Probability in the Engineering and Informational Sciences, **29**, 2015, 181–189.
doi:10.1017/S026996481400028X

ON THE ADAN–WEISS LOSS MODEL HAVING SKILL-BASED SERVERS AND LONGEST IDLE ASSIGNMENT RULE

Babak Haji
Department of Industrial and Systems Engineering, University of Southern California, Los Angeles, CA 90089, USA
E-mail: bhaji@usc.edu

We consider a queueing loss system with heterogeneous skill based servers with arbitrary distributions. We assume Poisson arrivals, with each arrival having a vector indicating which of the servers are eligible to serve it. Arrivals can only be assigned to a server that is both idle and eligible. We assume arrivals are assigned to the idle eligible server that has been idle the longest and derive, up to a multiplicative constant, the limiting distribution for this system. We show that the limiting probabilities of the ordered list of idle servers depend on the service time distributions only through their means. Moreover, conditional on the ordered list of idle servers, the remaining service times of the busy servers are independent and have their respective equilibrium service distributions. We also provide an algorithm using Gibbs sampler Markov Chain Monte Carlo method for estimating the limiting probabilities and other desired quantities of this system.

1. INTRODUCTION

In [2], Adan and Weiss introduced a queueing loss model with n servers having arbitrary service distributions, $G_i, i = 1, \ldots, n$. In their model, customers arrive according to a Poisson process and arrivals are discriminating, in the sense that each has an eligibility vector $(X_1, \ldots, X_n)$ with $X_i = 1$ if server i is eligible to serve that customer and $X_i = 0$ if ineligible, $i = 1, \ldots, n$. The vectors of successive arrivals are independent and identically distributed. An arrival finding all its eligible servers busy is lost, otherwise the arrival is assigned to idle eligible server that has been idle the longest. Using a supplementary variable approach to make their model Markovian, Adan and Weiss derived the limiting distribution for this model, and in doing so showed that the limiting distribution of the ordered list of idle servers depend on the service distributions only through their means. In this paper, we obtain results of [2] using the method of stages. This method considers the service time at each server as a weighted average of different gamma distributions with the same rate. Our approach will be to conjecture the reverse chain, and then show the conjecture is correct by finding, up to a multiplicative constant, the limiting probabilities of the ordered list of idle servers. We also show that given the set of busy servers, the remaining service times are independent with their respective equilibrium distributions (This result is only implicitly

 0269-9648/15

noted in Adan and Weiss [2].) In practice, when n is large finding the limiting probabilities or any other quantity of interest becomes computationally intractable. Therefore, we will provide an algorithm for simulating this model using the Gibbs Sampler method to estimate these desired quantities.

Recently, Haji and Ross [4] studied a similar model with the additional assumption that a random eligibility vector $(X_1, \ldots, X_n)$ is exchangeable, but allowed for a general class of operating policies which they refer to as "idle time ordering policies". They showed that all idle time-ordering policies resulted in the same limiting probabilities for the ordered list of idle servers which depend on the service distributions only through their means. In addition, given the set of idle servers, $I = \{i_1, \ldots, i_k\}$, all $k!$ possible orderings are equally likely and the remaining service times of the busy servers are independent and have their respective equilibrium service distributions.

Servers that can only serve certain arrivals are refereed to as "skill-based servers". Other papers concerned with skill-based servers are Whitt [9], Adan, Hurkens and Weiss [1], Ross [5], Talreja et al. [6], and Visschers, Adan, and Weiss [7].

2. MODEL ANALYSIS

Define the "idle server vector" as 0 if there are currently no idle servers, or as $i_1, \ldots, i_k$ if there are currently k servers idle, where i_1 has been idle the longest, i_2 the second longest, and so on. Let $P_{j:i_1,\ldots,i_k}$ be the probability that an arrival will be assigned to server i_j given that the idle server vector is $i_1, \ldots, i_k$. That is,

$$P_{j:i_1,\ldots,i_k} = P\left(\sum_{m=1}^{j-1} X_{i_m} = 0,\, X_{i_j} = 1\right).$$

