DYNAMIC ROUTING OF CUSTOMERS WITH GENERAL DELAY COSTS

Nilay T. Argon*, Li Ding**, Kevin D. Glazebrook***, Serhan Ziya*

*Department of Statistics and Operations Research,
University of North Carolina at Chapel Hill,
Smith Building, CB# 3260, Chapel Hill, NC 27599, U.S.A.
e-mails: nilay@unc.edu, ziya@unc.edu

**School of Management, University of Edinburgh,
Edinburgh, EH8 9YJ, U.K.
e-mail: l.ding@sms.ed.ac.uk

***Department of Mathematics and Statistics
Lancaster University Management School, Bailrigg, Lancaster LA1 4YX, U.K.
e-mail: k.glazebrook@lancaster.ac.uk

Abstract

We consider a network of parallel service stations each modelled as a single server queue. Each station serves its own dedicated customers as well as generic customers who are routed from a central controller. We suppose that the cost incurred by a customer is an increasing function of her time spent in system. In a significant advance on most previous work, we do not require waiting costs to be convex, still less linear. With the objective of minimizing the long-run average cost, we develop heuristic routing policies and demonstrate their superior performance in an extensive numerical study.

Keywords: State-dependent routing; index policy; heuristic policy; stochastic dynamic programming; Markov decision processes.

1 Introduction

This paper is concerned with the dynamic routing of incoming customers to one of several parallel service stations, each of which has its own queue with an infinite buffer space. Our model develops those found most frequently in the literature in two key respects, both of which are critical for applications. First, in our setups there are two customer types which inhabit the system: *dedicated* and *generic*. Dedicated customers must be served at a designated station and routing decisions only concern the remaining generic traffic, which we shall assume can be served at any station. Plainly such a dichotomizing of the traffic is important in settings where specialist skills are required on occasion and are expensive to provide. Note that stations are not necessarily separated geographically since customers requiring different skills will naturally form separate queues. Examples include call centers [7] and emergency response systems [32]. Such models are also relevant to dynamic routing within a grid environment [1] and to load balancing in broadcast communication networks [37].

Both customer types incur costs which depend upon the time spent waiting in the system. Our objective is to develop routing policies which will minimize the long-run average total cost or which will come close to doing so. This brings us to the second key feature of our model, namely that we shall impose only very minor conditions upon the structure for waiting costs. Most previous work which has considered the dynamic service and routing control of queueing systems has imposed requirements that waiting costs be linear in the time spent in system. In some exceptions, this has been generalized to an assumption that costs be convex non-decreasing (see, for example, [14], [35], [40], and [45]). Van Mieghem [45] discusses the limited appeal of linear waiting costs, but we would argue that an assumption of

convexity may also be unrealistic in many settings. Convex waiting costs are unable to take adequate account of such commonplace features as delivery time/service level guarantees (see, for example, [10], [20], and [24]) nor in an emergency response setting of the absolute imperative of treating patients with critical trauma injuries quickly (the "golden hour" rule, see, for example, [36]).

State-dependent routing problems have received much attention in the literature. The only general settings in which it has proved possible to elucidate simple structure for optimal policies are those involving homogeneous stations. Here "Shortest Delay Routing" (SDR) has been shown to be optimal under a range of conditions. See, for example, Weber [47], Winston [52], Johri [23], Hordijk and Koole [21], Menich and Serfozo [34], Sparaggis, Towsley, and Cassandras [39], Towsley, Sparaggis, and Cassandras [44], and Koole, Sparaggis, and Towsley [27]. However, Whitt [49] shows that SDR may not be optimal even for very simple cases. Also see Foschini and Salz [12], Houck [22], Wein [48], Laws [31], Kelly and Laws [25], Foley and McDonald [11], Teh and Ward [41], and Tezcan [42] for more on the performance of SDR and analytical results for SDR under various asymptotic regimes. Recently, there also appeared a significant amount of work on routing problems where customers wait in a single line and are routed to one of the servers as servers become available. The majority of this work is motivated by the routing problem in call centers. See, for example, Van Mieghem [45], Gans and Zhou [13], Harrison and Zeevi [19], Mandelbaum and Stolyar [33], Armony and Maglaras [4, 5], Wallace and Whitt [46], Armony [3], and Bassamboo, Harrison, and Zeevi [6].

