ET 56

05348

1988

SERIE RESEARCH MEMORANDA

ERROR BOUNDS FOR COMPARING OPEN AND CLOSED QUEUEING NETWORKS WITH AN APPLICATION TO PERFORMABILITY ANALYSIS

Nico M. van Dijk

Research Memorandum 1988-56 Dec.'88

VRIJE UNIVERSITEIT
FACULTEIT DER ECONOMISCHE WETENSCHAPPEN
EN ECONOMETRIE
AMSTERDAM

ERROR BOUNDS FOR COMPARING
OPEN AND CLOSED QUEUEING NETWORKS
WITH AN APPLICATION TO PERFORMABILITY ANALYSIS

Nico M. van Dijk
Free University, Amsterdam, The Netherlands

Abstract A condition is provided from which an error bound and rate of convergence can be concluded when comparing the performance of an open and closed analog of a possibly non-product form queueing network. The result is applied to a Jackson network with breakdowns, such as arising in performability analysis. An explicit simple error bound is obtained of order 1/M with M the size of finite source input.

Key-words Error bound * Open-closed queueing network * performability.

1 Introduction

Ever since Erlang's and Engset's classical results for open and closed single service facility models, the issue of whether an open or closed modeling approach is to be used for a given service network such as a computer or communication system has been actual (cf. [30]). At first glance an open modeling seems more natural as jobs that are usually generated exteriorly travel through a network in order to receive a certain amount of services after which it departs the system. Systems that are closed by nature such as Engset's classical machine interference system or relatedly a maintenance model for a fixed number of devices, seem much more special. However, typical present day service networks such as a central processor system with a fixed number of input sources or a manufacturing sytem with a transport device are to be regarded as closed, since upon completion of a job's services a new job is instantly inserted (cf. [15], [30], [31]). Also, communication or broadcasting systems, like ALOHA or CSMA networks (cf. [10], [15]), can be analyzed as a closed two station service network with jobs representing transmitters that are idle (not transmitting) while at station 1 and busy (transmitting) while at station 2 (cf. [24]).

Beyond the natural modeling issue, however, there are other reasons for preferencing an open or closed modeling. For open networks that exhibit a product form, the "stationary independence" of the stations enables an analysis or computational procedure per station. Closed product form expressions, in contrast, can be computationally unattractable as the partitioning constant is to be calculated. Various techniques, such as mean value analysis (cf. [12]), asymptotic analysis (cf. [9]), or statistical mechanics (cf. [11]) can thus be involved. On the other hand, analytic results for closed form systems might lead to limiting results for open systems (cf. Barbour [1]). For non-product systems, open modeling can be handy for applying simple results as Little's formula, but generally closed networks are then more appropriate as computational procedures usually require a finite or truncated state space.

This paper therefore is concerned with estimating how much a performance measure of interest like a throughput or system utilization differs for open and closed versions of a particular service network. No conditions such as a product form structure will be imposed, so that the analysis applies to networks with phenomena as blocking, synchronous servicing or breakdowns. A general condition will be provided from which an error bound or rate of convergence as the input source tends to infinity can be concluded (section 2).

This result can be seen as a combined perturbation and state space truncation result. Though state space truncation is a common feature for practical computations, explicit error bounds or rates of convergence are hardly available in the literature. Convergence results of state space truncations are extensively studied with most notably references as [6] or [16]. The latter reference also reports simple error bounds but these are just robust bounds and do not secure an order of accuracy. Similar statements hold for the somewhat related results of [2] and [14]. In the particular context of approximating open models by closed, convergence results have been established by [1] and [30]. Among various other results such as on monotonicity, throughput bounds and insensitivity, the latter reference also includes some error bounds for approximating some special Erlang models.

The condition in this paper is straightforward and directly yields error bounds for arbitrary networks. The verification of this condition basically comes down to providing estimates for so-called bias terms of reward structures. This in turn can frequently be transformed to monotonicity results, for which various proof-techniques that have been developed over the last couple of years (cf. [17], [19], [21], [23], [27], [29]), might be applicable.

The results will be illustrated for a particular network of practical interest: a queueing network subject to breakdowns. Breakdowns are a main problem of concern in developing and evaluating computer, communication or manufacturing systems. Particularly, performance evaluation of computer systems with disturbances, interruptions or breakdowns currently receives considerable attention under the name of "performability" (cf. [8], [18]). As breakdown systems do not generally exhibit a product form expression, numerical or approximate computations are thus involved so that error bounds for closed modeling of open systems become most relevant. To this end, the necessary condition will be verified by establishing monotonicity results and estimates for the bias terms of the given reward structure. An explicit error bound for the throughput of the given breakdown network will so be derived (section 3).

For expository convenience and to highlight the essential features, the presentation will be restricted to exponential queueing networks with non-distinguishable jobs and in which only one job can change at a time. Extensions to multi-class, non-exponential and batch service networks however will briefly be argued as being essentially similar (section 4).

2 General model and main comparison result

Consider an arbitrary single class exponential queueing network such as illustrated below with either a Poisson arrival input with parameter λ (hereafter called the open case) or a finite source input with M sources and exponential holding times with parameter $\gamma = \lambda/M$ (hereafter called the closed case).

