Hindawi Publishing Corporation
Journal of Applied Mathematics and Decision Sciences
Volume 2007, Article ID 68280, 10 pages
doi:10.1155/2007/68280

Research Article

A Paradox in a Queueing Network with State-Dependent Routing and Loss

Ilze Ziedins

Received 25 May 2007; Accepted 8 August 2007

Recommended by Paul Cowpertwait

Consider a network of parallel finite tandem queues with two stages, where each arrival attempts to minimize its own cost due to loss. It is known that the user optimal and asymptotic system optimal policies may differ—we give examples showing that they may differ for finite systems and that as the service rate is increased at the second stage the user optimal policy may change in such a way that the total expected cost due to loss increases.

1. Introduction and background

Queueing networks often exhibit seemingly paradoxical behaviour, where adding capacity either in the form of extra capacity on links or at nodes, or even extra links or routes, does not always lead to an improvement in performance, and may even lead to a severe degradation in performance. The classic and very well-known example of this is Braess's paradox, which has been much studied, both in the traffic literature, and in the queueing theory literature (see [1, 2] for the original paper and [3] for a comprehensive list of references maintained by Braess). This was one of the examples mentioned by Professor Jeff Hunter in his inaugural lecture. It is therefore a great pleasure to write on a different kind of paradox in this festschrift for Jeff.

In addition to Braess's paradox, there are several other well-known paradoxes (see, e.g., Arnott and Small [4] and the references therein). The paradoxes can be ascribed to the difference between system optimal and perceived user optimal behaviour—that is, if individuals behave selfishly in a way that minimizes their own measure of cost (e.g., the delay in transit through a network), then the system as a whole can suffer, and flows can alter in such a way that all individuals see worse performance. This is most clearly

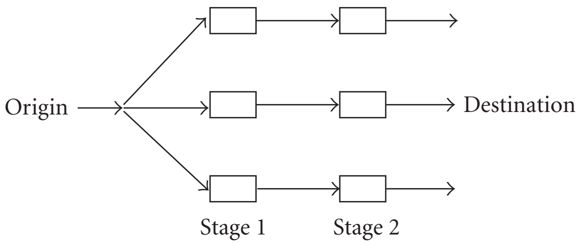

Figure 1.1. Three parallel two-stage tandem queues.

seen in traffic and transportation networks, and many of the early papers were directed to this application (see, e.g., [1, 5–8]). More recently, the phenomenon of selfish routing has become increasingly important in the context of telecommunication and computer networks (e.g., [9–13]).

The model considered here consists of a system of K parallel finite tandem queues with two stages, with a stream of arrivals who can be sent to any one of the queues. Figure 1.1 illustrates such a system with 3 parallel tandem queues. The objective is to minimize the cost due to loss, rather than minimize delay, the usual performance measure. Loss occurs when an individual attempts to enter a finite queue, but is unable to do so because it is full—that individual is then lost to the system. This assumption, while not at all realistic for traffic networks, is realistic for computer networks, where packets attempting to join a buffer that is full are lost and, if necessary, resent from the origin. This model was earlier studied in [14] (which obtained some properties of user optimal routing policies) and [15] (which obtained the asymptotic as $K \to \infty$ system optimal routing policies); and these papers should be referred to for a more complete list of earlier, relevant references. In [14] it was shown that user optimal policies may be paradoxical in the sense that arrivals may choose a queue with greater occupancy to minimize their probability of loss. The main contribution of this paper is to give an example showing that this paradoxical behaviour may then have further consequences—under user optimal routing it is possible for the expected cost, both to the system as a whole, and to the individual user, to *increase* when the service rate at the second queue is increased.

