traffic packets. According to the traffic matrix and the single-buffer assumption, the ratio of the effective arrival rate of high-traffic packets to that of low-traffic packets is equal to m_1/m_2. Their results obviously are not consistent with this fact.

In this paper we present alternative models to achieve more accurate estimates. In our new models, the destinations of blocked packets in buffers of nodes at stage 1 (and perhaps stage 2) are memorised or stored. Compared with the models adopted by Garg and Huang, our new models are only slightly more complicated. The

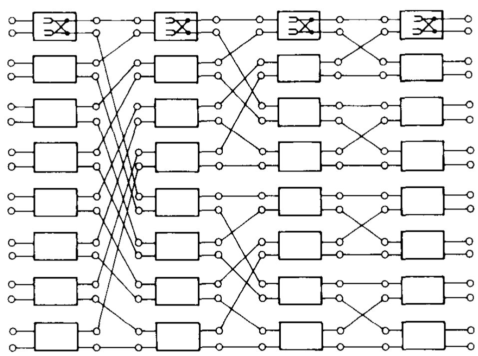

Fig. 1 *Four-stage single-buffered banyan network*

Paper 7695E (C3), first received 2nd January and in revised form 10th July 1990

The author is with the Department of Communication Engineering, National Chiao Tung University, Hsinchu, Taiwan 30050, Republic of China

same forms of nonuniform traffic matrices discussed in Reference 7 are analysed in this paper using our new models. Little's formula [10] is used to compute the average packet delays. As will be seen later, the normalised throughputs obtained from our new models are consistent with the fact mentioned in the last paragraph.

2 Nonuniform traffic matrices: form I

The first form of nonuniform traffic matrices we will analyse looks like

$$D = \begin{bmatrix} D_1(m_1) & D_2(m_2) \\ D_2(m_2) & D_1(m_1) \end{bmatrix}$$

where D represents the traffic matrix and is nonuniform, but the submatrices D_1 and D_2 are uniform. The row sums of D_1 and D_2 are equal to m_1 and m_2, respectively, and moreover, $m_1 + m_2$ is equal to 1. It should be pointed out that the (i, j)th element of the traffic matrix D represents the probability that a packet originated at the ith inlet is destined to the jth outlet.

Under nonuniform traffic matrices of form I, the sources and destinations of a banyan network are both clustered into two groups. Source (destination) groups 1 and 2 consist of sources (destinations) located in the upper and the lower half of the banyan network, respectively. A packet, originated at source group 1, is destined to destination group 1 with probability m_1 (called high-traffic) and is destined to destination group 2 (called low-traffic) with probability m_2. Conversely, a packet, originated at source group 2, is destined to destination group 1 with probability m_2 (low-traffic) and is destined to destination group 2 (high-traffic) with probability m_1. As a consequence, the ratio of the effective arrival rate of high-traffic packets to that of low-traffic packets is equal to m_1/m_2. Traffic matrices of form I may appear in several application areas. For example, in a communication network, the upper half and the lower half sources and destinations may come from different geographical areas and hence the traffic flows inside the same area may be different from that between different areas. As a result, a nonuniform traffic matrix of form I can be used to describe the traffic pattern if both traffic flows inside the same area and between different areas are uniform.

For simplicity, the network is assumed operating in a synchronous mode. A network cycle is defined as a time unit required for a packet to move forward one stage. Notice that a packet in a buffer at stage k is able to move forward one stage in the lth cycle if it is not blocked at stage k, and its destination buffer space at stage $k + 1$ is available at the end of the lth cycle. As usual, the inlets are assumed to generate packets independently with identical rates denoted by ρ. The following variables, which will be used to derive the equilibrium equations, are defined at the end of network cycles at the steady state.

Consider a node at stage 1 and let

a_0 = Prob{both buffers are empty}
a_{1h} = Prob{one buffer is empty and the other has a high-traffic packet}
a_{1l} = Prob{one buffer is empty and the other has a low-traffic packet}
a_{hl} = Prob{one buffer has a low-traffic packet and the other has a high-traffic packet}
a_{2h} = Prob{both buffers have high-traffic packets}
a_{2l} = Prob{both buffers have low-traffic packets}
r_h = Prob{a high-traffic packet is able to move to stage 2}
r_l = Prob{a low-traffic packet is able to move to stage 2}
t_h = Prob{the high-traffic packet is able to move to stage 2 but the low-traffic packet is not, conditioning on one buffer has a high-traffic packet and the other has a low-traffic packet}
t_l = Prob{the low-traffic packet is able to move to stage 2 but the high-traffic packet is not, conditioning on one buffer has a high-traffic packet and the other has a low-traffic packet}
t_{hl} = Prob{both high-traffic and low-traffic packets are able to move to stage 2, conditioning on one buffer has a high-traffic packet and the other has a low-traffic packet}

For nodes at stage k, $2 \leqslant k \leqslant n$, let

$p_{h0}(k)$ = Prob{high-traffic buffer is empty}
$p_{l0}(k)$ = Prob{low-traffic buffer is empty}
$p_h(k) = 1 - p_{h0}(k)$
$p_l(k) = 1 - p_{l0}(k)$
$r_h(k)$ = Prob{a packet in a high-traffic buffer is able to move forward}
$r_l(k)$ = Prob{a packet in a low-traffic buffer is able to move forward}
$q_h(k)$ = Prob{a packet is ready to come to a high-traffic buffer}
$q_l(k)$ = Prob{a packet is ready to come to a low-traffic buffer}.