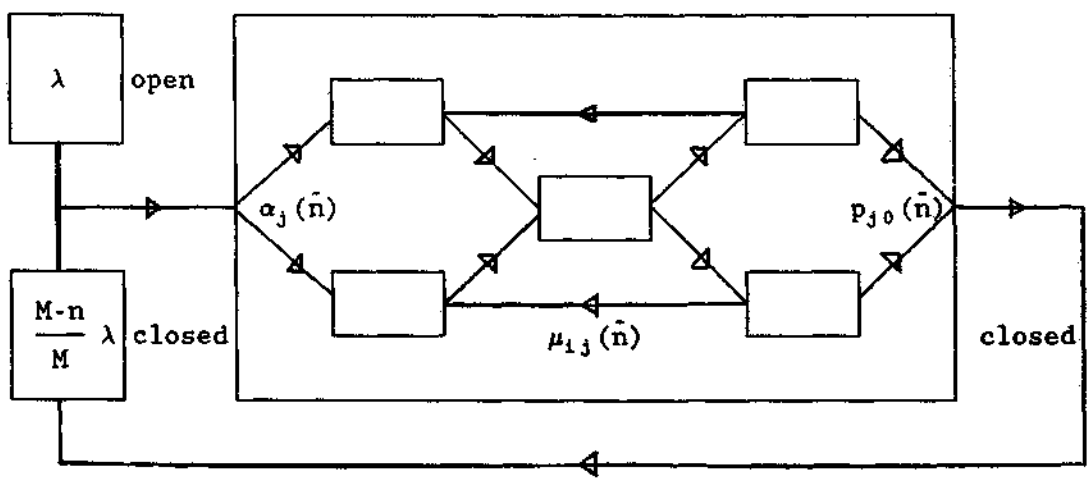

FIGURE 1

More precisely, let the network have N service stations and denote by $\bar{n}$ =

$(n_1, n_2, \ldots, n_N)$ that n_i jobs are present at station i, $i = 1,\ldots,N$. Further, for $i = 1,\ldots,N$, let $\bar{n}+e_i$ and $\bar{n}-e_i$ be the state equal to $\bar{n}$ with one job more respectively less, for $n_i>0$, at station i and introduce $\bar{n}+e_0=\bar{n}$. Consequently, the state $\bar{n}-e_i+e_j$ is the state equal to $\bar{n}$ with one job moved from station i to j. Then the transition rates for a change from a state $\bar{n}$ into a state $\bar{n}-e_i+e_j$ are given below, where $j = 0$ corresponds to a departure from the system at node i while $i = 0$ corresponds to an arrival at the system at station j. Herein as well as throughout we use $n = n_1 + \ldots + n_N$.

$$q(\bar{n},\bar{n}-e_i) = \mu_i(\bar{n}) \qquad (i = 1,..,N)$$

$$q(\bar{n},\bar{n}-e_i+e_j) = \mu_{ij}(\bar{n}) \qquad (i,j = 1,\ldots,N)$$

$$q(\bar{n},\bar{n}+e_j) = \begin{cases} \lambda\ \alpha_j(\bar{n}) & \text{for the open case} \\ [M-n][\lambda/M]\alpha_j(\bar{n}) & \text{for the closed case.} \end{cases} \tag{2.1}$$

In words that is, a job leaves station i and leaves the system or routes to station j, with state dependent intensities $\mu_i(\bar{n})$ and $\mu_{ij}(\bar{n})$ respectively, while upon arrival of a job at the system it enters station j with state dependent probability $\alpha_j(\bar{n})$. Here it is noted that we allow $j = i$ or $j = 0$, so that blocking phenomena can be modeled. For example, with probability $\alpha_0(\bar{n})$ an arrival is rejected and lost.

Without loss of generality assume that the open and closed model are irreducible at state spaces S and $\hat{S} \subset S$ with stationary distributions $\pi(.)$ and $\hat{\pi}(\bar{n})$ respectively. Throughout we will implicitly assume that the open and closed model are considered restricted to their irreducible sets. Now, let $r(\bar{n})$ be some reward rate whenever the system is in state $\bar{n}$. Then,

$$g = \Sigma_{\bar{n}}\ \pi(\bar{n})\ r(n) \tag{2.2}$$

and

$$\hat{g} = \Sigma_{\bar{n}}\ \hat{\pi}(\bar{n})\ r(\bar{n}) \tag{2.3}$$

represent the performance measure under the reward rate r for the open and closed case respectively. Here, without restriction of generality, we assume these performance measures to be well-defined. Without knowing or using the stationary probabilities $\pi(.)$ and $\bar{\pi}(.)$, we now wish to derive an error bound for the difference $|\bar{g} - g|$. To this end, we assume that for some finite number Q:

$$Q \geq \sup_{\bar{n}} \; [\lambda + \Sigma_i \; \mu_i(\bar{n}) + \Sigma_{i,j} \; \mu_{ij}(\bar{n})] \tag{2.4}$$

so that we can apply the standard uniformization technique (cf. [20], p.110) to transform the continuous-time models in discrete-time models. More precisely, we define one-step transition probabilities $p(\bar{n},\bar{n}-e_i+e_j)$ and $\bar{p}(\bar{n},\bar{n}-e_i+e_j)$ for the open and closed case by:

$$\bar{p}(\bar{n},\bar{n}-e_i) = p(\bar{n},\bar{n}-e_i) = \mu_i(\bar{n})/Q$$

$$\bar{p}(\bar{n},\bar{n}-e_i+e_j) = p(\bar{n},\bar{n}-e_i+e_j) = \mu_{ij}(\bar{n})/Q \qquad (j\neq i,\ j,i\neq 0)$$