The model discussed here differs in some respects from those commonly studied in the literature. The most obvious difference is that it is a queueing system with loss. Paradoxes in loss networks have been studied [9, 10], but networks with finite queues and loss have been seen more rarely. The usual performance measure of interest is delay, and in most papers it is then also assumed that queues are infinite when studying paradoxes. Another difference is that here we concentrate on state-dependent routing, whereas routing paradoxes have most commonly been studied with probabilistic routing (see, e.g., [11, 16, 17] for some examples of other studies with state-dependent routing). For the model in this paper, with state-dependent routing, it has already been shown in [14] that when individuals are able to choose a route to minimize their own loss, they may choose a route where the occupancy is greater, particularly at the first stage of the tandem. Here, we illustrate by example the difference in the expected cost between the user optimal policy and the system optimal policy when the number of queues, K, is finite, and we compare both with the asymptotically optimal cost. We then give an example showing that when the service rate at the second stage of the tandem is increased, the expected cost under the user optimal policy may increase, rather than decrease as might be expected.

In Section 2, we give a detailed description of the parallel tandem queue model, and we give some of the results that we will be using from [14, 15]. In Section 3, we give examples comparing the performance of state-dependent user optimal and system optimal routing schemes, looking both at asymptotic results taken from [15] and results for a finite system. We conclude with a short discussion in Section 4.

2. Definitions and preliminaries

This section gives a detailed description of the network and outlines some results from previous papers, in particular [14, 15].

Consider a network with K parallel tandem queues. Each tandem queue has two single server finite queues in sequence, which we refer to as stages, the first stage having service rate μ_1 and capacity C_1 and the second service rate μ_2 and capacity C_2. Arrivals to the system are as a Poisson process with rate $K\lambda$ so that the arrival rate scales with the size of the system. An arrival can be routed to any one of the K queues, but once it has joined a tandem queue, it cannot change to a different queue, so there is no interaction between the queues beyond that induced by any controls over the routing of arrivals. We assume that all interarrival times and service times are exponentially distributed and independent of one another, so the system can be modelled as a Markov process.

Since the queues are finite, not all arrivals will necessarily be accepted into the system. Moreover, since the second queue in the tandem is finite, it is possible for an individual to finish service at the first queue, but find that the second queue is full, so that they are unable to join it—we assume that in that case, the individual leaves the system, and is lost, without completing service at both queues.

We assume that the objective is to minimize the cost due to losing or blocking individuals. In a system such as this, it may also be possible to consider minimizing the delay, conditional on not being lost, but we do not do so here. Previous analyses of paradoxes in queueing networks (not loss networks) have often focussed on routing or other controls that minimize or equalize delay rather than loss, because the common assumption is that the queues have infinite capacity—however, in practice, queues are often finite, so that minimizing loss in networks with finite queues also needs to be considered.

Let d_1 be the cost of losing an individual on entry to queue 1, and d_2 the cost of losing an individual on entry to queue 2. We will examine both system optimal and user optimal policies, as well as a range of intermediate policies.

Under probabilistic routing, each arrival chooses queue k with probability p_k, $1 \le k \le K$, where $\sum_k p_k = 1$ independently of all other routing decisions, service times, and arrival times. In that case it is sufficient to consider the tandem queues separately, with arrival rate $K\lambda p_k$ at the kth tandem queue. The state for a single queue is then given by (i, j), where i is the occupancy at the first queue, and j the occupancy at the second. This single queue has state space $S_1 = \{(i,j) : 0 \le i \le C_1,\ 0 \le j \le C_2\}$ and transition rates

$$(i,j) \longrightarrow \begin{cases} (i+1,j) & K\lambda p_k \quad \text{if } i < C_1, \\ (i-1,j+1) & \mu_1 \quad \text{if } i > 0,\ j < C_2, \\ (i-1,j) & \mu_1 \quad \text{if } i > 0,\ j = C_2, \\ (i,j-1) & \mu_2 \quad \text{if } j > 0. \end{cases} \tag{2.1}$$

Let the equilibrium distribution for a single tandem queue with these transition probabilities be denoted by $\tilde{\pi}(n)$, $n \in S_1$. The system optimal policy is found by minimizing a weighted sum of the cost of loss for the individual queues. In contrast, the user optimal policy is one where the costs of loss at all queues that are in use are equalized. For this model, under probabilistic routing, the system optimal and user optimal policies coincide with $p_k = 1/K$, $1 \le k \le K$ and the flow of arrivals is divided equally between the tandem queues, so that each tandem queue has arrival rate λ.