In a recent study, Stolyar [40] considers a similar routing problem where customers are routed to one of the multiple parallel stations as they arrive to the system. Stolyar's model is

more general than ours in that customers belong to different classes and service rates depend on both the customer's class identity and the station serving the customer. However, the author imposes some restrictive conditions on the waiting cost structures. More specifically, he assumes that the waiting cost function for each queue is continuous, strictly increasing and convex. Stolyar proves that a routing policy, which he calls the *MinDrift* rule, is asymptotically optimal in the heavy traffic regime under a complete resource pooling condition. In Section 6, we test the performance of the Mindrift rule along with the performance of our heuristics under various traffic regimes and two different waiting cost functions.

For complex routing problems which do not assume station homogeneity, full characterization of optimal policies based, for example, on direct application of dynamic programming (DP) methodology is unrealistic. The focus for such systems has mostly been on the development of effective routing heuristics, many of which are based on calibrations of the constituent service stations which utilize queue length information. A simple and standard proposal is to route each generic customer in an individually optimal way to whichever station has the smallest expected cost for that individual. In Section 6, this is referred to as the greedy heuristic (GH). Two approaches to heuristic development which have proved effective are those based on DP policy improvement (see, for example, Krishnan and Ott [30], Krishnan [28, 29], Whittle [51], Sassen, Tijms, and Nobel [38], Ansell, Glazebrook, and Kirkbride [1], Bhulai and Koole [9], and Bhulai [8]) and on Lagrangian relaxations (see, for example, Niño-Mora [35] and Whittle [50, 51]). The second of these has appeared highly promising, but its effective deployment has been severely inhibited by a prior requirement to establish a structural property which Whittle [50] called *indexability*. While this property appears natural, it is little understood and is often extremely difficult to prove. There is a

developing literature centered around this very issue in a range of application areas of which routing is but one (see, for example, Glazebrook, Niño-Mora, and Ansell [15], Ansell, Glazebrook, Niño-Mora, and O'Keeffe [2], Glazebrook, Mitchell, and Ansell [16], and Glazebrook, Kirkbride, and Ouenniche [17]).

A major research achievement of this paper is the demonstration that we do indeed have indexability in our very general setting for both costs and traffic. In particular, we are able to show that the cost convexity assumptions of previous work are not needed for the parallel service system considered in this paper. Indexability is proved in Section 4 and the structure of the resulting routing heuristics is established. To provide a strong comparator for these Lagrangian heuristics, routing policies based on DP policy improvement are developed in Section 3. In Section 5, the precise nature of the station calibrations which our routing heuristics utilize is developed for some non-standard but practically important cost structures while Section 6 contains the results of an extensive investigation into heuristic performance.

2 Model description

We consider $N \geq 2$ parallel service stations each modelled as a single server queue. For each queue n, service times of customers are independent and exponentially distributed with a finite mean $1/\mu_n$. Each queue n has its own "dedicated" customers that arrive according to a Poisson process with rate η_n. These dedicated customers cannot be routed to any other station. There are also "generic" customers who arrive as a Poisson stream with rate λ and who can be served at any station. The routing decisions concern these customers only. Our analysis will still go through if the system controller can also choose to reject generic customers by paying a fixed charge per rejected customer. However, to

keep the presentation simpler, we assume that each arriving customer should be admitted to the system. The routing decision for a generic customer is made upon her arrival and this decision is irreversible. We assume that each queue operates under a First-Come-First-Served (FCFS) discipline, and hence no distinction is made between the dedicated customers and generic customers once admitted to the system. We introduce parameters

$$\theta_n = \lambda + \eta_n,\ \beta_n = \theta_n/\mu_n,\ \alpha_n = \eta_n/\mu_n,$$

and we assume that $\alpha_n < 1$ for all n and $\lambda + \sum_{n=1}^{N} \eta_n < \sum_{n=1}^{N} \mu_n$ so that there exist routing policies under which the system is stable.

Each customer incurs a cost that is a function of the time the customer spends in the system. Write ξ_{nk} for $n \in \{1, 2, \ldots, N\}$ and $k \in \{1, 2, \ldots\}$ for the total system time for the kth customer admitted to queue n. Also let $C_n(\cdot)$ denote the cost function for queue n, i.e., $C_n(t)$ is the cost incurred by a customer who is given service in station n and whose total system time is t. Our objective is to minimize the expected long-run average total cost for this system, which can be written as

$$\limsup_{t\to\infty} \frac{\mathbb{E}\left[\sum_{n=1}^{N} \sum_{k=1}^{M_n(t)} C_n(\xi_{nk})\right]}{t},$$

where $M_n(t)$ is the total number of customers who have been served in station n by time t.