but

$$\begin{cases} p(\bar{n},\bar{n}+e_j) = \lambda\alpha_j(\bar{n})/Q & (j = 0,1,\ldots,N) \\ \bar{p}(\bar{n},\bar{n}+e_j) = [M-n][\lambda/M]\alpha_j(\bar{n})/Q & (j = 0,1.\ldots,N) \end{cases}$$

and

$$\begin{cases} p(\bar{n},\bar{n}) = 1+\mu_{ii}(\bar{n})/Q - [\Sigma_i \mu_i(\bar{n})+\Sigma_{i,j} \; \mu_{ij}(\bar{n})+\lambda]/Q \\ \bar{p}(\bar{n},\bar{n}) = 1+\mu_{ii}(\bar{n})/Q - [\Sigma_i \mu_i(\bar{n})+\Sigma_{i,j} \; \mu_{ij}(\bar{n})+\lambda[M-n]M^{-1}]/Q \end{cases} \tag{2.5}$$

From now on, we always use an upper bar "-" symbol to indicate the closed case and the symbol "(-)" to indicate both the open and closed case. Now in order to compare the values g and $\bar{g}$, let the operators $\overset{(-)}{T}$ and $\{\overset{(-)}{T}_t \mid t=0,1,2,\ldots\}$ for arbitrary functions f be given by

$$\begin{aligned} \overset{(-)}{T} f(\bar{n}) &= \Sigma^N_{i,j=0} \overset{(-)}{p}(\bar{n},\bar{n}-e_i+e_j)\; f(\bar{n}-e_i+e_j) \\ \overset{(-)}{T}_{t+1} &= \overset{(-)}{T}\; \overset{(-)}{T}_t \qquad (t = 0,1,2,\ldots) \\ \overset{(-)}{T}_0 &= I \end{aligned} \tag{2.6}$$

and define the reward functions $\{V_t \mid t=0,1,2,\ldots\}$ at S and $\{\bar{V}_t \mid t=0,1,2,\ldots\}$ at $\bar{S}$ by

$$\overset{(-)}{V}_N = \Sigma_{t=0}^{N-1} \overset{(-)}{T}_t r/Q \tag{2.7}$$

Then by virtue of the uniformization technique (cf. [20], p. 110) and the irreducibility assumption of S and $\bar{S}$, we obtain by standard Markov reward theory (cf. [13])

$$\overset{(-)}{g} = \lim_{N\to\infty} \frac{Q}{N} \overset{(-)}{V}_N (\bar{n}) \tag{2.8}$$

for arbitrary $\bar{n} \in \bar{S} \subset S$. The following key-result can now be proven. It provides a pair of conditions that guarantees an error bound for $|\bar{g} - g|$. These conditions will be argued and illustrated later on.

Theorem 2.1 Suppose that for some state $\bar{\ell} \in \bar{S}$, some nonnegative function μ, some constants B and C and all $t\geq 0$, $\bar{n} \in S$:

$$\bar{T}_t \ \mu(\bar{\ell}) \leq B \tag{2.9}$$

$$|n \ \Sigma_j \ \alpha_j(\bar{n})[V_t(\bar{n}+e_j) - V_t(\bar{n})]| \leq \mu(\bar{n})C. \tag{2.10}$$

Then

$$|(\bar{V}_N - V_N)(\bar{\ell})| \leq \lambda N \ B \ C \ / \ [MQ] \tag{2.11}$$

$$|\bar{g} - g| \leq \lambda \ B \ C \ / \ M. \tag{2.12}$$

Proof Clearly, (2.12) follows from (2.8) and (2.11). By virtue of (2.6) and (2.7), we have for all $t \geq 0$:

$$\overset{(-)}{V}_{t+1} = r/Q + \overset{(-)}{T} \overset{(-)}{V}_t .$$

Since $\bar{T}$ remains restricted to $\bar{S} \subset S$, we can thus write:

$$(\bar{V}_N - V_N)(\bar{\ell}) = (\bar{T}\,\bar{V}_{N-1} - TV_{N-1})(\bar{\ell})$$

$$= (\bar{T} - T)\ V_{N-1}(\bar{\ell}) + \bar{T}(\bar{V}_{N-1} - V_{N-1})(\bar{\ell})$$

$$= \Sigma_{t=0}^{N-1}\ \bar{T}_t\ [(\bar{T}-T)V_{N-t-1}](\bar{\ell}) + \bar{T}_N(\bar{V}_0 - V_0)(\bar{\ell}) \tag{2.13}$$

where the last term follows by iteration. From the uniformization construction (2.5) and the definition (2.6), however, we readily obtain for any $\bar{n} \in \bar{S}$ and $s \leq N$:

$$(\bar{T}-T)V_s(\bar{n}) = n\ \lambda[MQ]^{-1}\ \Sigma_j\ \alpha_j(\bar{n})[V_s(\bar{n}+e_j)-V_s(\bar{n})] \tag{2.14}$$

By substituting (2.14) in (2.13), recalling that $\overset{(-)}{V_0}(.)=0$, taking absolute values and noting that expectation operators $\bar{T}_t$ are monotone operators, we thus obtain from (2.10):

$$|(\bar{V}_N - V_N)(\bar{\ell})| \leq \lambda C[MQ]^{-1}\ \Sigma_{t=0}^{N-1}\ \bar{T}_t\ \mu(\bar{\ell}).$$

Condition (2.9) completes the proof. □

Remark 2.2 (μ-function) Typically one can think of the μ-function to be a polynomial in n. For example, with the terms $V_t(\bar{n}+e_j)-V_t(\bar{n})$ uniformly bounded in all t, n and j, we can take $\mu(\bar{n}) = (1+n)$. Condition (2.9) then simply requires the expected system size to remain bounded over time. This is most natural in practice.