Notice that r_h and r_l should be interpreted differently from $r_h(k)$ and $r_l(k)$ $(k \geqslant 2)$. The value of r_h (or r_l) is equal to the probability that the destination buffer space at stage 2 is available at the end of a network cycle. However, to compute the value of $r_h(k)$ (or $r_l(k)$), we have to multiply the probability that the destination buffer space at stage $k + 1$ is available (this probability is equal to 1 for $k = n$) with the probability that the high-traffic (low-traffic) packet is not blocked at stage k. Given the above notation, the set of equilibrium equations for the system can be derived. For convenience, let

$$A = a_0 + a_{1h} r_h + a_{1l} r_l + a_{hl} t_{hl}$$

i.e. A is equal to the probability that both buffers of a node at stage 1 are available at the end of a network cycle. Then we have

$$a_0 = A(1 - \rho)^2$$

$$a_{1h} = A[2\rho(1 - \rho)m_1] + [a_{1h}(1 - r_h) + a_{hl} t_l + a_{2h} r_h](1 - \rho)$$

$$a_{1l} = A[2\rho(1 - \rho)m_2] + [a_{1l}(1 - r_l) + a_{hl} t_h + a_{2l} r_l](1 - \rho)$$

$$a_{hl} = A(\rho^2 m_1 m_2) + a_{1h}(1 - r_h)\rho m_2 + a_{1l}(1 - r_l)\rho m_1 + a_{hl}(t_h \rho m_1 + t_l \rho m_2 + 1 - t_h - t_l - t_{hl}) + a_{2h}(r_h \rho m_2) + a_{2l}(r_l \rho m_1)$$

$$a_{2h} = A\rho^2 m_1^2 + a_{1h}(1 - r_h)\rho m_1 + a_{hl}(t_l \rho m_1) + a_{2h}(r_h \rho m_1 + 1 - r_h)$$

$$a_{2l} = 1 - a_0 - a_{1h} - a_{1l} - a_{hl} - a_{2h}$$

$$r_h = p_{h0}(2) + p_h(2) r_h(2)$$

$$r_l = p_{l0}(2) + p_l(2) r_l(2)$$

$$t_h = r_h(1 - r_l)$$

$$t_l = r_l(1 - r_h)$$

$$t_{hl} = r_h r_l$$

$$q_h(2) = a_{1h} + a_{hl} + a_{2h}$$

$$q_l(2) = a_{1l} + a_{hl} + a_{2l}$$

$$q_h(k) = 0.75p_h^2(k-1) + p_{h0}(k-1)p_h(k-1)$$

$$q_l(k) = 0.75p_l^2(k-1) + p_{l0}(k-1)p_l(k-1)$$

$$r_h(k) = [p_{h0}(k) + 0.75p_h(k)] \times [p_{h0}(k+1) + p_h(k+1)r_h(k+1)]$$

$$r_l(k) = [p_{l0}(k) + 0.75p_l(k)] \times [p_{l0}(k+1) + p_l(k+1)r_l(k+1)]$$

$$r_h(n) = p_{l0}(n) + 0.75p_l(n)$$

$$r_l(n) = p_{h0}(n) + 0.75p_h(n)$$

$$p_{h0}(k) = [1 - q_h(k)][p_{h0}(k) + p_h(k)r_h(k)]$$

$$p_{l0}(k) = [1 - q_l(k)][p_{l0}(k) + p_l(k)r_l(k)]$$

normalised high throughput:

$$S_h = p_h(n)r_h(n)$$

normalised low throughput:

$$S_l = p_l(n)r_l(n)$$

average high delay:

$$D_h = (1/S_h)\left[a_{1h} + a_{hl} + 2a_{2h} + \sum_{k=2}^{n} p_h(k)\right]$$

average low delay:

$$D_l = (1/S_l)\left[a_{1l} + a_{hl} + 2a_{2l} + \sum_{k=2}^{n} p_l(k)\right]$$

It can be seen that our new model for form I nonuniform traffic matrices is different from the one used by Garg and Huang only in the description of buffer conditions of nodes at stage 1. The destinations of blocked packets residing in the buffers of nodes at stage 1 are memorised in our model. The buffers at other stages can obviously be similarly modelled. However, our experience show that memorising the destinations of all the blocked packets will dramatically increase computational complexity without significant improvement. For example, for a four-stage banyan network with $m_1 = m_2 = 0.5$, the normalised throughput and the average packet delay are 0.523 and 5.96, respectively, for the proposed model; and are 0.511 and 6.10, respectively, for the more complicated model.