Let $c_n(i)$ denote the expected cost that will be incurred by a customer who joins queue n when there are i customers already waiting (including the customer in service) at station n. If we deem that (expected) costs are incurred as the customers join, this problem can be formulated as a semi-Markov decision process. Decision epochs are the times when generic customers arrive and the system state can be described by an N-dimensional vector $\boldsymbol{x} = (x_1, x_2, \ldots, x_N)$, where x_n denotes the number of customers waiting in station n

(including the customer in service). Action space is $\mathcal{A} = \{a_1, a_2, \ldots, a_N\}$, where a_n is the action of sending a customer to queue n. Once the problem is formulated, then it can in principle be solved to arbitrary accuracy by applying a (finite) truncation of the state space followed by deployment of standard DP methodologies. However, such an approach is computationally intractable for systems of realistic size. It is natural that interest should focus on the development of effective heuristics for routing. This is the subject of the following two sections.

3 Heuristic 1: Single-step policy improvement method

We develop the *single-step policy improvement heuristic* by applying a single step of the policy improvement algorithm to an optimal static Bernoulli routing policy. A static Bernoulli routing policy simply routes each incoming customer to queue n with some probability p_n independently of other customers and the system state. We shall develop the heuristic in two stages. First we shall determine an optimal static policy. This will then be followed by a single DP policy improvement step.

An optimal static policy:

A static routing policy routes each incoming generic customer to one of the service stations according to some probability distribution $\boldsymbol{p} = (p_1, p_2, \ldots, p_N)$ independently of other customers and the system state, where $p_n \geq 0$ for $n = 1, 2, \ldots, N$, and $\sum_{n=1}^{N} p_n = 1$. Let the set $\mathbb{P}$ be defined as

$$\mathbb{P} = \{\boldsymbol{p} : p_n \geq 0,\ \eta_n + \lambda p_n < \mu_n,\ 1 \leq n \leq N,\ \text{and}\ \sum_{n=1}^{N} p_n = 1\}.$$

We have $\mathbb{P} \neq \emptyset$ since we assume that $\lambda + \sum_n \eta_n < \sum_n \mu_n$.

Under any static policy $\boldsymbol{p} \in \mathbb{P}$, each service station n is an independent M/M/1 queue with arrival rate $\Lambda_n = \eta_n + \lambda p_n$ and service rate μ_n. Then, using standard results, the expected long-run average cost $TC(\boldsymbol{p})$, can be written as:

$$TC(\boldsymbol{p}) = \sum_{n=1}^{N} (\eta_n + \lambda p_n) \sum_{i=0}^{\infty} c_n(i)(1-\rho_n)\rho_n^i, \tag{1}$$

where $\rho_n = (\eta_n + \lambda p_n)/\mu_n$.

In order to proceed, we need the following assumption:

Assumption 3.1 *There exists a* $\boldsymbol{p} \in \mathbb{P}$ *such that* $\sum_{n=1}^{N} \sum_{i=0}^{\infty} c_n(i)\rho_n^i < \infty$ *where* $\rho_n = (\eta_n + \lambda p_n)/\mu_n$.

Assumption 3.1 is certainly satisfied if the $c_n(\cdot)$'s are all polynomially bounded.

Clearly, we can restrict ourselves to the set $\mathbb{P}$ to find the optimal static policy. Hence, an optimal static routing policy is any vector $\boldsymbol{p}^\star \in \mathbb{P}$ that minimizes (1), i.e.

$$\boldsymbol{p}^\star \in \left\{ \bar{\boldsymbol{p}} \in \mathbb{P} : TC(\bar{\boldsymbol{p}}) \leq TC(\boldsymbol{p}) \ \forall \boldsymbol{p} \in \mathbb{P} \right\}. \tag{2}$$

Policy improvement step:

We now apply a policy improvement step to an optimal static policy. Using the approach of Section 3.6 of Tijms [43], it can be easily shown that the policy improvement step yields an index policy. To see that, we first define $\Delta(\boldsymbol{x}, a_n)$ to be the difference in total expected costs that would be caused by sending the first incoming customer to queue n and then following an optimal static policy rather than using an optimal static policy throughout when the system is in state $\boldsymbol{x}$. Since each queue operates independently under a static policy, this

difference can be expressed as

$$\begin{aligned}\Delta(\boldsymbol{x}, a_n) &= \sum_{k=1}^{N} p_k^\star [D_n(x_n) - D_k(x_k)] \\ &= D_n(x_n) - \sum_{k=1}^{N} p_k^\star D_k(x_k) \end{aligned} \tag{3}$$

where $D_n(i)$ denotes the additional long-run cost of admitting an extra customer to queue n operating under the optimal static routing policy when i customers are already present. In (3), the summation term does not depend on the action a_n taken. Therefore, the action a_n that minimizes $\Delta(\boldsymbol{x}, a_n)$ can simply be found by choosing n to minimize $D_n(x_n)$. In other words, the policy improvement step gives us an index policy that routes an arriving generic customer to queue $n^\star$ such that

$$D_{n^\star}(x_{n^\star}) = \min_{n \in \{1,2,\ldots,N\}} \{D_n(x_n)\},$$

when the system state is $\boldsymbol{x} = (x_1, x_2, \ldots, x_N)$.