Under state-dependent routing the analysis is considerably more complicated and it is necessary to consider the state of the whole network simultaneously. Now, let n_{ij} be the number of tandem queues with occupancy i at the first queue, and j at the second queue. Then $n = (n_{ij} : 0 \le i \le C_1,\ 0 \le j \le C_2)$ is a Markov process with state space $S = \{n : \sum_{ij} n_{ij} = K,\ n_{ij} \in \{0,1,2,\dots,K\},\ 0 \le i \le C_1,\ 0 \le j \le C_2\}$ and transition rates partly depending on the routing rule. Denote by $r_n(i,j)$ the probability an arrival is sent to a tandem queue in state (i,j) if the network is in state n, and let $r_n(b)$ be the probability that an arrival is lost in state n, where $r_n(b) + \sum_{ij} r_n(i,j) = 1$ for all $n \in S$. Let $R = \{r_n(i,j), r_n(b);\ n \in S,\ 0 \le i \le C_1,\ 0 \le j \le C_2,\ 0 \le r_n(i,j), r_n(b) \le 1\}$ denote a particular state-dependent admission and routing policy. Note that for a finite system, since this is a Markov decision process, the system optimal policy will have $r_n(i,j), r_n(b) \in \{0,1\}$. Given some linear ordering of the states (i,j), let e_{ij} denote the $(C_1+1) \times (C_2+1)$ unit vector with the ijth entry equal to 1, and the remaining entries equal to 0. Then the transition rates under policy R are given by

$$n \longrightarrow \begin{cases} n - e_{ij} + e_{i+1,j} & K\lambda r_n(i,j) \quad \text{for } (i,j) \in S_1, \\ n - e_{ij} + e_{i-1,j+1} & n_{ij}\mu_1 \quad \text{if } i > 0,\ j < C_2, \\ n - e_{ij} + e_{i-1,j} & n_{ij}\mu_1 \quad \text{if } i > 0,\ j = C_2, \\ n - e_{ij} + e_{i,j-1} & n_{ij}\mu_2 \quad \text{if } j > 0. \end{cases} \tag{2.2}$$

We denote by $\pi_R(n)$, $n \in S$ the equilibrium distribution under a given policy R.

The state space grows rapidly as the capacities C_1, C_2, and the number of queues increase. The system optimal policy can be found using the theory of Markov decision processes, but apart from some special cases (e.g., when $C_1 = C_2 = 1$), the exactly optimal policy will, in general, not only require considerable computational effort to calculate, but also, just as importantly, substantial effort to implement. In Section 3, we therefore limit ourselves to state-dependent policies that are relatively easy both to analyse and to implement (although note that the asymptotic result given below gives optimality over all state-dependent policies). The system optimal policy minimizes over all policies R the expected cost per queue per unit time, which is given by

$$\lambda \sum_{n:n\in S} \pi_R(n) r_n(b) + \mu_1 \sum_{n:n\in S} \pi_R(n) \sum_i n_{i,C_2}/K. \tag{2.3}$$

The user optimal policy, however, is one that chooses the route that will give the lowest expected cost due to loss for an arrival. This can be calculated explicitly (see [14]), and the details are not given here, although the calculations are done for the examples in Section 3.

In the numerical examples below, in addition to giving exact results for the system with a small number of queues, found by calculating the equilibrium distribution numerically, we also give the asymptotic costs and policy, as the number of queues becomes large. The following results from [15], which are obtained using the methods of [18, 19], give the basis for obtaining the asymptotic results given below. In the following, instead of considering n as the state, we instead consider $\mathbf{x}^K = n/K$. Here, x_{ij}^K is the proportion of tandem queues in state (i, j).