Notice that Little's formula is applied to compute the average packet delays. At the steady state, the effective arrival rate of high-traffic packets to the banyan network is equal to $2^{n-1}S_h$ and the average number of high-traffic packets in system is equal to

$$2^{n-1}(a_{1h} + a_{hl} + 2a_{2h}) + 2^{n-1}\sum_{k=2}^{n} p_h(k)$$

Hence the average delay for high-traffic packets, denoted by D_h, is equal to

$$(1/S_h)\left[a_{1h} + a_{hl} + 2a_{2h} + \sum_{k=2}^{n} p_h(k)\right]$$

The average delay for low-traffic packets can be similarly derived. Notice again that, in our new model (and the one used by Garg and Huang), buffers at different stages are assumed to be independent. Without this assumption, the modelling complexity is expected to increase dramatically.

3 Nonuniform traffic matrices: form II

In this Section, we will consider nonuniform traffic matrices of the following form:

$$D = [D_1(m_1) \vdots D_2(m_2)]$$

The nonuniform traffic matrix D is partitioned vertically and the submatrices D_1 and D_2 are uniform. The row sums of D_1 and D_2 are equal to m_1 and m_2, respectively, and $m_1 + m_2$ is equal to 1. Traffic matrices of form II can also appear in several application areas. For example, in a multiprocessor system, it may happen that the variables stored in the upper half memory modules are accessed more frequently than those stored in the lower half ones. In reality, the traffic matrix D studied here can be partitioned at the centre or not at the centre. We consider in the following two possible cases seaparately.

Case 1: partitioned at the centre: For this case, the outlets are clustered into two equal-size groups. A packet is called a high-traffic (or low-traffic) packet if its destination is located in the upper (lower) half. Similarly, the ratio of the normalised high-traffic throughput to the normalised low-traffic throughput should be equal to m_1/m_2.

Given the same variables defined in the previous Section, the set of equilibrium equations for this case can be derived. In reality, the set of equilibrium equations for this case, except the two for $r_h(n)$ and $r_l(n)$, are exactly the same as those for form I nonuniform traffic matrices. The expressions for $r_h(n)$ and $r_l(n)$ should be interchanged, i.e.

$$r_h(n) = p_{h0}(n) + 0.75p_h(n)$$

and

$$r_l(n) = p_{l0}(n) + 0.75p_l(n).$$

The normalised throughputs and the average delays of high-traffic and low-traffic packets can be computed using the same formulas given in the previous section.

Case 2: partitioned not at the centre: For the second case of form II nonuniform traffic matrices, the matrix D is partitioned not at the centre. Here we consider traffic matrices with the following partition:

$$D = [\underset{\text{group of upper 1/4 destinations}}{D_1(m_1)} \vdots \underset{\text{group of lower 3/4 destinations}}{D_2(m_2)}]$$

Fig. 2 illustrates such a partition for a four-stage banyan network. For convenience, packets destined to the upmost 1/4 destinations, the succeeding 1/4 destinations, and the lower 1/2 destinations are called high-traffic,

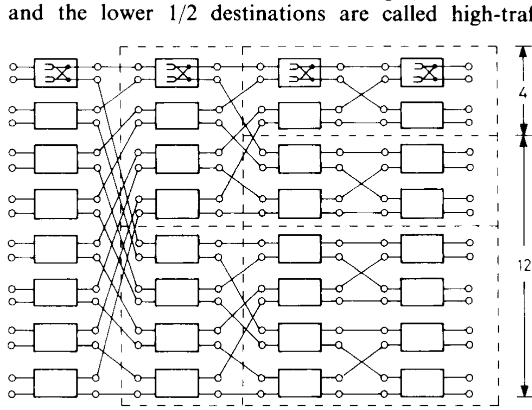

Fig. 2 *Illustration of a partition of form II, case 2*

middle-traffic, and low-traffic packets, respectively. The relationships among the normalised high-traffic, middle-traffic, and low-traffic throughputs will be discussed later. To obtain performance estimates under nonuniform traffic matrices of this case, more notation, in addition to those introduced in Section 2, are needed.