Let $\beta_0 = 0$, and for $k > 0$ let $\beta_{i_1,\ldots,i_k} = P(\sum_{m=1}^{k} X_{i_m} > 0)$; hence, $\beta_{i_1,\ldots,i_k}$ is the probability that an arrival finding the idle server vector $i_1, \ldots, i_k$ will be served. Note that $\sum_{j=1}^{k} P_{j:i_1,\ldots,i_k} = \beta_{i_1,\ldots,i_k}$ and

$$P_{j:i_1,\ldots,i_k} = \beta_{i_1,\ldots,i_j} - \beta_{i_1,\ldots,i_{j-1}}.$$

Example 1: If we assume that $X_1, \ldots, X_n$ are independent and $X_i = 1$ with probability p_i, $i = 1, \ldots, n$, then $\beta_{i_1,\ldots,i_k} = 1 - \prod_{m=1}^{k}(1 - p_{i_m})$ and $P_{j:i_1,\ldots,i_k} = p_{i_j} \prod_{m=1}^{j-1}(1 - p_{i_m})$.

A random variable X is called general Erlang, $GE(N, \mu)$, if

$$X = \sum_{i=1}^{N} W_i$$

where W_i, $i = 1, 2, 3, \ldots$, are *iid* exponential random variables with rate μ and N is a positive integer valued random variable that is independent of the W_i's. Let G_e be the equilibrium distribution of X.

Lemma 1: *G_e is the distribution function of a $GE(N_e, \mu)$ random variable, where*

$$P(N_e = j) = P(N \geq j)/E(N). \qquad (1)$$

Proof: See Haji and Ross [4]. ■

In our model, we suppose that the service distribution of server i is general Erlang $GE(N_i, \mu_i)$, $i = 1, \ldots, n$, and we analyze our model as a continuous time Markov chain. To do so, we define the "state vector" as $(0, \mathbf{r})$ with $\mathbf{r} = (r_1, \ldots, r_n)$, if there are currently no idle servers and server i has r_i remaining exponential stages with rate μ_i in order to complete its service; or as $(i_1, \ldots, i_k, \mathbf{r})$, if $i_1, \ldots, i_k$ is the current idle server vector and each server i has r_i exponential stages to complete, where $r_i = 0$ for all $i \in \{i_1, \ldots, i_k\}$.

Proposition 1: *For general Erlang service times, $GE(N_i, \mu_i)$ $i = 1, \ldots, n$, where N_i is a random variable with $p_i(j) = P(N_i = j)$, the stationary probability of the state vector $(i_1, \ldots, i_k, \mathbf{r})$ has the following form:*

$$P(i_1, \ldots, i_k, \mathbf{r}) = \frac{\mu_{i_k} \cdots \mu_{i_1}}{\lambda^k \beta_{i_1} \cdots \beta_{i_1,\ldots,i_k}} P(0, \mathbf{1}) \prod_{m \notin \{i_1,\ldots,i_k\}} P(N_m \geq r_m),$$

where $P(0, \mathbf{1})$ is such that

$$P(0, \mathbf{1})(1 + \sum_{(i_1,\ldots,i_k,\mathbf{r})} \frac{\mu_{i_k} \cdots \mu_{i_1}}{\lambda^k \beta_{i_1} \cdots \beta_{i_1,\ldots,i_k}} \prod_{m \notin \{i_1,\ldots,i_k\}} P(N_m \geq r_m)) = 1.$$

Proof: Using longest idle rule, $P_{j:i_1,\ldots,i_k}$ is the probability that the idle time-ordering policy assigns an arrival to server i_j when the state is $(i_1, \ldots, i_k, \mathbf{r})$. For states

$$\begin{aligned}
\mathbf{x} &= (i_1, \ldots, i_k : r_1, \ldots, r_n), \\
\mathbf{x}^* &= (i_1, \ldots, i_k : r_1, \ldots, r_j - 1, \ldots, r_n), \\
\mathbf{x}^+ &= (i_1, \ldots, i_k, i_{k+1} : r_1, \ldots, r_{i_{k+1}-1}, 0, r_{i_{k+1}+1}, \ldots, r_n), \\
\mathbf{x}^- &= (i_1, \ldots, i_{j-1}, i_{j+1}, \ldots, i_k : r_1, \ldots, r_{i_j-1}, s, r_{i_j+1}, \ldots, r_n),
\end{aligned}$$

