ON JACKSON'S PRODUCT FORM WITH
'JUMP-OVER' BLOCKING

N.M. van Dijk

Research Memorandum 1988-4  Feb.'88
Abstract

It is shown that Jackson's product form is retained for queueing networks with capacity constraints under the assumption of a "jump-over" blocking protocol.

Key-words:

Queueing networks, product form, "jump-over" blocking.

* This paper has been written during a visit to The University of Adelaide, Australia
1. Introduction

Queueing networks with finite capacity constraints are generally known to fail Jackson's celebrated product form (cf. [4]). A major exception is the reversible routing case (cf. [4], [5], [8], [9], [10]). Another exception seems to arise when jobs encountering a saturated node jump over this node as if they were being served at an infinite speed. Despite the intuitive appeal of the latter statement (cf. Remark 2.2), however, a formal proof seems not be available in the literature.

This note, therefore, aims to provide a simple and straightforward proof for the “jumping-over” protocol to retain Jackson's product form. To highlight the essential feature of “jump-over” blocking, the presentation is restricted to the standard closed and exponential Jackson network formulation. The extension to open, non-exponential and more complex product form situations is immediate as will be briefly argued (cf. Remark 2.3).

2. Model and product form result

Consider a standard closed exponential Jackson network with $N$ nodes, a fixed number of jobs and routing probabilities $p_{ij}, j = 1, ..., N$ after a service completion at node $i$, $i = 1, ..., N$. Node $j$, $j = 1, ..., N$ has a service speed $f_j(n_j)$ when $n_j$ jobs are present and exponential services with parameter $\mu_j$. In addition, however, node $j$ has a capacity constraint $B_j$. That is, no more than $B_j$ jobs are allowed at the same time at node $j$, $j = 1, ..., N$.

A job which requests service at node $j$ while node $j$ is saturated (that is, $B_j$ jobs are already present) will instantly be routed to another node $\ell$ with probability $p_{j\ell}$ as if it is served by node $j$ at an infinite speed. This “jumping-over” may have to be repeated at node $\ell$ and will continue until the job gets accepted at some node.

Let $\vec{n} = (n_1, n_2, ..., n_N)$ denote the state vector with $n_j$ jobs at node $j$, $j = 1, ..., N$. The notation $\vec{n} + e_i - e_j$ is used to indicate the state which differs from $\vec{n}$ only in that there is one job more at node $i$ and one job less at node $j$. Without loss of generality assume that $p_{ii} = 0$, $i = 1, ..., N$. Further, let the routing matrix $(p_{ij})$ be irreducible so that there exists a unique probability distribution $\{\lambda_j, ..., \lambda_N\}$ which satisfies the traffic equations:

$$\lambda_j = \Sigma_i \lambda_i p_{ij} \quad (i = 1, ..., N). \quad (2.1)$$

**Theorem 2.1**

With $c$ a normalizing constant the steady state distribution is given by

$$\pi(\vec{n}) = c \prod_{j=1}^{N} (\lambda_j/\mu_j)^{n_j} \left[ \prod_{k=1}^{n_j} f_j(k) \right]^{-1} \quad (n_i \leq B_i, \quad i = 1, ..., N) \quad (2.2)$$
Proof

We need to verify the global balance equations for any state \( \bar{n} \) stating that the total rate out of state \( \bar{n} \) due to a change at any of the nodes \( j = 1, ..., N \) is equal to the rate into state \( \bar{n} \) due to a change at any of the nodes \( j = 1, ..., N \). For this however, it suffices to verify that for each node \( j = 1, ..., N \) separately

\[
\text{the rate out of state } \bar{n} \text{ due to a departure at node } j = \\
\text{the rate into state } \bar{n} \text{ due to an arrival at node } j. \tag{2.3}
\]