In order to describe the policy, we require an expression for $D_n(\cdot)$. It will simplify matters if we now focus on a single station and drop the identifier n. Let $\Lambda = \eta + \lambda p^\star$ denote the total arrival rate to this station (the total arrival rate of dedicated and generic customers assigned to this queue under the optimal static policy). Define μ to be the service rate, $\rho = \Lambda/\mu$ to be the traffic intensity, where $\rho < 1$, and $c(i)$ to be the expected cost for a customer who sees i customers upon arrival.

Utilizing standard arguments which exploit the fact that the system regenerates upon every return to the empty state, it can be shown that

$$D(i) = c(i) + (K_{i+1} - gT_{i+1}) - (K_i - gT_i), \text{ for } i \geq 0, \tag{4}$$

where K_i is the expected total cost incurred until the first time there is no customer in the system starting with i customers, T_i is the expected time until the first time the station is empty starting with i customers, and g is the long-run average cost for the queue under the static policy. Note that $K_0 = 0$ and $T_0 = 0$.

To make sense of (4), note that $K_i - gT_i$ is the relative cost (or bias) of starting the queue with i customers rather than none. Hence the difference between the last two terms in (4) is simply the difference in total expected long run costs between starting the queue with $i+1$ rather than i customers. However, this difference ignores the cost to be incurred by the customer to be admitted. Adding the expected cost for that customer, $c(i)$, to this difference gives us the additional cost of admitting one more customer, which is $D(i)$. Now, we can rewrite (4) as follows:

$$D(i) = c(i) + K_{i+1} - K_i + g(T_i - T_{i+1}), \text{ for } i \geq 0. \tag{5}$$

To find an expression for K_i, we first note that

$$K_i = \frac{\Lambda}{\Lambda + \mu}(c(i) + K_{i+1}) + \frac{\mu}{\Lambda + \mu}K_{i-1} \text{ for } i \geq 1.$$

Rearranging terms, we obtain

$$k_{i+1} = \rho^{-1}k_i - c(i),$$

where $k_i = K_i - K_{i-1}$. Using this relationship recursively, we get

$$k_i = \frac{k_1 - \sum_{j=1}^{i-1} c(j)\rho^j}{\rho^{i-1}}, \text{ for } i \geq 2. \tag{6}$$

Note that $k_1 = K_1 - K_0 = K_1$. Hence, to give an explicit expression for k_i, it remains to determine K_1.

In order to derive K_1, consider the M/M/1 queue under study. This queue evolves in *cycles*, i.e., independent and identically distributed (i.i.d.) renewal intervals, which start when a customer arrives to an empty system. Each cycle consists of a *busy period* (there are customers in the system and the server is working) followed by an *idle period* (no customers in the system, and therefore the server is idle). Let γ_m denote the total cost incurred over cycle m. Clearly, γ_m is i.i.d. for all m. Let $\Gamma = E[\gamma_1]$.

Recall that g is the long-run average cost for this system. Then, using the renewal reward theorem and the fact that expected length of a busy period is $1/(\mu - \Lambda)$ (see, e.g., Section 2.11 in [18]), we can show the following:

$$\Gamma = \frac{g}{\Lambda(1-\rho)}.$$

Then,

$$K_1 = \Gamma - c(0),$$

since K_1 does not include expected costs for the customer initially in the system. Then, using (6) and the fact that $g = \Lambda \sum_{j=0}^{\infty} c(j)(1-\rho)\rho^j$, after a few algebraic manipulations, we can establish that

$$k_i = \frac{\sum_{j=i}^{\infty} c(j)\rho^j}{\rho^{i-1}}, \text{ for } i \geq 1. \tag{7}$$

Finally, since $T_{i+1} - T_i$ is equal to the expected length of a "busy period," we have $T_{i+1} - T_i = 1/(\mu - \Lambda)$. Then, using (7), we deduce that

$$D(i) = c(i) + k_{i+1} - g\frac{1}{\mu - \Lambda} = \sum_{j=0}^{\infty}(c(j+i) - \rho c(j))\rho^j, \tag{8}$$

for $i \geq 0$. (An alternative way of obtaining (8) from (4) is by solving for the relative values $K_i - gT_i$ directly, utilizing the form of solution to classes of second-order difference equations given as Corollary 3.3 in Bhulai [8].)