Consider the sequence of networks indexed by K, the Kth network operating under any admissible acceptance and routing policy (an admissible policy must be nonanticipating). Let $Kw_{ij}^K(t)$ be the number of arrivals that have been accepted at a tandem queue in state (i, j) by time t, $0 \le i \le C_1$, $0 \le j \le C_2$ and let $Kw_b^K(t)$ be the number of arrivals that have been lost at entry (i.e., not accepted into the system) by time t. Then $\{(\mathbf{x}^K, \mathbf{w}^K)(\cdot)\}$ is relatively compact and the limit of any convergent sequence has the following properties.

(1) $\sum_{ij} x_{ij}(t) = 1$ for all $t \ge 0$.

(2) $x_{ij}(t) \ge 0$ for all $t \ge 0$, $0 \le i \le C_1$, $0 \le j \le C_2$.

(3) There exists $\mathbf{z}(\cdot)$ such that, almost surely, $z_{ij}(t), z_b(t) \ge 0$, and

$$\begin{aligned} x_{ij}(t) = x_{ij}(0) + \int_0^t \big(&z_{i-1,j}(s) I_{\{i\neq 0\}} - z_{ij}(s) I_{\{i\neq C_1\}} \\ &+ \mu_1 x_{i+1,j-1}(s) I_{\{i\neq C_1, j\neq 0\}} + \mu_1 x_{i+1,j}(s) I_{\{i\neq C_1, j=C_2\}} \\ &+ \mu_2 x_{i,j+1}(s) I_{\{j\neq C_2\}} - x_{ij}(s)\big(\mu_1 I_{\{i\neq 0\}} + \mu_2 I_{\{j\neq 0\}}\big)\big)\, ds, \end{aligned} \tag{2.4}$$

$$\lambda = \sum_{ij} z_{ij}(t) + z_b(t),$$

for all $t \ge 0$, $0 \le i \le C_1$, $0 \le j \le C_2$.

The equilibrium distribution is then a solution to the following system of equations:

$$\begin{aligned} &z_{ij} + x_{ij}\big(\mu_1 I_{\{i\neq 0\}} + \mu_2 I_{\{j\neq 0\}}\big) \\ &\quad = z_{i-1,j} I_{\{i\neq 0\}} + \mu_1 x_{i+1,j-1} I_{\{i\neq C_1, j\neq 0\}} \\ &\qquad + \mu_1 x_{i+1,j} I_{\{i\neq C_1, j=C_2\}} + \mu_2 x_{i,j+1} I_{\{j\neq C_2\}}, \quad 0 \le i \le C_1,\ 0 \le j \le C_2, \end{aligned} \tag{2.5}$$

$$\lambda = \sum_{ij} z_{ij} + z_b, \quad \sum_{ij} x_{ij} = 1, \quad x_{ij}, z_{ij}, z_b \ge 0.$$

These equations are balance equations for the asymptotic system. The z_{ij} here give the rate at which arrivals are entering queues in state (i, j) under the given policy, while z_b gives the rate at which arrivals are blocked.

In [15] these equations are constraints for the linear optimisation problem

$$\text{minimize } F(\mathbf{x}, z_b) = d_1 z_b + d_2 \mu_1 \sum_{i=1}^{C_1} x_{i,C_2}. \tag{2.6}$$

This can be solved to find the asymptotically optimal value of the objective function, and hence derive the asymptotically optimal control. However, the balance equations above can more generally be used to find the asymptotic costs for any routing policy of interest. In particular, we will give the asymptotic costs for the two main policies of interest, which are to accept all arrivals if possible, and to accept arrivals only if they can be routed to a tandem queue that has total occupancy less than C_2 (i.e., for the designated tandem queue, $n_1 + n_2 < C_2$).