the infinitesimal rates of the resultant continuous time Markov chain are

$$\begin{aligned}
q_{\mathbf{x},\mathbf{x}^*} &= \mu_j, && \text{for } r_j > 1, \\
q_{\mathbf{x},\mathbf{x}^+} &= \mu_{i_{k+1}}, && \text{for } r_{i_{k+1}} = 1, \\
q_{\mathbf{x},\mathbf{x}^-} &= \lambda P_{j:i_1,\ldots,i_k} p_{i_j}(s).
\end{aligned}$$

We now make the following conjecture about the reverse process.

(a) It is a queueing model with n servers all of whom are eligible to serve any arriving customer.

(b) The state is $\mathbf{x} = (i_1, \ldots, i_k : r_1, \ldots, r_n)$ if $(i_1, \ldots, i_k)$ is the current idle server vector; r_j is the current stage of server j if that server is busy, and $r_j = 0$ if j is idle.

(c) An arrival to server j begins in stage 1; and the time it takes server j to complete a stage is exponential with rate μ_j, $j = 1, \ldots, n$.

(d) Customer at server j leaves the system upon completion of stage m with probability $\lambda_j(m) = \frac{p_j(m)}{\sum_{k \geq m} p_j(k)}$; otherwise it goes to stage $m+1$ with probability $\bar{\lambda}_j(m) = 1 - \lambda_j(m)$.

(e) The arrival rate of customers when the order list of idle servers is $i_1, \ldots, i_k$ is $\lambda \beta_{i_1,\ldots,i_k}$, and the arriving customer is assigned to server i_k.

(f) If server r becomes idle when the idle server vector is $i_1, \ldots, i_k$ then the new idle server vector becomes $i_1, \ldots, i_{j-1}, r, i_j, \ldots, i_k$ with probability

$$\frac{P_{j:i_1,\ldots,i_{j-1},r,i_j,\ldots,i_k}\beta_{i_1}\cdots\beta_{i_1,\ldots,i_k}}{\beta_{i_1}\cdots\beta_{i_1,\ldots,i_{j-1}}\beta_{i_1,\ldots,i_{j-1},r}\beta_{i_1,\ldots,i_{j-1},r,i_{j+1}}\cdots\beta_{i_1,\ldots,i_{j-1},r,i_{j+1},\ldots,i_k}}.$$

With $\mathbf{x}, \mathbf{x}^*, \mathbf{x}^+$ and $\mathbf{x}^-$ as defined in the preceding, the infinitesimal rates of the reversed chain under our conjecture are

$$\begin{aligned} q^*_{\mathbf{x}^*,\mathbf{x}} &= \bar{\lambda}_j(r_j - 1)\mu_j, \\ q^*_{\mathbf{x}^+,\mathbf{x}} &= \lambda\beta_{i_1,\ldots,i_{k+1}}, \\ q^*_{\mathbf{x}^-,\mathbf{x}} &= \mu_{i_j}\lambda_{i_j}(s)\frac{P_{j:i_1,\ldots,i_k}\beta_{i_1}\cdots\beta_{i_1,\ldots,i_{j-1}}\beta_{i_1,\ldots,i_{j-1},i_{j+1}}\cdots\beta_{i_1,\ldots,i_{j-1},i_{j+1},\ldots,i_k}}{\beta_{i_1}\cdots\beta_{i_1,\ldots,i_k}}. \end{aligned}$$

In order to verify our conjecture we first need to show that when in state $(i_1, \ldots, i_k, \mathbf{r})$ the rates at which the forward and the reverse process leave that state are equal.