Thus, we have proved the following theorem.

Theorem 3.1 ***(Policy Improvement Heuristic)*** *The heuristic dynamic routing policy that is developed by applying a policy improvement step to an optimal static policy* $\boldsymbol{p}^\star = (p_1^\star, p_2^\star, \ldots, p_N^\star)$ *operates as follows: Upon arrival of a generic customer, if the system state is* $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$*, then route the customer to any station* $n^\star$ *for which*

$$D_{n^\star}(x_{n^\star}) = \min_{k \in \{1,2,\ldots,N\}} \{D_k(x_k)\},$$

where

$$D_k(x_k) = \sum_{j=0}^{\infty} (c_k(j + x_k) - \rho_k c_k(j))\rho_k^j \tag{9}$$

and $\rho_k = (\eta_k + \lambda p_k^\star)/\mu_k$.

The index function (9) is insightful. To see that, we can use (7) and (8) to rewrite (9) as

$$D_k(x_k) = c_k(x_k) + \frac{1}{\mu_k - \Lambda_k} \left(\frac{\Lambda_k}{\rho_k^{x_k+1}} \sum_{i=x_k+1}^{\infty} c_k(i)\rho_k^i(1-\rho_k) - \Lambda_k \sum_{i=0}^{\infty} c_k(i)\rho_k^i(1-\rho_k) \right). \tag{10}$$

The first term on the right-hand side of (10), $c_k(x_k)$, is the expected cost for a customer who will be sent to station k, and might be interpreted as the *internal cost* for the customer. However, the customer also incurs costs for the other incoming customers who will be routed to the same station and the second term of (10) takes this cost into account. The term $1/(\mu_k - \Lambda_k)$ is the expected length of a "busy period" caused by the customer under the optimal static policy. In other words, it is the expected time until the first occasion for which the number of customers in the station is x_k again. The first term in the parentheses in (10) is the long-run average cost for the station when the average is taken over the times when there are more than x_k customers in the station under the optimal static policy, while the second term in the parentheses is simply the long-run average cost for station k. Their difference can be interpreted as the per unit time cost of keeping the queue above level x_k. Multiplying this difference by the expected length of a busy period we find the *total cost* the admitted customer imposes on the system. Hence, in a sense, the expression that follows $c_k(x_k)$ in (10) is the *external cost* of admitting the customer to station k.

4 Heuristic 2: Lagrangian relaxation method

In development of an alternative collection of routing heuristics, we firstly propose a relaxation of the optimal routing problem of Section 2 in which the class of policies is expanded to those which route each incoming generic customer to *any number of service stations*, subject to an overall admission rate of λ. Expressed differently, arriving generic customers are routed to one service station *on average*. Let $\mathcal{U}$ be the class of stationary policies which upon each generic arrival, increase the queue length at any number of chosen stations by one and let $u \in \mathcal{U}$ be a class member. Let $\widetilde{C}_n(u)$ be the time average cost rate incurred at station n and $\widetilde{A}_n(u) = \lambda - \widetilde{R}_n(u)$ be the generic admission rate (with $\widetilde{R}_n(u)$ denoting the corresponding rejection rate) at station n under control $u \in \mathcal{U}$.

A Lagrangian approach to the above relaxed problem yields

$$\widehat{C}(W) = \min_{u \in \mathcal{U}} \sum_{n=1}^{N} \{\widetilde{C}_n(u) + W\widetilde{R}_n(u)\} - W(N-1)\lambda, \tag{11}$$

where in (11) W is a Lagrange multiplier which has an economic interpretation as a charge for rejecting a single generic customer. Plainly, $\widehat{C}(W) \leq C^{opt}$ for all $W \in \mathbb{R}$ where C^{opt} is the value of the original minimization problem. Further (11) admits a stationwise decomposition, expressed by

$$\widehat{C}(W) = \sum_{n=1}^{N} \widehat{C}_n(W) - W(N-1)\lambda,$$

where

$$\widehat{C}_n(W) = \min_{u \in \mathcal{U}_n} \{\widetilde{C}_n(u) + W\widetilde{R}_n(u)\}, \text{ for } n = 1, 2, \ldots, N. \tag{12}$$

The optimization problem in (12) relates to station n alone with $\mathcal{U}_n$ the class of stationary controls for determining whether to admit (action a) or reject (action r) each generic arrival,

the goal of optimization being the minimization of an aggregate of customer waiting cost ($\widetilde{C}_n(u)$) and rejection charges ($W\widetilde{R}_n(u)$). Call this station n problem $P_n(W)$.