For the upper half nodes at the second stage, let

b_0 = Prob{both buffers are empty}
b_{1h} = Prob{one buffer is empty and the other has a high-traffic packet}
b_{1m} = Prob{one buffer is empty and the other has a middle-traffic packet}
b_{hm} = Prob{one buffer has a high-traffic packet and the other has a middle-traffic packet}
b_{2h} = Prob{both buffers have high-traffic packets}
b_{2m} = Prob{both buffers have middle-traffic packets}
x_h = Prob{a high-traffic packet is able to move to stage 3}
x_m = Prob{a middle-traffic packet is able to move to stage 3}
y_h = Prob{the high-traffic packet is able to move to stage 3 but the middle-traffic packet is not, conditioning on one buffer has a high-traffic packet and the other has a middle-traffic packet}
y_m = Prob{the middle-traffic packet is able to move to stage 3 but the high-traffic packet is not, conditioning on one buffer has a high-traffic packet and the other has a middle-traffic packet}
y_{hm} = Prob{both the high-traffic and the middle-traffic packets are able to move to stage 3, conditioning on one buffer has a high-traffic packet and the other has a middle-traffic packet}.

For nodes at stage k, $3 \leqslant k \leqslant n$, let

$p_{m0}(k)$ = Prob{middle-traffic buffer is empty}
$p_m(k) = 1 - p_{m0}(k)$
$q_m(k)$ = Prob{a packet is ready to come to a middle-traffic buffer}

Finally, let

$$f_1 = m_1 + m_2/3$$
$$f_2 = 2m_2/3$$
$$g_1 = m_1/f_1$$
$$g_2 = 1 - g_1$$
$$B = b_0 + b_{1h}x_h + b_{1m}x_m + b_{hm}y_{hm}$$

and

$$\rho_1 = a_{1h} + a_{hl} + a_{2h}$$

Then the set of equilibrium equations for this case are

$$a_0 = A(1-\rho)^2$$
$$a_{1h} = A[2\rho(1-\rho)f_1] + [a_{1h}(1-r_h) + a_{hl}t_l + a_{2h}r_h](1-\rho)$$
$$a_{1l} = A[2\rho(1-\rho)f_2] + [a_{1l}(1-r_l) + a_{hl}t_h + a_{2l}r_l](1-\rho)$$
$$a_{hl} = A(\rho^2 f_1 f_2) + a_{1h}(1-r_h)\rho f_2 + a_{1l}(1-r_l)\rho f_1 + a_{hl}[t_h\rho f_1 + t_l\rho f_2 + 1 - t_h - t_l - t_{hl}] + a_{2h}(r_h\rho f_2) + a_{2l}(r_l\rho f_1)$$
$$a_{2h} = A\rho^2 f_1^2 + a_{1h}(1-r_h)\rho f_1 + a_{hl}(t_l\rho f_1) + a_{2h}(r_h\rho f_1 + 1 - r_h)$$
$$a_{2l} = 1 - a_0 - a_{1h} - a_{1l} - a_{hl} - a_{2h}$$
$$\rho_1 = q_h(2) = a_{1h} + a_{hl} + a_{2h}$$
$$r_h = b_0 + b_{1h}(0.5 + 0.5x_h) + b_{1l}(0.5 + 0.5x_l) + b_{hl}(0.5x_h + 0.5x_l) + b_{2h}(0.5x_h) + b_{2l}(0.5x_l)$$
$$r_l = p_{l0}(2) + p_l(2)r_l(2)$$
$$b_0 = B(1-\rho_1)^2$$
$$b_{1h} = B[2\rho_1(1-\rho_1)g_1] + [b_{1h}(1-x_h) + b_{hm}y_m + b_{2h}x_h](1-\rho_1)$$
$$b_{1m} = B[2\rho_1(1-\rho_1)g_2] + [b_{1m}(1-x_m) + b_{hm}y_h + b_{2m}x_m](1-\rho_1)$$
$$b_{hm} = B(\rho_1^2 g_1 g_2) + b_{1h}(1-x_h)\rho_1 g_2 + b_{1m}(1-x_m)\rho_1 g_1 + b_{hm}(y_h\rho_1 g_1 + y_m\rho_1 g_2 + 1 - y_h - y_m - y_{hm}) + b_{2h}(x_h\rho_1 g_2) + b_{2m}(x_m\rho_1 g_1)$$
$$b_{2h} = B\rho_1^2 g_1^2 + b_{1h}(1-x_h)\rho_1 g_1 + b_{hm}(y_m\rho_1 g_1) + b_{2h}(x_h\rho_1 g_1 + 1 - x_h)$$
$$b_{2m} = 1 - b_0 - b_{1h} - b_{1m} - b_{hm} - b_{2h}$$
$$q_m(k) = 0.75p_m^2(k-1) + p_{m0}(k-1)p_m(k-1)$$
$$r_m(k) = [p_{m0}(k) + 0.75p_m(k)] \times [p_{m0}(k+1) + p_m(k+1)r_m(k+1)]$$
$$r_m(n) = p_{m0}(n) + 0.75p_m(n)$$
$$x_h = p_{h0}(3) + p_h(3)r_h(3)$$
$$x_m = p_{m0}(3) + p_m(3)r_m(3)$$
$$y_h = x_h(1-x_m)$$
$$y_m = x_m(1-x_h)$$
$$y_{hm} = x_h x_m$$

The equilibrium equations for $p_{h0}(k)$, $p_{l0}(k)$, $r_h(k)$, $r_l(k)$, $q_h(k)$, $q_l(k)$, r_h, r_l, t_h, t_l, and t_{hl} are exactly the same as those given in the previous Section. Notice that the equilibrium equation for $p_{h0}(2)$ given there is replaced by a set of equations.