Lemma 2: *The rates at which the forward and the reverse process leave the state $(i_1, \ldots, i_k, \mathbf{r})$ are equal; i.e.*

$$\sum_{\mathbf{x}^k \neq \mathbf{x}} q_{\mathbf{x},\mathbf{x}^k} = \sum_{\mathbf{x}^k \neq \mathbf{x}} q^*_{\mathbf{x},\mathbf{x}^k}.$$

Proof: Using the infinitesimal rates of the forward and the reverse process the preceding equality can be written as

$$\begin{aligned} &\sum_{j=1}^{k}\lambda P_{j:i_1,\ldots,i_k} + \sum_{i_j\notin\{i_1,\ldots,i_k\}}\mu_{i_j} \\ &\quad = \lambda\beta_{i_1,\ldots,i_k} + \sum_{i_j\notin\{i_1,\ldots,i_k\}}\sum_{m=1}^{k+1}\frac{P_{m:\mathbf{x}^m}\beta_{i_1}\cdots\beta_{i_1,\ldots,i_k}}{\beta_{i_1}\cdots\beta_{i_1,\ldots,i_{m-1}}\beta_{i_1,\ldots,i_{m-1},i_j}\cdots\beta_{i_1,\ldots,i_{m-1},i_j,i_m,\ldots,i_k}}\mu_{i_j}, \end{aligned}$$

where $\mathbf{x}^m = (i_1, \ldots, i_{m-1}, i_j, i_m, \ldots i_k)$ and $P_{m:\mathbf{x}^m} = P_{m:i_1,\ldots,i_{m-1},i_j,i_m,\ldots i_k}$. Since $\sum_{j=1}^k P_{j:i_1,\ldots,i_k} = \beta_{i_1,\ldots,i_k}$, to show that the preceding equality holds it suffices to show that

$$\sum_{m=1}^{k+1}\frac{P_{m,\mathbf{x}^m}\beta_{i_1}\cdots\beta_{i_1,\ldots,i_k}}{\beta_{i_1},\cdots\beta_{i_1,\ldots,i_{m-1}}\beta_{i_1,\ldots,i_{m-1},i_j}\cdots\beta_{i_1,\ldots,i_{m-1},i_j,i_m,\ldots,i_k}} = 1. \quad \mathbf{(2)}$$

This will be proven by first showing that for $1 \leq n \leq k$,

$$\sum_{m=1}^{n}\frac{P_{m,\mathbf{x}^m}\beta_{i_1},\ldots\beta_{i_1,\ldots,i_k}}{\beta_{i_1}\cdots\beta_{i_1,\ldots,i_{m-1}}\beta_{i_1,\ldots,i_{m-1},i_j}\cdots\beta_{i_1,\ldots,i_{m-1},i_j,i_m,\ldots,i_k}} = \frac{\beta_{i_1,\ldots,i_n}\cdots\beta_{i_1,\ldots,i_k}}{\beta_{i_j,i_1,\ldots,i_n}\cdots\beta_{i_j,i_1,\ldots,i_k}}. \quad \mathbf{(3)}$$

We will prove (3) by induction. It is clear that for $n = 1$ this is true. Now assuming that (3) is true for n we will show that it is also true for $n + 1$. Namely,

$$\begin{aligned}
&\sum_{m=1}^{n+1} \frac{P_{m,\mathbf{x}^m}\beta_{i_1}\cdots\beta_{i_1,\ldots,i_k}}{\beta_{i_1}\cdots\beta_{i_1,\ldots,i_{m-1}}\beta_{i_1,\ldots,i_{m-1},i_j}\cdots\beta_{i_1,\ldots,i_{m-1},i_j,i_m,\ldots,i_k}}\\
&\quad= \sum_{m=1}^{n} \frac{P_{m,\mathbf{x}^m}\beta_{i_1}\cdots\beta_{i_1,\ldots,i_k}}{\beta_{i_1}\cdots\beta_{i_1,\ldots,i_{m-1}}\beta_{i_1,\ldots,i_{m-1},i_j}\cdots\beta_{i_1,\ldots,i_{m-1},i_j,i_m,\ldots,i_k}}\\
&\qquad+ \frac{P_{n+1,\mathbf{x}^{n+1}}\beta_{i_1}\cdots\beta_{i_1,\ldots,i_k}}{\beta_{i_1}\cdots\beta_{i_1,\ldots,i_n}\beta_{i_1,\ldots,i_n,i_j}\cdots\beta_{i_1,\ldots,i_n,i_j,i_{n+1},\ldots,i_k}}\\
&\quad= \frac{\beta_{i_1,\ldots,i_n}\cdots\beta_{i_1,\ldots,i_k}}{\beta_{i_j,i_1,\ldots,i_n}\cdots\beta_{i_j,i_1,\ldots,i_k}} + \frac{P_{n+1,\mathbf{x}^{n+1}}\beta_{i_1,\ldots,i_{n+1}}\cdots\beta_{i_1,\ldots,i_k}}{\beta_{i_1,\ldots,i_n,i_j}\cdots\beta_{i_1,\ldots,i_n,i_j,i_{n+1},\ldots,i_k}}\\
&\quad= \frac{\beta_{i_1,\ldots,i_{n+1}}\cdots\beta_{i_1,\ldots,i_k}}{\beta_{i_j,i_1,\ldots,i_{n+1}}\cdots\beta_{i_j,i_1,\ldots,i_k}},
\end{aligned}$$