Remark 2.3 (Bounded bias-terms) From Markov reward theory it is standardly known that the so-called bias-terms $|V_t(\bar{n}+e_j)-V_t(\bar{n})|$ can be estimated from above for any given $\bar{n}$ and j and uniformly in t as based upon mean first-passage time results (cf. [7], [13], [26]) provided the reward rate r is bounded. Particularly, with the state space S being finite, an estimate C can then be provided uniformly in t, $\bar{n}$ and j depending on mean first passage times. For the unbounded case similar estimates can be provided under appropriate conditions (cf. [26]). For multi-dimensional applications, however, mean first passage times are extremely hard to obtain (cf. [7]) and even not robust bounds seem to be generally available. For the application in section 3, therefore, we will estimate these bias terms in a direct manner based on monotonicity results.

Remark 2.4 (Unbounded reward rate) Note that no assumption has been made as to the reward rate r. Particularly, we can think of unbounded functions such as

$$r(\bar{n}) = n$$

for evaluating the mean system size. Other possibilities of interest are:

$$r(\bar{n}) = 1_{\{B\}}(\bar{n}),$$

where $1_{\{B\}}(.)$ denotes an indicator function of a subset B, so as to calculate the steady state probability of a set B or

$$r(\bar{n}) = \Sigma_i \ \mu_i(\bar{n})$$

by which we compute the output rate and thus the throughput of the system.

Remark 2.5 (Mixed open and closed networks) Note that no conditions are imposed other than the irreducibility assumption and the uniformly bounded transition rates. Particularly, we may have mixed networks with a fixed number of jobs traveling within one subset of stations that cannot leave the system and jobs that enter from outside which travel through another subset of stations after which they leave the system. As the transition rates such as the service and/or arrival rates, however, depend on the total system state, these disjoint station clusters cannot be dealt with in isolation and are to be regarded as one system. This can be practical for modeling breakdowns, priority jobs or multiple job-types, as will be illustrated by the breakdown application in the next section.

3 Application: A Jackson queueing network with breakdowns

In order to illustrate how the conditions (2.9) and (2.10) can be verified for concrete networks as well as to provide a result of practical interest, this section is concerned with applying theorem 2.1 to queueing networks with breakdowns. Such systems do not exhibit the celebrated product form expression but are of interest such as for evaluating the reliability or efficiency of a computer system which from time to time can be down. For expository convenience we restrict ourselves to the "simple" case of a standard Jackson network with one type of breakdown. Similar results, however, can be expected along the same lines for other more complex networks and multiple breakdowns.

3.1 Model

Consider a Jackson queueing network with N service stations which is subject to breakdowns, independently of whether the system works or not, at an exponential rate with parameter ν_1. A breakdown renders the total system inoperative for an exponential period with parameter ν_0. The service speed of station i when n_i customers are present is given by $\mu_i(n_i)$ which is assumed to be non-decreasing in n_i. Upon service completion at station i a job routes to another station j with probability p_{ij} or leaves the system with probability $p_{i0} = [1-\Sigma_{j\le N} p_{ij}]$, $i=1,\dots,N$. Arrivals at the system are generated either by a Poisson process with parameter λ (the open case) or by a finite source input with M sources and exponential source parameter $[\lambda/M]$ (the closed case). Upon arrival a job is assigned station j with probability $p_{0j} = \alpha_j$. Without restriction of generality it is assumed that the routing matrix $(p_{ij})_{i,j=0}^{N}$ is irreducible. The above description is visualized for an example in figure 2 below.