3. Examples

Consider a system of parallel tandem queues with $C_1 = C_2 = 2$. If arrivals are accepted into the system, they will be routed to a queue in one of the states $(0,0), (1,0), (0,1), (1,1), (0,2), (1,2)$. In the state-dependent case, under user optimal routing, arrivals choose the queue that will minimize their own cost. Let $p_d(\mathbf{n})$ be the probability that an arrival joining a tandem queue in state $\mathbf{n} = (n_1, n_2)$ will reach its destination (the success probability). When $C_1 = C_2 = 2$,

$$\begin{aligned} p_d(1,1) &= 1 - \left(\frac{\mu_1}{\mu_1 + \mu_2}\right)^2, \\ p_d(1,2) &= 1 - \left(\frac{\mu_1}{\mu_1 + \mu_2}\right)^2 \left(1 + \frac{\mu_2}{\mu_1 + \mu_2}\right), \\ p_d(0,2) &= 1 - \frac{\mu_1}{\mu_1 + \mu_2} \end{aligned} \tag{3.1}$$

with, trivially, $p_d(0,0) = p_d(1,0) = p_d(0,1) = 1$ and $p_d(2,0) = p_d(2,1) = p_d(2,2) = 0$ (see [14] for details). We see immediately that $p_d(1,1) > p_d(1,2) > p_d(0,2)$. When $\mu_1 = \mu_2 = 1$, for instance, $p_d(1,1) = 3/4$, $p_d(1,2) = 5/8$ and $p_d(0,2) = 1/2$. Thus, somewhat paradoxically, an arrival wishing to minimize their own blocking probability at the second stage would prefer to join a queue in state $(1,2)$, rather than one in state $(0,2)$, even though the number of individuals in the former is greater. A queue in state $(1,1)$ is preferred to a queue in state $(0,2)$. In both cases the increased delay for the new arrival allows additional time for individuals ahead of them in the tandem to leave, thus reducing the blocking probability for the new arrival. Under user optimal routing, in general, arrivals may join queues that are in state (i,j) provided $d_1 > d_2(1 - p_d(i,j))$. Thus when $\mu_1 = \mu_2 = 1$, for instance, they may join queues in state $(1,1)$ if $d_2 < 4d_1$, in state $(1,2)$ if $d_2 < d_1 8/3$, and in state $(0,2)$ if $d_2 < 2d_1$.

The policy under user optimal routing is in strong contrast to the asymptotically system optimal policy, which is to accept arrivals if possible when $d_2 < d_1$, and otherwise to only accept arrivals into queues in one of the states $(0,0)$, $(1,0)$, or $(0,1)$, that is, a queue in such a state that the probability the arrival is lost at the second stage is 0 (see [15] for details). The results in that paper also yield the asymptotic average costs (to first order). Let

$$\lambda^* = \mu_1 \frac{1 + (\mu_1/\mu_2)}{1 + (\mu_1/\mu_2) + (\mu_1/\mu_2)^2}. \tag{3.2}$$

Then the asymptotic average costs of the two policies are as follows.

(1) If $\lambda < \lambda^*$, then all arrivals (to first order) can be routed to queues where there is no blocking, and the average cost is 0.

(2) For the policy that accepts all arrivals if possible the average cost is $d_2(\lambda - \lambda^*)$ when $\lambda^* < \lambda < \mu_1$, and $d_1(\lambda - \mu_1) + d_2(\mu_1 - \lambda^*)$ when $\lambda > \mu_1$.

(3) For the policy that only accepts arrivals into queues with occupancy less than $C_1 + C_2$, the average cost is $d_1(\lambda - \lambda^*)$ for $\lambda > \lambda^*$.

For a finite number of queues, as already observed, state-dependent optimal policies can be found using the theory of Markov decision processes, but are complex. Instead we consider a number of policies intermediate between the two asymptotically optimal ones. The costs of these are calculated numerically, and although closed form expressions can be given, we do not do so here, since they are tedious and not at all illuminating. In the examples below, where $C_1 = C_2 = 2$, we consider the following policies. A policy here consists of a list of possible states for queues into which an arrival can be accepted, listed in order of preference with the most preferred first. The policies compared below are

(1) $(0,0), (0,1), (1,0)$,

(2) $(0,0), (0,1), (1,0), (1,1)$,

(3) $(0,0), (0,1), (1,0), (0,2)$,

(4) $(0,0), (0,1), (1,0), (1,1), (1,2)$,

(5) $(0,0), (0,1), (1,0), (1,1), (0,2)$,

(6) $(0,0), (0,1), (1,0), (1,1), (1,2), (0,2)$.