The normalised throughputs and the average packet delays for an n-stage banyan network under nonuniform traffic matrices of form II, case 2, can be computed by normalised high throughput:

$$S_h = p_h(n)r_h(n)$$

normalised middle throughput:

$$S_m = p_m(n)r_m(n)$$

normalised low throughput:

$$S_l = p_l(n)r_l(n)$$

average high delay:

$$D_h = (1/S_h)\left[2(a_{1h} + a_{hl} + 2a_{2h})g_1 + b_{1h} + b_{hm} + 2b_{2h} + \sum_{k=3}^{n} p_h(k)\right]$$

average middle delay:

$$D_m = (1/S_m)\left[2(a_{1h} + a_{hl} + 2a_{2h})g_2 + b_{1m} + b_{hm} + 2b_{2m} + \sum_{k=3}^{n} p_m(k)\right]$$

average low delay:

$$D_l = (1/S_l)\left[a_{1l} + a_{hl} + 2a_{2l} + \sum_{k=2}^{n} p_l(k) \right]$$

Notice that, for this case, the destinations of blocked packets in buffers at stage 1 and upper half buffers at stage 2 are memorised to achieve more accurate performance estimates. However, the destinations of blocked packets in other buffers are not memorised to reduce the modelling complexity. According to the traffic matrix, the normalised middle-traffic throughput is equal to the normalised low-traffic throughput and the normalised high-traffic throughput is g_1/g_2 times that of the normalised middle-traffic throughput. Again, Little's formula is used to compute average packet delays.

Clearly, the set of equilibrium equations for other nonuniform traffic matrices of similar forms can be similarly derived. For example, the equilibrium equations for the following traffic matrix:

$$D = [\ \overset{\text{group of upmost 1/4 destinations}}{D_1(m_1)} \ \vdots\ \overset{\text{group of succeeding 1/4 destinations}}{D_2(m_2)} \ \vdots\ \overset{\text{group of lower 1/2 destinations}}{D_3(m_3)}\ ]$$

are exactly the same as those we listed above for this case. The only differences are now f_1 and f_2 should be replaced by $m_1 + m_2$ and m_3, respectively.

4 Numerical and simulation results

In this Section, we focus our attention on the performance of a four-stage banyan network under nonuniform traffic patterns studied in the previous Sections. Simulations are performed to verify our results.

Example 1 (form I): Figs. 3*a* and *b* show, respectively, the normalised throughputs and the average packet delays against input rate ρ for nonuniform traffic matrices of form I (the pair (m_1, m_2) is equal to (0.75, 0.25) or (0.6, 0.4)). Our results are consistent with the fact that the ratio of the normalised throughput of high-traffic packets to that of low-traffic packets is equal to m_1/m_2. Besides, it can be seen that analytic results are close to simulation results, especially when ρ is small. The performance degradation caused by nonuniform traffic flow can be observed in Figs. 3*c* and *d* (analytic results are used). One can see that a higher degree of nonuniformity results in a more serious performance degradation.

Example 2 (form II, case 1): The normalised throughputs and the average packet delays as functions of input rate ρ for nonuniform traffic matrices of form II, case 1, are plotted in Figs. 4*a–d*. Again, our results are consistent with the fact that the ratio of the normalised throughput