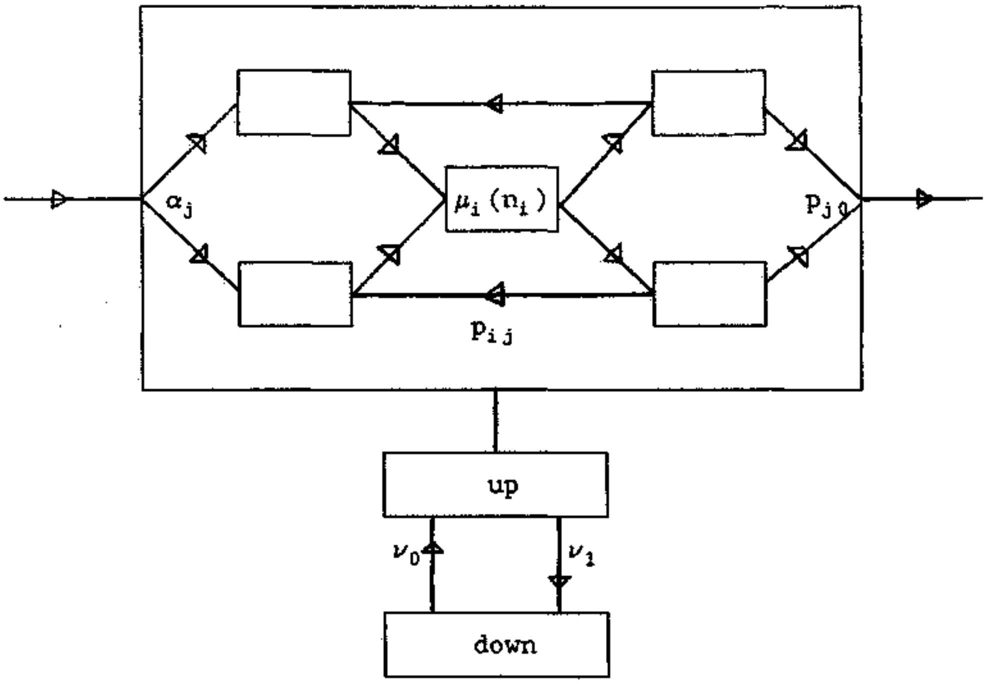

FIGURE 2

The breakdown description above is known in the literature as the "independent breakdown" case in contrast with the "active breakdown" case in which the system can go down only when it is working (cf. [5], p. 101). Either case appears to be as equally untractable analytically. For the simple case of a single multi-server facility without arrival blocking and either type of breakdown, closed form expressions have been obtained only for the generating function of the queue length (cf. [5], p. 103). For the general case as considered above no closed form expression has been reported and certainly a product form will not hold (cf. [4]). The throughput of the open case is equal to λ assuming that the system is stable. For the practically more interesting closed case, however, there exists no counter part. To this end, we will apply the results of section 2.

3.2 Reformulation

First we need to reformulate our system in the setting of section 2. Therefore, as per the figure above and also referring to remark 2.4, consider two special stations N+1 and N+2 and one special job which is always present alternately visiting these two stations with respective exponential holding times with parameters ν_1 and ν_0. When the job is at station N+1 the system is up while when at station N+2 the system is down. The state of the system is denoted by $(\bar{n},\theta)$ representing that n_i jobs are at station i, $i=1,\ldots,N$ while $\theta = n_{N+1}$ indicates that the system is down for $\theta = 0$ and up for $\theta = 1$. The corresponding transition rates for this system with N+2 stations as per (2.1) are parametrized by

$$\begin{cases} \mu_{ij}(\bar{n},\theta) = 1_{\{\theta=1\}}\mu_i(n_i)p_{ij} & (j\le N)\ (i\le N) \\ \mu_i(\bar{n},\theta) = 1_{\{\theta=1\}}\mu_i(n_i)p_{i0} & (i\le N) \\ p_{N+1,N+2} = p_{N+2,N+1} = 1 & \\ \mu_{N+1}(\bar{n},\theta) = 1_{\{\theta=1\}}\nu_1 & \\ \mu_{N+2}(\bar{n},\theta) = 1_{\{\theta=0\}}\nu_0 & \\ \alpha_j(\bar{n},\theta) = \alpha_j & \end{cases} \tag{3.1}$$

where $1_{\{A\}}=1$ if event A is satisfied and 0 otherwise. The breakdown system as described above is hereby transformed to a stochastically equivalent system in the setting of section 2 with N+2 stations and Poisson arrivals with parameter λ (open case) or a finite source input with M sources and exponential holding rate λ/M per source (closed case). As performance measure of interest we wish to evaluate the throughput of the system. To this end, we set

$$r(\bar{n},\theta) = 1_{\{\theta=1\}}\Sigma_i\mu_i(n_i)p_{i0} \tag{3.2}$$

Further, we assume that for some finite Q:

$$Q \ge \sup_{\bar{n}}\ [\lambda + \nu_0 + \nu_1 + \Sigma_i\ \mu_i(n_i)]. \tag{3.3}$$

The transition probabilities $p(.,.)$ for the open case and $\bar{p}(.,.)$ for the closed case as according to the uniformization (2.5) and the parametrization (3.1) are now given by

$$^{(\bar{p})}([\bar{n},1],\ [\bar{n},0] = \nu_1/Q$$

$$^{(\bar{p})}([\bar{n},0],\ [\bar{n},1] = \nu_0/Q$$

$$^{(\bar{p})}([\bar{n},1],\ [\bar{n}-e_i,1]) = \mu_i(n_i)p_{i0}/Q$$

$$^{(\bar{p})}([\bar{n},1],\ [\bar{n}-e_i+e_j,1]) = \mu_i(n_i)p_{ij}/Q$$

but

$$\begin{cases} p([\bar{n},\theta],\ [\bar{n}+e_j,\theta]) = \lambda\alpha_j/Q \\ \bar{p}([\bar{n},\theta],\ [\bar{n}+e_j,\theta]) = [M\text{-}n][\lambda/M]\alpha_j/Q \end{cases}$$

and

$$\begin{cases} p([\bar{n},\theta],\ [\bar{n},\theta]) = 1 - \{\Sigma_i\mu_i(n_i) - \nu_\theta - \lambda\}/Q \\ \bar{p}([\bar{n},\theta],\ [\bar{n},\theta]) = 1 - \{\Sigma_i\mu_i(n_i) - \nu_\theta - [M\text{-}n][\lambda/M]\}/Q \end{cases} \qquad (3.4)$$

The operator $\bar{T}$ for the closed system and the functions $\{V_t \mid t=0,1,2,\ldots\}$ for the open system are correspondingly defined by (2.6) and (2.7) with r given by (3.2).