the last equality holds because $P_{n+1,\mathbf{x}^{n+1}} + \beta_{i_1,\ldots,i_n} = \beta_{i_j,i_1,\ldots,i_n}$. Using (3) it is easy to show that (2) holds. That is,

$$\begin{aligned}
&\sum_{m=1}^{k+1} \frac{P_{m,\mathbf{x}^m}\beta_{i_1}\cdots\beta_{i_1,\ldots,i_k}}{\beta_{i_1}\cdots\beta_{i_1,\ldots,i_{m-1}}\beta_{i_1,\ldots,i_{m-1},i_j}\cdots\beta_{i_1,\ldots,i_{m-1},i_j,i_m,\ldots,i_k}}\\
&\quad= \sum_{m=1}^{k} \frac{P_{m,\mathbf{x}^m}\beta_{i_1}\cdots\beta_{i_1,\ldots,i_k}}{\beta_{i_1}\cdots\beta_{i_1,\ldots,i_{m-1}}\beta_{i_1,\ldots,i_{m-1},i_j}\cdots\beta_{i_1,\ldots,i_{m-1},i_j,i_m,\ldots,i_k}} + \frac{P_{k+1,\mathbf{x}^{k+1}}\beta_{i_1}\cdots\beta_{i_1,\ldots,i_k}}{\beta_{i_1}\cdots\beta_{i_1,\ldots,i_k}\beta_{i_1,\ldots,i_k,i_j}}\\
&\quad= \frac{\beta_{i_1,\ldots,i_k}}{\beta_{i_j,i_1,\ldots,i_k}} + \frac{P_{k+1,\mathbf{x}^{k+1}}}{\beta_{i_1,\ldots,i_k,i_j}} = 1.
\end{aligned}$$

Therefore, Lemma 2 is proven. ■

Now that we showed the rates which the forward and the reverse process leave state $(i_1, \ldots, i_k, \mathbf{r})$ are equal our conjecture will be verified if we can find probabilities $P(\mathbf{x})$, $\sum_{\mathbf{x}} P(\mathbf{x}) = 1$, such that

$$\begin{aligned}
P(\mathbf{x})q_{\mathbf{x},\mathbf{x}^*} &= P(\mathbf{x}^*)q^*_{\mathbf{x}^*,\mathbf{x}} && \text{for } r_j > 1,\\
P(\mathbf{x})q_{\mathbf{x},\mathbf{x}^+} &= P(\mathbf{x}^+)q^*_{\mathbf{x}^+,\mathbf{x}} && \text{for } r_{i_{k+1}} = 1,\\
P(\mathbf{x})q_{\mathbf{x},\mathbf{x}^-} &= P(\mathbf{x}^-)q^*_{\mathbf{x}^-,\mathbf{x}}.
\end{aligned}$$