For instance, policy (2) is to send an arrival to a queue in state $(0,0)$ if possible, otherwise to a queue in state $(0,1)$, otherwise to a queue in state $(1,0)$, and finally, if there is no queue in any of these three states, to a queue in state $(1,1)$. If there are no queues in any of these four states, then the arrival is lost. Some candidate policies have been omitted from the list. The policy $(0,0), (0,1), (1,0), (1,2)$ has the same average cost as policy (1), since the state $(1,2)$ for a single queue is transient under this policy (to see this, observe that to reach the state $(1,2)$ from any of the other three states included in this policy, it needs to pass through a state with $n_1 + n_2 = 2$, but no such state is included in this policy). Also, in policies (4), (5), and (6) we have assumed that queues in state $(1,1)$ are preferred to queues in state $(0,2)$.

The first example has $C_1 = C_2 = 2$, with $\mu_1 = \mu_2 = 1$. In Figure 3.1 we plot the expected cost per unit time for each of the six policies when $d_1 = 1$ and $d_2 = 3$ for a system of four queues. We note that policy (3) is the user optimal policy in this case, although the system optimal policy is to only accept arrivals into tandem queues that have occupancy no more than 1. For comparison purposes, the asymptotic cost as the number of queues, $K \to \infty$, under the asymptotically optimal policy (policy (1)) is also given. The expected cost is lowest for policy (1), and highest for policy (6), with policy (2) having lower cost than policy (4), which has lower cost again than policies (3) and (5). Thus, in accordance with

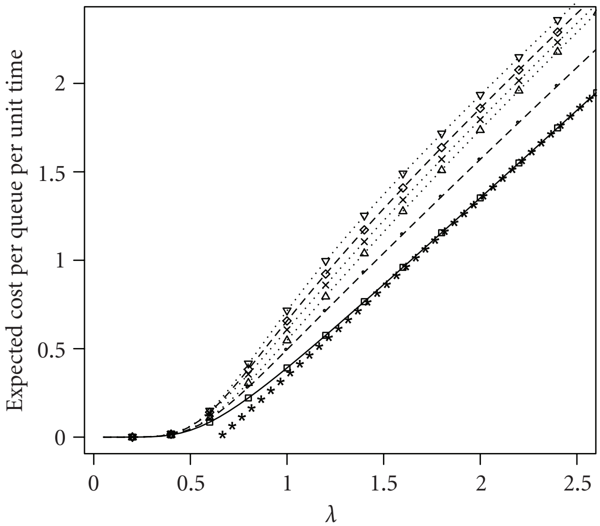

FIGURE 3.1. Expected cost per queue per unit time, $C_1 = C_2 = 2$, $\mu_1 = \mu_2 = 1$, $d_1 = 1$, $d_2 = 3$. Plotting symbols for each policy are 1 □, 2 ∘, 3 ×, 4 △, 5 ⋄ 6 ▽, asymptotically optimal policy ⋆. Four tandem queues.