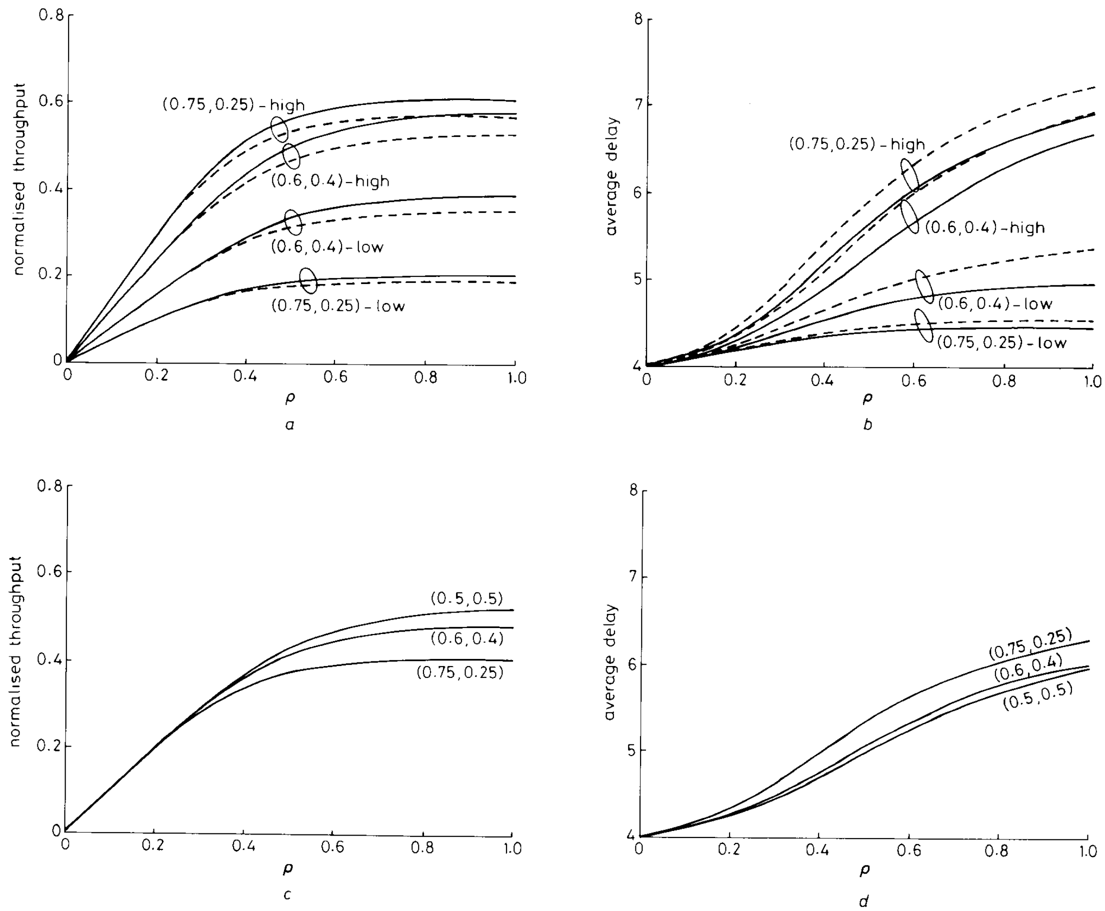

Fig. 3 *Normalised throughputs and average packet delays against input rate for nonuniform traffic matrices of form I*

a Normalised throughput against ρ
b Average packet delay against ρ
c Normalised throughput for the total network against ρ
d Average packet delay for the total network against ρ
——— analytic
– – – simulation

of high-traffic packets to that of low-traffic packets is equal to m_1/m_2. Furthermore, a higher degree of nonuniformity results in a greater loss of performance.

Example 3 (form II, case 2): Figs. 5*a–e* show the results for nonuniform traffic matrices of form II, case 2. The same conclusions as we stated in examples 1 and 2 can be drawn for this case.

It is noted that the performance estimates obtained from analytic models are consistently better than those obtained by simulations. The reason is that, in our models, buffers at different stages are assumed to be independent. However, in a real system, if the buffers at stage i contain blocked packets, then it is likely that packets in buffers at previous stages will also be blocked. If we consider the simulation results as true values, then the maximal relative error for the above three examples is about 10%. An approach to obtain better estimates of the normalised throughputs (and the average packet delays) is to first find the ratios of the normalised throughputs (or the average packet delays) obtained from simulation to those obtained from analytic models at $\rho = 0.5$ and then multiply these ratios to analytic values for $0.5 \leqslant \rho \leqslant 1$. By doing such, a maximal relative error less than 5% can be achieved at the expense of one simulation.

5 Conclusions

We have presented in this paper alternative models to achieve more accurate performance estimates of single-buffered banyan networks under certain nonuniform traffic patterns. Our models are only slightly more complicated than those adopted in Reference 7. For example, for an n-stage banyan network with a nonuniform traffic matrix of form I, the number of unknowns to be solved are equal to $8n - 5$ and $8n + 3$ for the model used in Reference 7 and our model, respectively. More complicated models in which the destination of each blocked packet is memorised can be used to obtain even better estimates. However, our experience shows that the more complicated models require much higher computational loads than the models we adopted whereas no significant improvement can be achieved. An approach is proposed to compensate for the independence assumption among buffers at different stages at the expense of one simulation. How to accurately model the dependence among buffers remains to be explored.