Thus, we must find probabilities that satisfy

$$P(\mathbf{x})\mu_j = P(\mathbf{x}^*)\bar{\lambda}_j(r_j - 1)\mu_j \qquad \text{for } r_j > 1, \tag{4}$$

$$P(\mathbf{x})\mu_{i_{k+1}} = P(\mathbf{x}^+)\lambda\beta_{i_1,\ldots,i_{k+1}} \qquad \text{for } r_{i_{k+1}} = 1, \tag{5}$$

$$\begin{aligned}
P(\mathbf{x})\lambda P_{j,k}p_{i_j}(s) &= P(\mathbf{x}^-)\mu_{i_j}\lambda_{i_j}(s)\\
&\quad\times \frac{P_{j:i_1,\ldots,i_k}\beta_{i_1}\cdots\beta_{i_1,\ldots,i_{j-1}}\beta_{i_1,\ldots,i_{j-1},i_{j+1}}\cdots\beta_{i_1,\ldots,i_{j-1},i_{j+1},\ldots,i_k}}{\beta_{i_1}\cdots\beta_{i_1,\ldots,i_k}}.
\end{aligned} \tag{6}$$

https://doi.org/10.1017/S026996481400028X Published online by Cambridge University Press

Rewriting the first and the second equations we have,

$$P(\mathbf{x}^*) = \frac{1}{\bar{\lambda}_j(r_j-1)}P(\mathbf{x}), \quad \text{for } r_j > 1, \tag{7}$$

$$P(\mathbf{x}^+) = \frac{\mu_{i_{k+1}}}{\lambda\beta_{i_1,\ldots,i_{k+1}}}P(\mathbf{x}). \tag{8}$$

By first iterating over (8) and then (7) we have,

$$P(\mathbf{x}) = \frac{\mu_{i_k}\cdots\mu_{i_1}}{\lambda^k\beta_{i_1}\cdots\beta_{i_1,\ldots,i_k}}P(0,\mathbf{1})\prod_{m\notin\{i_1,\ldots,i_k\}}\prod_{i=1}^{r_m-1}\bar{\lambda}_m(i).$$

Using that $\prod_{i=1}^{n-1}\bar{\lambda}_m(i) = P(N_m \geq n)$, the preceding gives

$$P(\mathbf{x}) = \frac{\mu_{i_k}\cdots\mu_{i_1}}{\lambda^k\beta_{i_1}\cdots\beta_{i_1,\ldots,i_k}}P(0,\mathbf{1})\prod_{m\notin\{i_1,\ldots,i_k\}}P(N_m \geq r_m).$$

As it is straightforward to verify that the preceding, with $P(0,\mathbf{1})$ chosen to make the probabilities sum to 1, satisfy the reversibility Eqs. (4)–(6), the proposition is proven. ■

THEOREM 1: *Suppose service distributions are $G_1,\ldots,G_n$, and let $E[S_j]$ be the mean of the distribution G_j. If $i_1,\ldots,i_k$ is the idle server vector in steady state, then*

$$P(i_1,\ldots,i_k) = \frac{1}{\lambda^k\beta_{i_1}\cdots\beta_{i_1,\ldots,i_k}E(S_{i_1})\cdots E(S_{i_k})}P(0),$$

where $P(0)$ is the probability that all the servers are busy. Furthermore, given that idle server vector is $i_1,\ldots,i_k$,

(a) *The limiting probabilities of the idle server vector $i_1,\ldots,i_k$ depend on the service distributions only trough their means;*
(b) *The remaining service times of the busy servers are independent and are distributed according to their respective equilibrium service distributions;*
(c) *The amount of service time already provided on their current customers by the busy servers are independent and are distributed according to their respective equilibrium service distributions.*

PROOF: To begin, suppose that the service distribution of server i is general Erlang $GE(N_i,\mu_i)$, $i=1,\ldots,n$. Let $P(i_1,\ldots,i_k)$ be the steady-state probability that $i_1,\ldots,i_k$ is the idle server vector. Using Proposition 1 and summing $P(i_1,\ldots,i_k,\mathbf{r})$ over all the consistent vectors $\mathbf{r}$ (that is all $\mathbf{r}$ such that $r_j = 0, j\in\{i_1,\ldots,i_k\}$) yields

$$\begin{aligned}P(i_1,\ldots,i_k) &= \sum_{\mathbf{r}}\frac{\mu_{i_k}\cdots\mu_{i_1}}{\lambda^k\beta_{i_1}\cdots\beta_{i_1,\ldots,i_k}}P(0,\mathbf{1})\prod_{m\notin\{i_1,\ldots,i_k\}}P(N_m \geq r_m)\\ &= \frac{\mu_{i_k}\cdots\mu_{i_1}}{\lambda^k\beta_{i_1}\cdots\beta_{i_1,\ldots,i_k}}P(0,\mathbf{1})\sum_{\mathbf{r}}\prod_{m\notin\{i_1,\ldots,i_k\}}P(N_m \geq r_m).\end{aligned}$$