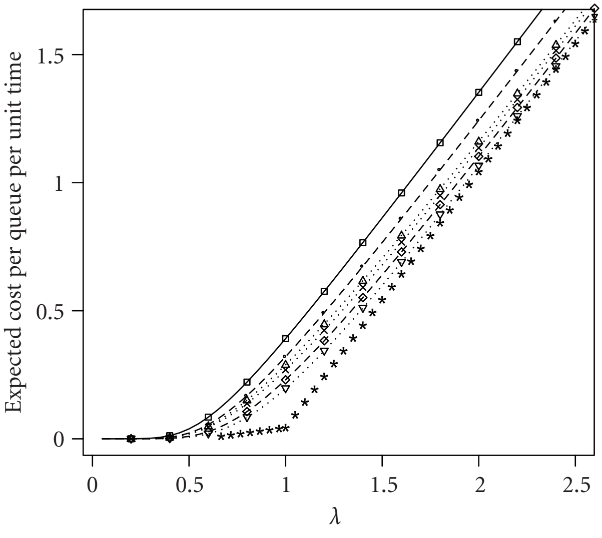

FIGURE 3.2. Expected cost per queue per unit time, $C_1 = C_2 = 2$, $\mu_1 = \mu_2 = 1$, $d_1 = 1$, $d_2 = 0.1$. Plotting symbols for each policy are 1 □, 2 ∘, 3 ×, 4 △, 5 ⋄ 6 ▽, asymptotically optimal policy ⋆. Four tandem queues.

the user optimal policy, sending arrivals to queues in state (0,2) also gives the highest average cost. However, when $d_2 < d_1$, this is largely reversed. Figure 3.2 gives a similar plot, but now with $d_2 = 0.1$, and we see that the policies are reversed, with policy (6) having the lowest cost, and policy (1) the highest.

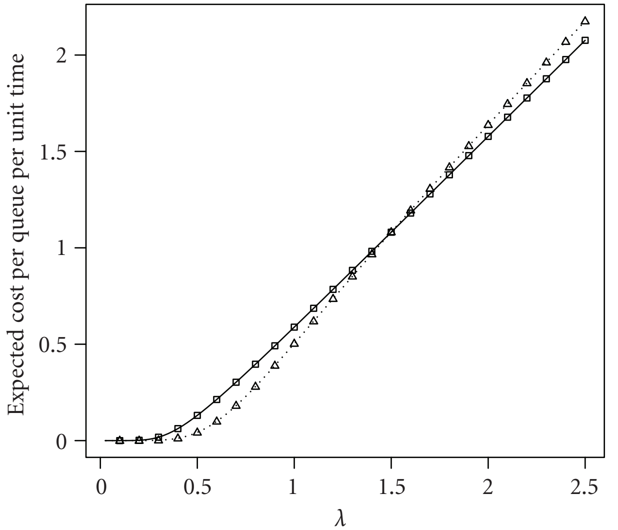

FIGURE 3.3. Expected cost per queue per unit time, $C_1 = C_2 = 2$, $\mu_1 = 1$, $d_1 = 1$, $d_2 = 2.5$ under user optimal policies when $\mu_2 = 0.5, 1.0$. Plotting symbols: $\mu_2 = 0.5$ $\square$, $\mu_2 = 1.0$ $\triangle$. Four queues.

Finally, Figure 3.3 plots the expected cost under the user optimal policy when $d_1 = 1$ and $d_2 = 2.5$ for $C_1 = C_2 = 2$, $\mu_1 = 1$, and $\mu_2 = 0.5, 1.0$. For both values of μ_2, the system optimal policy is policy (1). When $\mu_2 = 0.5$, the user optimal policy coincides with the system optimal policy, however, when $\mu_2 = 1.0$, policy (4) is user optimal since the expected cost of using queues in states $(1,1)$ and $(1,2)$ is less than d_1 in this case. We see from the plot that if arrivals follow user optimal policies, increasing the service rate at stage 2 of the tandem queues gives a higher expected cost overall for high values of λ.

These examples have all had unrealistically small capacities. This has been because the state space grows rapidly with C_1, C_2, and K. In all cases above, the equilibrium distribution, when the number of queues is finite, has been calculated explicitly to obtain the expected costs (rather than estimating from simulation). However, we conjecture that the finding in this case will carry over to larger capacities, that is, for sufficiently high arrival rates, as μ_2 increases, the expected cost may also increase, when d_2 is greater than d_1 (note that $d_2 > d_1$ is a reasonable scenario for a system where there may be a greater cost attached to losing an individual on whom some service has already been expended).