We now focus exclusively on station n and the corresponding optimization problems $P_n(W)$, $W \in \mathbb{R}^+$. It will ease notation if, until further notice, we drop the station identifier n from the notation and refer to $P(W)$, $W \in \mathbb{R}^+$. In order to describe solutions to the $P(W)$, we develop the sequence $\{W(i), i \in \mathbb{N}\}$ of reals (where $\mathbb{N}$ is the set of non-negative integers) and the corresponding sequence $\{u(i),\ i \in \mathbb{N}\}$ of monotone admission control policies as follows:

$$W(i) = \sum_{j=0}^{i} \beta^j \big\{\alpha(1-\alpha) \sum_{k=i+1}^{\infty} c(k)\alpha^{k-i-1} + c(i)(1-\alpha) - c(j)\beta\big\} + c(i)\beta^{i+1},\ i \in \mathbb{N}, \quad (13)$$

and

$$u(i) \text{ chooses } r \iff \text{ queue length at arrival epoch } \geq i,\ i \in \mathbb{N}. \quad (14)$$

It will facilitate the analysis if we make the following technical assumptions:

Assumption 4.1 *The sequence of expected costs* $\{c(i),\ i \in \mathbb{N}\}$ *is non-decreasing such that*

(i) $c(i) \to \infty$, $i \to \infty$;

(ii) $\sum_{i=0}^{\infty} c(i)\alpha^i < \infty$.

Lemma 4.1 is a straightforward consequence of Assumption 4.1.

Lemma 4.1 $\{W(i),\ i \in \mathbb{N}\}$ *is an increasing sequence of positive reals such that* $W(i) \to \infty$, $i \to \infty$.

The following key result describes optimal policies for $P(W)$ for all $W \geq 0$. We extend (13) by adopting the notational convention $W(-1) = 0$. The proof of Theorem 4.1 will utilize two following lemmas.

Theorem 4.1 *Policy* $u(i)$ *is optimal for* $P(W)$ *for* $W \in [W(i-1), W(i))$, $i \in \mathbb{N}$.

Fix $W \in [W(i-1), W(i))$ for some $i \in \mathbb{N}$ and choose $M \in \mathbb{N}$ with $M \geq i+1$. We develop a finite state approximation to $P(W)$ by requiring that generic customers be rejected at queue lengths M and above. Let $P(W, M)$ be the optimization problem which seeks an optimal control under this constraint. We shall establish Theorem 4.1 by demonstrating that $u(i)$ is optimal for $P(W, M)$, $W \in [W(i-1), W(i))$, $M \geq i+1$, and then considering the limit $M \to \infty$.

Since the control $u(i)$ operates identically whether applied to $P(W)$ or to $P(W, M)$, for $i \leq M-1$, we may write $g(i)$ unambiguously for the *long-run average cost rate* and $w(i, j)$ for the *bias in state* j under the operation of policy $u(i)$, $M-1 \geq j \geq 0$.

Lemma 4.2 *Control $u(i)$ will be optimal for $P(W, M)$, $i \leq M-1$ if*

$$\begin{array}{rcl} c(j) + w(i, j+1) - w(i, j) & \leq & W, \ 0 \leq j \leq i-1, \ \textit{and} \\ c(j) + w(i, j+1) - w(i, j) & \geq & W, \ M-1 \geq j \geq i. \end{array} \tag{15}$$

Proof: Under action a (accept generic customers) in state $j \geq 1$, possible state transitions are to $j+1$ and $j-1$ with rates θ and μ, respectively. The corresponding rates under action r are η and μ. There are no service completions in state 0. Adopting a uniformization with common transition rate $\theta + \mu$, we develop the average cost optimality equations for $P(W, M)$ as

$$\begin{array}{rcl} g^* + (\theta+\mu) w^*(j) & = & \min\{\theta c(j) + \theta w^*(j+1) + \mu I(j \geq 1) w^*(j-1) + \mu I(j=0) w^*(0); \\ & & \eta c(j) + W(\theta - \eta) + \eta w^*(j+1) + (\theta - \eta) w^*(j) + \\ & & \mu I(j \geq 1) w^*(j-1) + \mu I(j=0) w^*(0)\}, \ \text{for } 0 \leq j \leq M-1, \end{array} \tag{16}$$

where g^* is the optimal long-run average cost, $w^*(j)$ is the bias under an optimal policy for initial state j, and the indicator function $I(A) = 1$ if A is true, and 0 otherwise.