3.3 Error bound

Now we need to verify the conditions of theorem 2.1. To this end, the following lemma, which is the most essential step, is concerned with estimating the bias terms in (2.10). First, we introduce the notation

$$\Delta_j V_t(\bar{n},\theta) = V_t(\bar{n}+e_j,\theta) - V_t(\bar{n},\theta) \qquad (j=1,\ldots,N),(t\geq 0) \qquad (3.5)$$

Lemma 3.1 For all $t \geq 0$ and $[\bar{n},\theta]$, we have

$$0 \leq \Delta_j V_t(\bar{n},\theta) \leq 1 \qquad (j=1,\ldots,N) \qquad (3.6)$$

Proof This will be given by induction to t. For $t = 0$, (3.6) trivially holds as $V_0(.) = 0$. Suppose that (3.6) holds for $t \le m$. Then by (2.7), (3.4) and (3.5):

$$\Delta_i V_{m+1}(\bar{n},\theta) =$$

$$([1_{\{\theta=1\}} \Sigma_j \mu_j(n_j) p_{j0} +$$

$$1_{\{\theta=1\}} [\mu_i(n_i+1) - \mu_i(n_i)] p_{i0} +$$

$$\lambda\ \Sigma_j \alpha_j V_m(\bar{n}+e_i+e_j, \theta) +$$

$$\nu_0\ 1_{\{\theta=0\}} V_m(\bar{n}+e_i, 1) + \nu_1\ 1_{\{\theta=1\}} V_m(\bar{n}+e_\ell, 0) +$$

$$1_{\{\theta=1\}}\ \Sigma_{j=1}^N\ \mu_j(n_j)\ \Sigma_{\ell=0}^N\ p_{j\ell}\ V_m(\bar{n}+e_i-e_j+e_\ell, 1) +$$

$$1_{\{\theta=1\}}\ [\mu_i(n_i+1) - \mu_i(n_i)]\ \Sigma_{\ell=0}^N\ p_{i\ell}\ V_m(\bar{n}+e_\ell, 1)]/Q +$$

$$[1 - (\lambda+\nu_\theta+1_{\{\theta=1\}} \Sigma_j \mu_j(n_j) + 1_{\{\theta=1\}} [\mu_i(n_i+1) - \mu_i(n_i)])/Q]\ V_m(\bar{n}+e_i, \theta))$$

$$- ([1_{\{\theta=1\}}\ \Sigma_j \mu_j(n_j) p_{j0} +$$

$$\lambda\ \Sigma_j \alpha_j\ V_m(\bar{n} + e_j, \theta) +$$

$$\nu_0\ 1_{\{\theta=0\}} V_m(\bar{n}, 1) + \nu_1 1_{\{\theta=1\}} V_m(\bar{n}, 0) +$$

$$1_{\{\theta=1\}}\ \Sigma_{\ell=1}^N \mu_j(n_j)\ \Sigma_{\ell=0}^N\ p_{j\ell}\ V_m(\bar{n}-e_j+e_\ell, 1) +$$

$$1_{\{\theta=1\}}\ [\mu_i(n_i+1) - \mu_i(n_i)]\ \Sigma_{\ell=0}^N\ p_{i\ell}\ V_m(\bar{n}, 1)]/Q +$$

$$[1 - (\lambda+\nu_\theta+1_{\{\theta=1\}}\ \Sigma_j \mu_j(n_j) + 1_{\{\theta=1\}} [\mu_i(n_i+1) - \mu_i(n_i)])/Q]\ V_m(\bar{n}, \theta))$$

$$= \{1_{\{\theta=1\}}\ [\mu_i(n_i+1)-\mu_i(n_i)]p_{i0}\ +$$

$$\lambda\ \Sigma_j \alpha_j \Delta_i V_m(\bar{n}+e_j, \theta)\ +$$

$$\nu_0\ 1_{\{\theta=0\}}\ \Delta_i V_m(\bar{n},1)\ +\ \nu_1\ 1_{\{\theta=1\}}\ \Delta_i V_m(\bar{n},1)\ +$$

$$1_{\{\theta=1\}}\ \Sigma_{j=1}^N\ \mu_j(n_j)\ \Sigma_{\ell=0}^N\ p_{j\ell}\ \Delta_i V_m(\bar{n}-e_j+e_\ell,1)\ +$$

$$1_{\{\theta=1\}}\ [\mu_i(n_i+1)-\mu_i(n_i)]\ \Sigma_{\ell=1}^N\ p_{i\ell}\ \Delta_\ell V_m(\bar{n},1)\ +$$

$$1_{\{\theta=1\}}\ [\mu_i(n_i+1)-\mu_i(n_i)]p_{i0}\ [V_m(\bar{n},1)-V_m(\bar{n},1)]\}/Q\ +$$

$$[1 - \{\lambda+\nu_\theta+1_{\{\theta=1\}}\ \Sigma_j \mu_j(n_j)+1_{\{\theta=1\}}\ [\mu_i(n_i+1)-\mu_i(n_i)]\}/Q]\ \Delta_i V_m(\bar{n},\theta). \tag{3.7}$$

Now recall that $\mu_i(n_i)$ is assumed to be non-decreasing and note that the one but last term of the latter expression is equal to 0. Then by substituting the induction hypothesis (3.6) for $t=m$, the lower estimate 0 of (3.6) for $t=m+1$ follows immediately. The upper estimate 1 of (3.6) for $t=m+1$ is guaranteed by letting the first term replace this one but one last term from which it is clear that all terms sum to 1. □

As will be shown shortly, lemma 3.1 will imply that condition (2.10) can be verified with a function μ which is linear in n. Therefore, let us investigate condition (2.9) with

$$\mu(\bar{n},\theta) = n \tag{3.8}$$

To this end, we first consider a modified version of the given breakdown system in the open case by letting arrivals be rejected and lost whenever the system is down ($\theta=0$). Then it can be concluded from literature (cf. [4]) or verified by direct substitution in the equilibrium balance equations, that this modified system has the product form:

$$\pi_M(\bar{n},\theta) = c[1/\nu_\theta]\ \prod_i\ [\lambda_i]^{n_i} / [\prod_{k=1}^{n_i}\ \mu_i(k)] \tag{3.9}$$

where c is a normalizing constant and $\{\lambda_i\}$ the unique solution of

$$\lambda_j = \lambda\alpha_j + \Sigma_i \lambda_i p_{ij} \quad (j \le N) \tag{3.10}$$