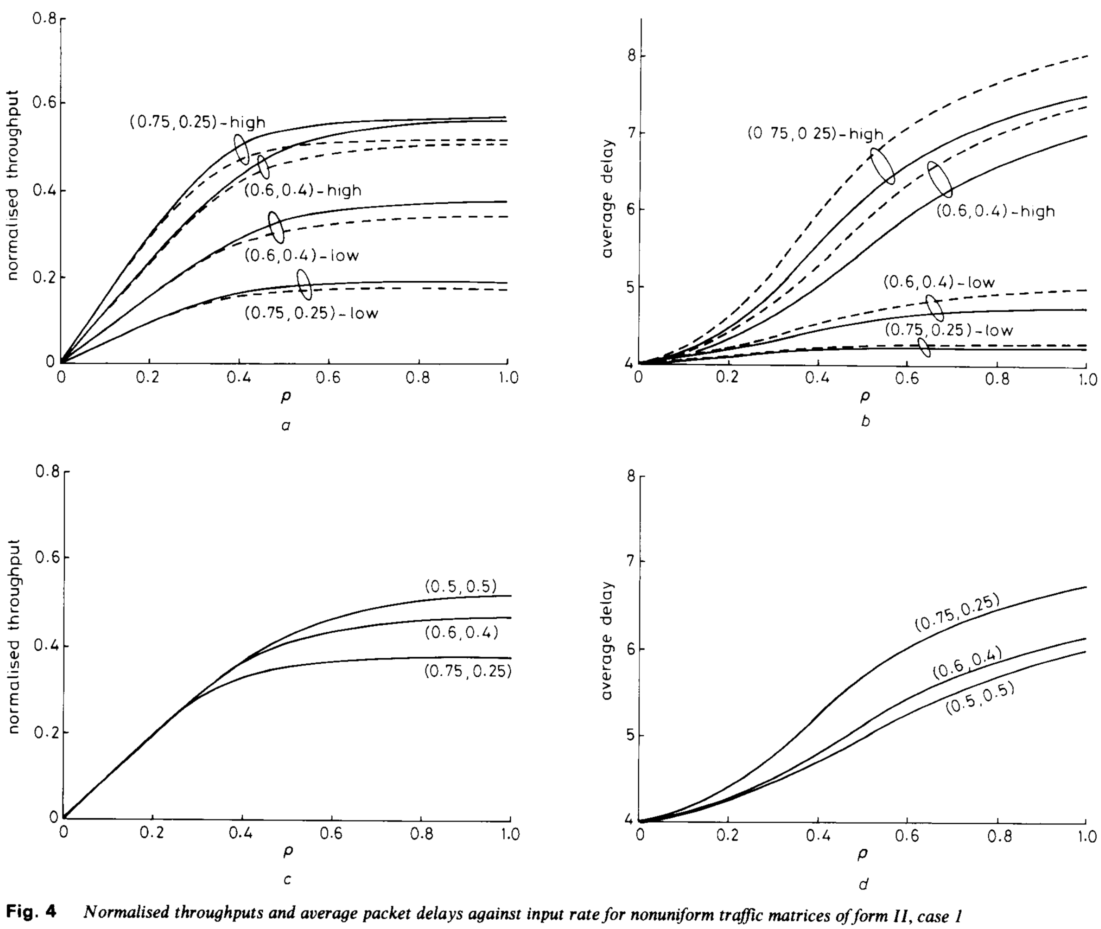

Fig. 4 *Normalised throughputs and average packet delays against input rate for nonuniform traffic matrices of form II, case 1*
a Normalised throughput against ρ
b Average packet delay against ρ
c Normalised throughput for the total network against ρ
d Average packet delay for the total network against ρ
—— analytic
- - - - simulation