Moreover, any stationary policy which takes actions to minimize the right side of (16)

for all $0 \leq j \leq M-1$ will be average cost optimal. Hence, policy $u(i)$ will be optimal for $P(W,M)$ if, for $j \leq i-1$,

$$\begin{aligned}
&\theta c(j) + \theta w(i,j+1) + \mu I(j \geq 1) w(i,j-1) + \mu I(j=0) w(i,0) \\
&\quad \leq \eta c(j) + W(\theta - \eta) + \eta w(i,j+1) + (\theta - \eta) w(i,j) + \mu I(j \geq 1) w(i,j-1) \\
&\qquad + \mu I(j=0) w(i,0)
\end{aligned} \tag{17}$$

with the reverse inequality for $M-1 \geq j \geq i$. The requirements in (15) follow simply. This concludes the proof. □

In order to utilize Lemma 4.2, we need the biases $w(i,j)$, $0 \leq j \leq M-1$. As in the calculations around (4) and (5) above, we deploy an argument based on the fact that the queue length process under $u(i)$ regenerates whenever the queue empties to assert that

$$w(i,j) = K(i,j) - g(i)T(i,j), \ 0 \leq j \leq M-1. \tag{18}$$

In (18), $K(i,j)$ is the total expected cost incurred (waiting costs and rejection charges) as the queue empties (for the first time) under policy $u(i)$ from an initial position in which j customers are present, and $T(i,j)$ is the corresponding expected time. From (18), it is easy to see that $w(i,0) = 0$. Evaluations of these quantities are given in Lemma 4.3.

Lemma 4.3

(i) For $i \in \mathbb{Z}^+$, we have that

(a) $T(i,j) = \sum_{k=0}^{j-1} \{(1-\beta^{i-1-k})(\mu-\theta)^{-1} + \beta^{i-1-k}(\mu-\eta)^{-1}\}$, $0 \leq j \leq i-1$;

(b) $T(i,j) = T(i,i-1) + (j-i+1)(\mu-\eta)^{-1}$, $i \leq j \leq M-1$;

(c) $K(i,j) = \sum_{k=0}^{j-1} \{W(\theta-\eta)\beta^{i-k}(\theta(1-\alpha))^{-1} + \sum_{l=k+1}^{i-1} c(l)\beta^{l-j} + \sum_{l=i}^{\infty} c(l)\beta^{i-k}\alpha^{l-i}\eta\theta^{-1}\}$, $0 \leq j \leq i-1$;

(d) $K(i,j) = K(i,i-1) + \sum_{k=0}^{j-i} \{W(\theta-\eta)(\mu-\eta)^{-1} + \sum_{l=k}^{\infty} c(l+i)\alpha^{l+1-k}\}$, $i \leq j \leq M-1$;

(e) $g(i) = K(i,1)\{\theta^{-1} + T(i,1)\}^{-1}$.

(ii) For the case $i = 0$, we have that

(a) $T(0,j) = j(\mu-\eta)^{-1}$, $0 \le j \le M-1$;
(b) $K(0,j) = jW(\theta-\eta)(\mu-\eta)^{-1} + \sum_{k=1}^{j}\sum_{l=k}^{\infty} c(l)\alpha^{l-k+1}$, $0 \le j \le M-1$;
(c) $g(0) = K(0,1)\{\eta^{-1} + T(0,1)\}^{-1}$.

Proof: Choose i and j such that $M-1 \ge j \ge i$. Under policy $u(i)$, the time taken for the queue length to drop from j to $j-1$ (for the first time) is stochastically identical to a busy period of an M/M/1 queue with arrival and service rates η and μ, respectively. It follows that

$$T(i,j) - T(i,j-1) = (\mu-\eta)^{-1}, \ i \le j \le M-1 \tag{19}$$

and Lemma 4.3 (i)(b) follows simply. Further if $i-1 \ge j \ge 1$, then by conditioning on the time of the first transition after time zero we infer that

$$(\theta+\mu)T(i,j) = 1 + \theta T(i,j+1) + \mu T(i,j-1), \ 1 \le j \le i-1. \tag{20}$$

Combining (20) with (19) and the fact that $T(i,0) = 0$ yields Lemma 4.3 (i)(a) simply.