4. Conclusions

The numerical examples of the previous section have shown that permitting otherwise indistinguishable arrivals to use queues in certain states may lead to greater expected costs, when arrivals attempt to minimize their own costs due to loss. The difference here between the expected cost under user optimal and system optimal policies can be substantial. Furthermore, increasing the service rate, as in the classical paradoxes, may lead to worse overall performance, if user optimal policies are permitted. We have given a numerical example where increasing the service rate at the second stage of the tandem leads to increased expected cost.

Acknowledgment

This paper was completed while the author was visiting Lund University on a Solander Fellowship.

References

[1] D. Braess, "Über ein Paradoxon aus der Verkehrsplanung," *Unternehmensforschung*, vol. 12, no. 1, pp. 258–268, 1968.

[2] D. Braess, A. Nagurney, and T. Wakolbinger, "On a paradox of traffic planning," *Transportation Science*, vol. 39, no. 4, pp. 446–450, 2005, English language translation of [1].

[3] http://homepage.ruhr-uni-bochum.de/Dietrich.Braess/#paradox.

[4] R. Arnott and K. Small, "The economics of traffic congestion," *American Scientist*, vol. 82, no. 5, pp. 446–455, 1994.

[5] A. Downs, "The law of peak-hour expressway congestion," *Traffic Quarterly*, vol. 16, no. 3, pp. 393–409, 1962.

[6] R. Steinberg and W. I. Zangwill, "The prevalence of Braess' paradox," *Transportation Science*, vol. 17, no. 3, pp. 301–318, 1983.

[7] J. M. Thomson, *Great Cities and Their Traffic*, Gollancz, London, UK, Peregrine edition, 1977.

[8] J. G. Wardrop, "Some theoretical aspects of road traffic research," *ICE Proceedings, Engineering Divisions*, vol. 1, no. 3, pp. 325–362, 1952.

[9] E. Altman, R. El Azouzi, and V. Abramov, "Non-cooperative routing in loss networks," *Performance Evaluation*, vol. 49, no. 1–4, pp. 257–272, 2002.

[10] N. G. Bean, F. P. Kelly, and P. G. Taylor, "Braess's paradox in a loss network," *Journal of Applied Probability*, vol. 34, no. 1, pp. 155–159, 1997.

[11] F. P. Kelly, "Network routing," *Philosophical Transactions of the Royal Society of London. Series A*, vol. 337, no. 1647, pp. 343–367, 1991.

[12] A. Orda, R. Rom, and N. Shimkin, "Competitive routing in multiuser communication networks," *IEEE/ACM Transactions on Networking*, vol. 1, no. 5, pp. 510–521, 1993.

[13] T. Roughgarden and É. Tardos, "How bad is selfish routing?" *Journal of the ACM*, vol. 49, no. 2, pp. 236–259, 2002.

[14] S. Spicer and I. Ziedins, "User-optimal state-dependent routeing in parallel tandem queues with loss," *Journal of Applied Probability*, vol. 43, no. 1, pp. 274–281, 2006.

[15] R. Sheu and I. Ziedins, "Asymptotically optimal control of parallel tandem queues with loss," The University of Auckland, preprint, 2007.

[16] B. Calvert, "The Downs-Thomson effect in a Markov process," *Probability in the Engineering and Informational Sciences*, vol. 11, no. 3, pp. 327–340, 1997.

[17] B. Calvert, W. Solomon, and I. Ziedins, "Braess's paradox in a queueing network with state-dependent routing," *Journal of Applied Probability*, vol. 34, no. 1, pp. 134–154, 1997.

[18] P. J. Hunt and C. N. Laws, "Asymptotically optimal loss network control," *Mathematics of Operations Research*, vol. 18, no. 4, pp. 880–900, 1993.

[19] C. N. Laws and Y. C. Teh, "Alternative routeing in fully connected queueing networks," *Advances in Applied Probability*, vol. 32, no. 4, pp. 962–982, 2000.

Ilze Ziedins: Department of Statistics, The University of Auckland, Private Bag 92019,
Auckland 1142, New Zealand
Email address: ilze@stat.auckland.ac.nz