To this end, consider a fixed state \( \bar{n} \) and a fixed node \( j \). First note that (2.3) is trivially satisfied when \( n_j = 0 \). Assume that \( 1 \leq n_j \leq B_j \) and that nodes \( \ell_1, \ell_2, ..., \ell_t \neq j \) are saturated where \( t = 0 \) is included. (That is, \( n_i = B_i \) for \( i = \ell_1, ..., \ell_t \) while \( n_i < B_i \) for \( i \neq j, \ell_1, ..., \ell_t \)). Now note that the nodes can always be renumbered such that nodes \( \ell_1, ..., \ell_t \), and \( j \) will get numbers \( 1, ..., t \) and \( t + 1 \) respectively. Without loss of generality we may thus assume \((\ell_1, ..., \ell_t) = (1, ..., t)\) and \( j = t + 1 \).

Now let \( \tilde{R} \) be the probability matrix of an absorbing Markov chain with state space \{1, ..., t, *\} defined by

\[
\tilde{R}_{**} = 1 \\
\tilde{R}_{ij} = r_{ij} = p_{ij} \quad (i, j = 1, ..., t) \\
\tilde{R}_{is} = \Sigma_{\ell=t+1, ..., NPi} \quad (i = 1, ..., t)
\]

Hence, \( \tilde{R} \) performs as the original routing matrix as long as none of the nodes \( t + 1, ..., N \) is entered while the nodes \( t + 1, ..., N \) are lumped together as one absorbing state. Further, let the \( t \times t \)-matrix \( R \) be the restriction of \( \tilde{R} \) which is equivalent to that of \( P = (p_{ij}) \) to the states \{1, ..., t\}, denote its \( n \)-th power by \( R^n \) and define the \( t \times t \)-matrix :

\[
F = I + R + R^2 + ... + R^k + ...
\]

restricted to the states \{1, ..., t\} where \( I \) is the identity matrix. Then by standard Markov chain theory, for \( s, \ell \in \{1, ..., t\} \) the entry \( F_{s\ell} \) corresponds to the expected number of visits to node \( \ell \) before a job gets into service at any of the nodes \( t + 1, ..., N \) when starting to "jump-over" nodes at node \( s \), \( F_{s\ell} \tilde{R}_{s*} \) is the probability that it actually enters one of these nodes \( t + 1, ..., N \) by a transition out of node \( \ell \) and \( p_{ti} / \tilde{R}_{s*} \) is the fraction of this probability by which it actually enters node \( \ell \in \{t + 1, ..., N\} \). (Here it is noted that \( F \) is well-defined and finite due to the irreducibility assumption of the routing matrix). With the above notation we are now ready to write out the rates into and out of node \( j \) in a convenient form.

The rate out of state \( \bar{n} \) due to a departure at node \( j \) is

\[
\pi(\bar{n}) \mu_j f_j(n_j) \tag{2.6}
\]
By recalling that in state \( \bar{n} \) the nodes 1, ..., \( t \) are assumed to be saturated while the nodes \( t + 2, ..., N \) are not as we assumed \( j = t + 1 \), the rate into state \( \bar{n} \) due to an arrival at node \( j \) can be written as:

\[
\sum_{i=1}^{t+2} \pi(\bar{n} + e_i - e_j)\mu_i f_i(n_i + 1)p_{ij} + \sum_{i=t+2}^{N} \pi(\bar{n} + e_i - e_j)\mu_i f_i(n_i + 1)p_{i\bar{n}} F_{\bar{n}j} p_{\bar{n}j} \tag{2.7}
\]

By substituting

\[
\pi(\bar{n} + e_i - e_j) = \pi(\bar{n}) \lambda_i \mu_j f_j(n_j) / [\lambda_j \mu_i f_i(n_i + 1)], \tag{2.8}
\]

as according to (2.2), and introducing the notation for \( s = 1, ..., t \):

\[
d_s = \left[ \sum_{i=t+1}^{N} \lambda_i p_{is} \right] \tag{2.9}
\]

and for \( \ell = 1, ..., t \):

\[
\pi_\ell = \sum_{s=1}^{t} d_s F_{s\ell}, \tag{2.10}
\]