As a consequence, the corresponding mean system load L_M of this modified system is given by

$$L_M = [\nu_0\nu_1/(\nu_0+\nu_1)]L \tag{3.11}$$

where L is the mean system load of the original system without breakdowns, that is as corresponding to the standard Jackson product form (3.9) without $[1/\nu_\theta]$ (or with $\nu_\theta = \infty$). As this system load L can easily be computed, the following lemma is useful. Let $\bar{0} = (0,\ldots,0)$, then

Lemma 3.2 For all $t \ge 0$:

$$\bar{T}_t\ \mu(\bar{0},1) \le L_M \tag{3.12}$$

Proof By virtue of the uniformization technique (cf. [20], p.110) one can directly conclude that the mean system load under the construction (3.4) with $\bar{p}$ is equal to that of the original closed continuous-time model, say denoted by $\bar{L}_o$. Clearly, as the state dependent arrival rates of the closed system are always less or equal than λ as in the open case, one can show by standard monotonicity proof techniques such as in [17], [19] or [29] that

$$\bar{L}_o \le L_o,$$

where L_o denotes the mean system loead of the original system in the open case. Finally, similarly to [21] and the proof of lemma 3.1, one can also show that

$$L_o \le L_M,$$

by which the proof is completed. □

We are now able to apply theorem 2.1.

Theorem 3.3 (Throughput error bound) With $\bar{g}$ and λ the throughput of the closed and open system respectively, we have

$$|\bar{g} - \lambda| \leq \lambda\ L_M\ /\ M \tag{3.13}$$

Proof This follows directly from lemma 3.1 and lemma 3.2 by applying theorem 2.1 with $\bar{\ell}=(\bar{0},1)$, $\mu(\bar{n},\theta)=n$, $\alpha_j(\bar{n},\theta)=\alpha_j$ for $j=1,,,.N$ and $\alpha_{N+1}(\bar{n},\theta) = \alpha_{N+2}(\bar{n},\theta) = 0$. Then (2.10) is guaranteed with C=1 by:

$$|\Sigma_{j=1}^{N}\ \alpha_j\ [V_t(\bar{n}+e_j,\theta) - V_t(\bar{n},\theta)]\ | \leq 1 \tag{3.14}$$

Example 3.4 (Single server stations) As a particular example, with each station $i \leq N$ a single server station with exponential service rate μ_i, (3.9)-(3.11) and (3.13) yield

$$|\bar{g} - \lambda| \leq \lambda\ M^{-1}[\nu_0\nu_1/(\nu_0+\nu_1)][\prod_{i=1}^{N} \rho_i/(1-\rho_i)] \quad (\rho_i=\lambda/\mu_i) \tag{3.15}$$

4 Some extensions

4.1 Multiclass networks The results can be extended to multiclass queueing networks if we model a different Poisson arrival stream or a finite source input for each separate job-class. Basically, condition (2.10) then needs to be replaced by

$$|n^r\ \Sigma_j\alpha_j^r(\bar{n})\ [V_t(\bar{n}+e_j^r) - V_t(\bar{n})]\ | \leq \mu(\bar{n})C \qquad (t\geq 0)(r=1,2,\ldots) \tag{4.1}$$

with $\bar{n} = ((n_1^1,n_1^2,\ldots),(n_2^1,n_2^2,\ldots),\ldots,(n_N^1,n_N^2,\ldots))$ denoting the number n_j^r of class r jobs at station j for all r and j, n^r the total number of class r jobs present, e_j^r representing an additional class r job at station j and $\alpha_j^r(\bar{n})$ the probability that upon arrival a class r job is assigned station j when the system is in state $\bar{n}$.

4.2 Non-exponential networks Theorem 2.1 is essentially based on only the differences in the reward structure (bias terms) due to a single arrival (see condition (2.10)). The internal network transition structure is thus not relevant other than for providing concrete estimates for these bias terms such as illustrated in section 3. Particularly, by relaxing the exponential service structure to include mixtures of Erlang distributions as service distributions and extending the state description by the number of residual exponential phases of service to be received, similar results can be expected. Since, however, mixtures of Erlang distributions arbitrarily closely approximate general nonnegative distributions (in the sense of weak convergence), by standard though highly technical weak convergence arguments on appropriate sample path spaces (e.g. [1], [3]) a non-exponential analog of theorem 2.1 seems possible. Most crucially, however, in the non-exponential case the verification of condition (2.10) or rather the estimation of the bias terms as in lemma 3.1 will become much more complex (cf. [23]).

4.3 Modified finite source inputs Slightly perturbed modifications of the given finite source input are also possible. In combination with perturbation results such as developed in [26], the error bound in theorem 2.1 can then be proven with an additional error bound due to the perturbation. For example, assume that the M sources have an exponential holding time with parameter $[\tilde{\lambda}/M]$ where $|\tilde{\lambda}-\lambda|\leq\varepsilon$. Then the error bound in (2.12) can be shown to be

$$(\lambda + \varepsilon)\ BC/M \tag{4.2}$$

4.4 Infinite service rates The boundedness assumption of the transition rates was made merely to justify a discrete time transformation based on the uniformization technique. This, however can be avoided in an approximative manner similarly to [22] in order to allow unbounded transition rates such as arising from infinite service rates as in infinite server stations.