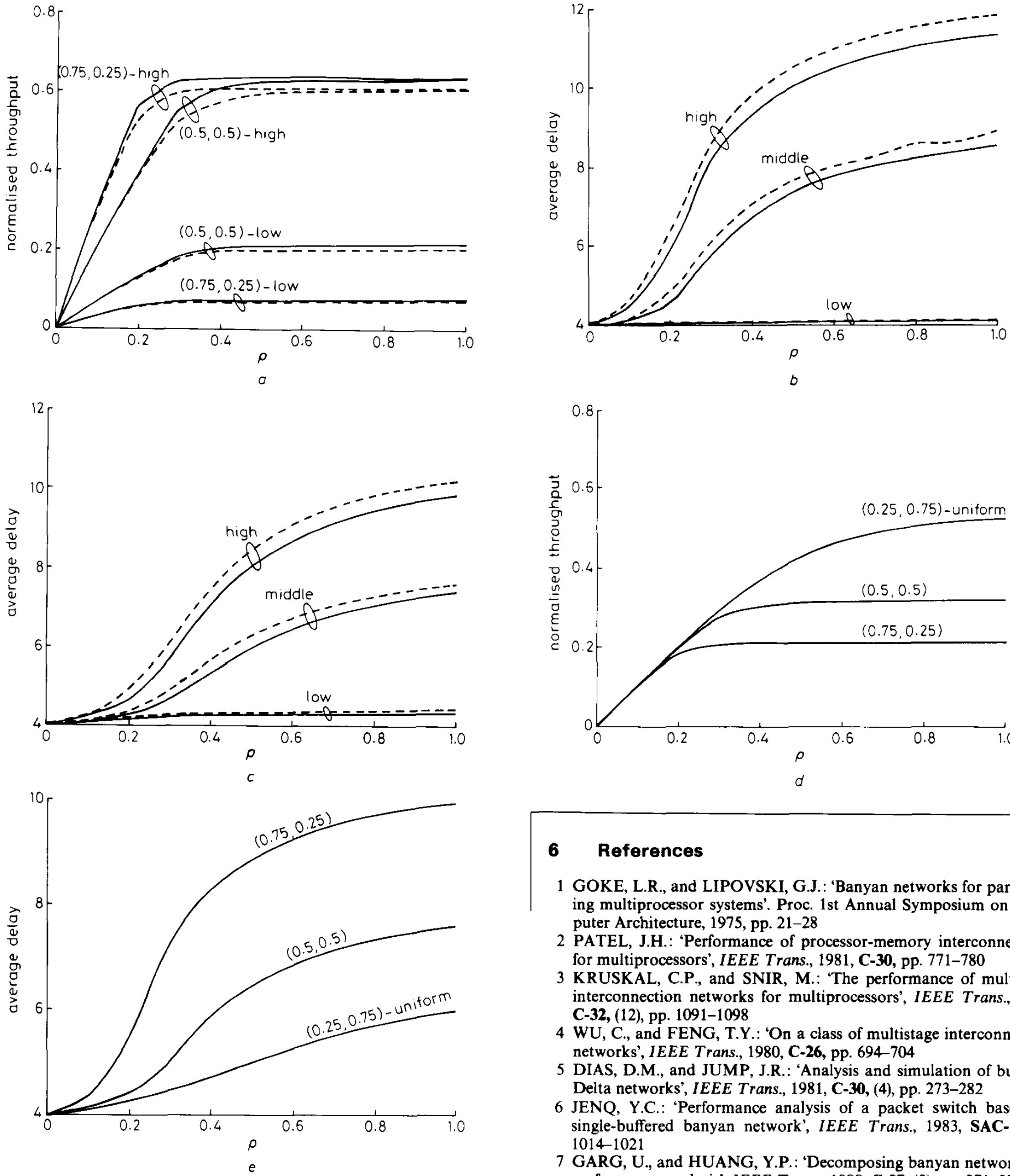

Fig. 5 *Normalised throughputs and average packet delays against input rate for nonuniform traffic matrices of form II, case 2*

a Normalised throughput against ρ
b Average packet delay against ρ with $(m_1, m_2) = (0.75, 0.25)$
c Average packet delay against ρ with $(m_1, m_2) = (0.5, 0.5)$
d Normalised throughput for the total network aginst ρ
e Average packet delay for the total network against ρ
——— analytic
- - - - simulation

6 References

1 GOKE, L.R., and LIPOVSKI, G.J.: 'Banyan networks for partitioning multiprocessor systems'. Proc. 1st Annual Symposium on Computer Architecture, 1975, pp. 21–28

2 PATEL, J.H.: 'Performance of processor-memory interconnections for multiprocessors', *IEEE Trans.*, 1981, **C-30**, pp. 771–780

3 KRUSKAL, C.P., and SNIR, M.: 'The performance of multistage interconnection networks for multiprocessors', *IEEE Trans.*, 1983, **C-32**, (12), pp. 1091–1098

4 WU, C., and FENG, T.Y.: 'On a class of multistage interconnection networks', *IEEE Trans.*, 1980, **C-26**, pp. 694–704

5 DIAS, D.M., and JUMP, J.R.: 'Analysis and simulation of buffered Delta networks', *IEEE Trans.*, 1981, **C-30**, (4), pp. 273–282

6 JENQ, Y.C.: 'Performance analysis of a packet switch based on single-buffered banyan network', *IEEE Trans.*, 1983, **SAC-1**, pp. 1014–1021

7 GARG, U., and HUANG, Y.P.: 'Decomposing banyan networks for performance analysis', *IEEE Trans.*, 1988, **C-37**, (3), pp. 371–376

8 KIM, H.S., and LEON-GARCIA, A.: 'Performance of buffered banyan networks under nonuniform traffic patterns'. IEEE Infocom, New Orleans, 1988, pp. 4A.4.1–4A.4.10

9 WU, L.T.: 'Mixing traffic in a buffered banyan network'. Proc. of 9th Data Communications Symposium, Whistler Mountain, British Columbia, Sept. 1985, pp. 134–139

10 KLEINROCK, L.: 'Queueing Systems, Vol. 1: Theory' (Wiley, New York, 1974)