Now if we let the set of the busy servers be $b_1, \ldots, b_{n-k}$ which is the compliment of the set of idle servers $i_1, \ldots, i_k$ we can write

$$P(i_1, \ldots, i_k) = \frac{\mu_{i_k} \cdots \mu_{i_1}}{\lambda^k \beta_{i_1} \cdots \beta_{i_1,\ldots,i_k}} P(0, \mathbf{1}) \sum_{r_{b_1}} \cdots \sum_{r_{b_{n-k}}} P(N_{b_1} \geq r_{b_1}) \ldots P(N_{b_{n-k}} \geq r_{b_{n-k}})$$

$$= \frac{\mu_{i_k} \cdots \mu_{i_1}}{\lambda^k \beta_{i_1} \cdots \beta_{i_1,\ldots,i_k}} P(0, \mathbf{1}) \prod_{m \notin \{i_1,\ldots,i_k\}} E(N_m). \quad \textbf{(9)}$$

Let $P(0)$ be the probability that all the servers are busy. Clearly,

$$P(0) = \sum_{\mathbf{r}} P(0, \mathbf{r}) = \sum_{\mathbf{r}} P(0, \mathbf{1}) \prod_{i=1}^{n} P(N_i \geq r_i) = P(0, \mathbf{1}) \prod_{i=1}^{n} E(N_i).$$

Hence, using the fact that the mean service time at server i is $E(S_i) = \frac{E(N_i)}{\mu_i}$, $i = 1, \ldots, n$, we can rewrite the limiting probabilities of the ordered list of idle servers (9) as follows:

$$P(i_1, \ldots, i_k) = \frac{1}{\lambda^k \beta_{i_1} \cdots \beta_{i_1,\ldots,i_k} E(S_{i_1}) \cdots E(S_{i_k})} P(0),$$

which shows that, conditional on the idle server vector $i_1, \ldots, i_k$, the limiting probabilities depend on the service distributions only through their means.

Moreover, it follows from Proposition 1 and (9) that

$$\frac{P(i_1, \ldots, i_k, \mathbf{r})}{P(i_1, \ldots, i_k)} = \prod_{m \notin \{i_1,\ldots,i_k\}} \frac{P(N_m \geq r_m)}{E(n_m)},$$

which, using Lemma 1, proves that conditional on the set of busy servers, their remaining service times are independent and are distributed according to their respective equilibrium service distributions. In addition, because the reverse chain has the same stationary probabilities as does the forward chain and as the interpretation of r_i for the reverse chain is that server i is currently at stage r_i, part (c) also follows. Hence, the theorem is proven when all service distributions are general Erlang type. Because any service distribution is the limit of a sequence of general Erlang distributions, see for instance Whitt [8], the approach of Barbour [3] can be used to establish necessary continuity arguments for extending the results for general Erlang service distributions to any arbitrary service distributions. Thus, the theorem is proven for arbitrary service distributions. ■

3. FINDING THE LIMITING PROBABILITIES USING MARKOV CHAIN MONTE CARLO METHODS

In practice, when the number of servers is large, the determination of the constant $P(0)$ is computationally intractable. Indeed, even if it were known, the derivation of other quantities of interest, such as average waiting time in the system, rate at which customers are lost, etc., remain computationally intractable. However, these quantities can be determined by using the Gibbs sampler Markov chain Monte Carlo method to generate a Markov chain whose limiting distribution is the stationary distribution of the ordered list of the set of idle servers. That is, by letting Y_i to be equal to m if server i is the mth server on the idle server vector and 0 if it is busy, we want to generate a Markov chain whose stationary distribution is

$$p(x_1,\ldots,x_n) = P(Y_i = x_i, i = 1,\ldots,n)$$
$$= \begin{cases} C\dfrac{1}{\lambda^k \beta_{i_1}\cdots\beta_{i_1,\ldots,i_k}\prod_{j=1}^{k} E(S_{i_j})}, & x_{i_m} = m, m \le k \text{ and } x_{i_m} = 0, m > k, \\ 0, & \text{otherwise.} \end{cases}$$