Evaluation The question of open or closed modeling for queueing networks can be adressed by providing error bounds for the performance in the open and closed case. Simple througput bounds for large closed networks or approximate computational procedures for infinite open systems can hereby be justified. A general condition is provided from which such error bounds can be concluded. The verification of this condition can be established by inductively proving monotonicity results of total reward structures. Explicit error bounds of order 1/M with M the number of input sources in the closed case can so be derived. Extensions such as to multiclass non-exponential service networks seem possible.

The results apply both to product and non-product form networks, for example with blocking phenomena or breakdowns such as arising in performability analysis.

References

[1] Barbour, A.D. (1982), Generalized semi-Markov schemes and open queueing networks, *J. Appl. Prob.* 19, 469-474.

[2] Courtois, P.J. and Semal, P. (1984), Error bounds for the analysis by decomposition of non-negative matrices. In: *Mathematical Computer Performance and Reliability*, G. Jazeolla, P.J. Courtois and A. Hordijk, North-Holland, 209-224.

[3] Hordijk, A. and Schassberger, R. (1982), Weak convergence of generalized semi-Markov processes, *Stochastic Process. & Appl.* 2, 271-291.

[4] Hordijk, A. and Van Dijk, N.M. (1983), Networks of queues, *Lecture Notes in* Control and Information Sciences, *Vol. 60, Springer, Berlin, 158-205.*

[5] Jaiswal, N.K. (1968), *Priority Queues*, Academic Press, New York.

[6] Kato, T. (1966), Perturbation theory for linear operators, Springer, Berlin.

[7] Kemeny, J.G., Snell, J.L. and Knapp, A.W. (1966), *Denumerable Markov Chains*, Van Nostrand, Princeton, NJ.

[8] Meyer, J.F. (1982), Closed-form solutions of Performability, *IEEE Transactions on Computers*, 31, 648-657.

[9] Mitra, D. and McKenna, J. (1986), Asymptotic Expansions for Closed Markovian Networks with State Dependent Service Rates, *J.A.C.M.* 33, 568-592.

[10] Nelson, R. and Kleinrock, L. (1985), Rude-CSMA: A multihop channel access protocol, *IEEE Commun.*, COM-33, 785-791.

[11] Pinsky, E. and Yemini, Y. (1984), A statistical mechanics of some interconnection networks, Performance '84, *North-Holland*.

[12] Reiser, M. and Lavenberg, S.S. (1980), Mean value analysis of closed multichain queueing networks, *JACM*, 27(2): 313-322.

[13] Ross, S.M. (1970), Applied probability models with optimization application, Holden-Day, San Francisco.

[14] Rohlicek, J.R. and Willsky, A.S. (1988), The reduction of perturbed Markov generators: an algorithm exposing the role of transient states, *J.A.C.M.* 35, 675-696.

[15] Schwartz, M. (1987), Telecommunication networks, *Addison Wesley*.

[16] Seneta, E. (1980), Non-negative matrices and Markov chains, Springer-Verlag.

[17] Shanthikumar, J.G. and Yao, D. (1987), Stochastic monotonicity of the queue lenghts in closed queueing networks, *Oper. Res.* 35, 583-588.

[18] Smith, R.M., Trivedi, K.S. and Ramesh, A.V. (1988), Performability analysis: Measures, an algorithm and a case study, *IEEE Trans. Computers*, 37, 406-417.

[19] Stoyan, D. (1983), Comparison methods for queues and other stochastic models, Wiley, New York.

[20] Tijms, H.C. (1986), *Stochastic Modelling and Analysis*. A computational approach. Wiley, New York.

[21] Van Dijk, N.M., (1988), Simple bounds for queueing systems with breakdowns, *Performance Evaluation* 8, 117-128.

[22] Van Dijk, N.M., (1988), On the finite horizon Bellman equation for controlled Markov jump models with unbounded characteristics: Existence and approximation, *Stochastic Procc. Appl.*, 28, 141-157.

[23] Van Dijk, N.M., A formal proof for the insensitivity of simple bounds for finite multi-server non-exponential tandem queues based on monotonicity results, *Stochastic Procc. Appl.*, 27, 261-277.

[24] Van Dijk, N.M. (1988) Product forms for random access schemes, Research Report, Free University, Amsterdam.

[25] Van Dijk, N.M. (1988) Perturbation theory for Markov reward processes with unbounded characteristics, *Adv. Appl. Prob.* 20, 99-111.

[26] Van Dijk, N.M. and Puterman, M.L. (1988), Perturbation theory for Markov reward processes with applications to queueing systems. *Adv. Appl. Prob.* 20, 79-98.

[27] Van Dijk, N.M. and Van der Wal, J. (1987), Simple bounds and monotonicity results for multi-server exponential tandem queues, To appear: *Queueing Systems*.

[28] Whitt, W. (1978), Approximations of dynamic programs I. *Math. Oper. Res.* 3, 231-243.

[29] Whitt, W. (1981), Comparing counting processes and queues, *Adv. Appl. Prob.* 13, 207-220.

[30] Whitt, W. (1984), Open and closed models for networks of queues, *AT&T Bell Laboratories Technical Journal*, 63, 1911-1979.

[31] Yao, D.D. and Buzacott (1987), Modeling a class of flexible manufacturing systems with reversible routing, *Oper. Res.*, 35, 87-93.