(2.7) can be rewritten as

\[
\pi(\bar{n}) \mu_j f_j(n_j) \lambda_j^{-1} \left\{ \sum_{i=t+2}^{N} \lambda_i p_{ij} + \sum_{\ell=1}^{t} \pi_\ell p_{\ell j} \right\} \tag{2.11}
\]

By virtue of the traffic equations (2.1) and recalling \( p_{jj} = 0 \), equality of (2.6) and (2.8) and thus relation (2.3) is verified if and only if

\[
\pi_\ell = \lambda_\ell \quad (\ell = 1, ..., t) \tag{2.12}
\]

To prove (2.12), write \( \tilde{\pi} = (\pi_1, ..., \pi_t) \) and \( \tilde{d} = (d_1, ..., d_t) \) and define

\[
\tilde{\pi}^n = \tilde{d}(I + R + ... + R^{n-1}) \tag{2.13}
\]

Then by (2.5), (2.10), (2.13) and recalling that \( F \) is finite, we obtain

\[
\tilde{\pi} = \lim_{k \to \infty} \tilde{\pi}^k = \lim_{k \to \infty} \tilde{\pi}^{k+1} = \lim_{k \to \infty} [\tilde{d} + \tilde{d}(I + R + ... + R^{k-1})R] = \tilde{d} + \pi R \tag{2.14}
\]

Consequently, by recalling (2.9) we have

\[
\pi_\ell = \sum_{i=t+1}^{N} \lambda_i p_{\ell i} + \sum_{s=1}^{t} \pi_s p_{s\ell} \quad (\ell = 1, ..., t) \tag{2.15}
\]


As these equations coincide with the traffic equations (2.1) which have a unique probability solution \( \{\lambda_1, ..., \lambda_N\} \), the equalities (2.12) are guaranteed. The proof is hereby completed.

**Remark 2.2**

A standard intuitive arguing for justifying the product form (2.1) is the following. Instead of assuming a capacity constraint \( B_j \) at node \( j \) let \( f_j(n_j) = M \) for \( n_j > N_j \) where \( M \) is assumed to be a very large but finite number. Jackson's standard product form without capacity constraint then applies. By letting \( M \to \infty \) that product form expression will then converge to the one coinciding with (2.2) and \( \pi(n) = 0 \) when \( n_j > B_j \) for some \( j \). Physically moreover, the system tends to behave as if jobs are jumping over saturated nodes within an almost negligible time. To formally justify this arguing, however, one has to prove that the corresponding queueing processes converge in a weak sense upon appropriate D-sample path spaces (cf. [3],[12]). This will involve various technical details. In particular, a difficulty will arise as D-spaces prohibit jumps of a fixed size to accumulate at one point as will be the case when more than one node gets saturated.

**Remark 2.3**

The essence of the proof comes down to the standard notion (2.3) of partial balance per node and the recurrence relation (2.14) with an appropriate initial distribution \( \bar{d} \) of the form (2.9) for the visit ratios \( \pi_\ell \) defined by (2.5) and (2.10) at all nodes \( \ell \) at which "jumping-over" occurs. The result of a product form with "jump-over" blocking is therefore extendable to any classical product form network without capacity constraints for which (2.3), or the more specified analogue when positions at a node are distinguished, holds. These include for instance open networks with Poisson arrival streams at any node and networks with symmetric nodes when positions at those nodes are regarded. (cf. [1],[2],[5],[8]). In particular, in the case of balanced positions such as in symmetric nodes, also insensitivity results to service distributional forms can so be obtained. (cf. [2],[6],[8],[11]).

**Acknowledgement**

I like to thank Bill Henderson and Peter Taylor of The University of Adelaide for useful comments. I am grateful to the Department of Applied Mathematics of this University for arranging a visit to this department during which this research has been developed. The research has been supported by an award of the Australian Research Grant Scheme.
References