When the current state of the Markov chain is $\mathbf{x} = (x_1,\ldots,x_n)$, the Gibbs sampler method chooses a coordinate that is equally likely to be any of $1,\ldots,n$. If coordinate j is chosen then with $r = \max_{i\neq j}\{x_i\}$ if $0 < x_j < r$ the next state is $\mathbf{x}$, otherwise the next state is $(x_1,\ldots,x_{j-1},0,x_{j+1},\ldots,x_n)$ with probability

$$\alpha = \frac{p(x_1,\ldots,x_{j-1},0,x_{j+1},\ldots,x_n)}{p(x_1,\ldots,x_{j-1},0,x_{j+1},\ldots,x_n) + p(x_1,\ldots,x_{j-1},r+1,x_{j+1},\ldots,x_n)}$$
$$= \frac{\lambda\beta_{i_1,\ldots,i_r,j}E(S_j)}{1+\lambda\beta_{i_1,\ldots,i_r,j}E(S_j)};$$

or it will be $(x_1,\ldots,x_{j-1},r+1,x_{j+1},\ldots,x_n)$ with probability $1-\alpha$.

The stationary distribution of the successive values of a Markov chain generated by the preceding is the limiting distribution of the ordered list of the set of idle servers. Thus, for instance, we can approximate the steady-state probability there are exactly k idle servers, call it $P(k)$, by the proportion of states $(x_1,\ldots,x_n)$ such that $\sum_{i=1}^{n} x_i = k(k+1)/2$. We can also approximate the steady-state probability that the ordered list of the set of idle servers is $i_1,\ldots,i_k$, call it $P(i_1,\ldots,i_k)$, by the proportion of states $(x_1,\ldots,x_n)$ such that $\sum_{j=1}^{k} x_{i_j} = k(k+1)/2$ and $\max_i(x_i) = k$. We can then use our estimates of $P(i_1,\ldots,i_k)$, $k = 0,\ldots,n$, to estimate $\sum_{k=1}^{n}\sum_{(i_1,\ldots,i_k)} P(i_1,\ldots,i_k)\beta_{i_1,\ldots,i_k}$, equal to the proportion of arrivals that enter the system.

Acknowledgments

This material is based upon work supported by the National Science Foundation under contract/grant number CMMI1233337 and by the US Army Research Laboratory and the US Army Research Office under grant W911NF-11-1-0015

References

1. Adan, I.J.B.F., Hurkens, C., & Weiss, G. (2010). A reversible Erlang loss system with multitype customers and multitype servers. *Probability in the Engineering and Informational Sciences* 24: 535–548.
2. Adan, I.J.B.F. & Weiss, G. (2012). A loss system with skill-based servers under assign to longest idle server policy. *Probability in the Engineering and Informational Sciences* 26: 307–321.
3. Barbour, A.D. (1976). Networks of queues and the method of stages. *Advances in Applied Probability* 8: 584–591.
4. Haji, B. & Ross, S.M. (2015). A queueing loss model with heterogeneous skill based servers under idle time ordering policies. *Journal of Applied Probability.* (accepted for publication)
5. Ross, S.M. (2014). Optimal server selection in a queueing loss model with heterogeneous exponential servers, discriminating arrivals, and arbitrary arrival times. *Journal of Applied Probability* 51: 880–884.

6. Talreja, R. & Whitt, W. (2007). Fluid models for overloaded multi-class many-service queueing systems with FCFS routing. *Management Science* 54: 1513–1527.
7. Visschers, J.W.C.H., Adan, I.J.B.F., & Weiss, G. (2012). A product form solution to a system with multi-type jobs and multi-type servers. *Queueing Systems* 70: 269–298.
8. Whitt, W. (1974). The continuity of Queues. *Advances in Applied Probability* 6: 175–183.
9. Whitt, W. (2006). A multi-class fluid model for a contact center with skill-based routing. *International Journal of Electronics and Communications* 60: 95–102.

https://doi.org/10.1017/S026996481400028X Published online by Cambridge University Press