Now choose i and j such that $M-1 \ge j \ge i$ and consider the system evolving from state j at time zero under $u(i)$, but where service is abandoned whenever the queue length drops to $j-1$. Hence, whenever the system enters $j-1$, it remains there for an exponentially distributed time with mean $1/\eta$ before returning to state j. An argument based on the fact that the system regenerates upon every entry into j yields the formula

$$\{K(i,j) - K(i,j-1) + W(\theta-\eta)\eta^{-1} + c(j-1)\}\big/\{(\mu-\eta)^{-1} + \eta^{-1}\} \tag{21}$$

for the overall cost rate for the system. But computation of the steady state distribution of the system state yields

$$W(\theta-\eta) + \sum_{k=j-1}^{\infty} \eta c(k)\alpha^{k+1-j}(1-\alpha) \tag{22}$$

as an alternative expression. Equating (21) and (22) yields

$$K(i,j) - K(i,j-1) = W(\theta-\eta)(\mu-\eta)^{-1} + \sum_{k=j}^{\infty} c(k)\alpha^{k-j+1},\ i \le j \le M-1, \tag{23}$$

from which Lemma 4.3 (i)(d) is easily deduced. Lemma 4.3 (i)(c) then follows via a recursion which modifies (20) suitably.

Lemma 4.3 (i)(e) is a trivial consequence of the fact that the system evolving under $u(i)$ regenerates upon every entry into state 1 and that the mean time between successive entries is $\theta^{-1} + T(i,1)$. This completes the proof of Lemma 4.3 (i). Part (ii) is dealt with similarly.

□

With Lemmas 4.2 and 4.3 in place, we now proceed to the proof of Theorem 4.1.

Proof of Theorem 4.1: Fix $i \in \mathbb{Z}^+$. We shall show that

$$c(j) + w(i,j+1) - w(i,j) \ge W,\ M-1 \ge j \ge i, \tag{24}$$

where $W \in [W(i-1), W(i))$. Other cases ($i \in \mathbb{Z}^+$, $j \le i-1$, and $i = 0$) are dealt with similarly.

Firstly, note that from (18), condition (24) becomes

$$c(j) + K(i,j+1) - K(i,j) - g(i)\{T(i,j+1) - T(i,j)\} \ge W,\ M-1 \ge j \ge i. \tag{25}$$

However, direct calculation based on (13) and Lemma 4.3 yields that, for any $W \in [W(i-1), W(i))$ we have that

$$g(i) \le g(j),\ M-1 \ge j \ge i. \tag{26}$$

Hence, since $T(i,j+1) - T(i,j) \ge 0$, in order to establish (25) it will be enough to show that

$$c(j) + K(i,j+1) - K(i,j) - g(j)\{T(i,j+1) - T(i,j)\} \ge W,\ M-1 \ge j \ge i. \tag{27}$$

Utilizing Lemma 4.3 (i) (b), (d), (e), the requirement in (27) becomes

$$\begin{aligned} c(j) + W(\theta-\eta)(\mu-\eta)^{-1} + \textstyle\sum_{k=j+1}^{\infty} c(k)\alpha^{k-j} - (1-\beta)\{W(\theta-\eta)\beta^j(1-\alpha)^{-1} \\ + \textstyle\sum_{k=0}^{j-1} c(k)\beta^k\theta + \sum_{k=j}^{\infty} c(k)\alpha^{k-j}\beta^j\eta\}\{\mu-\eta+\beta^j(\eta-\theta)\}^{-1} \geq W. \end{aligned} \tag{28}$$

Now multiply inequality (28) through by the positive quantity

$$\Upsilon(\theta,\mu,j) = \begin{cases} \{\mu-\eta+\beta^j(\eta-\theta)\}(\mu-\theta)^{-1}, & \mu \neq \theta, \\ 1+j(1-\alpha), & \mu = \theta, \end{cases} .$$

to obtain the inequality

$$W(j) \geq W \tag{29}$$

from (13). However, inequality (29) is guaranteed by Lemma 4.1 and the fact that $W \in [W(i-1), W(i))$. Inequality (24) must then follow.

We conclude from Lemma 4.2 that $u(i)$ is optimal for $P(W,M)$ for every $M \geq i-1$. By consideration of the limit $M \to \infty$, it is straightforward to deduce that $u(i)$ must be optimal for $P(W)$. This concludes the proof. □

We now restore the station suffix n and require that Assumption 4.1 hold for each station. Hence for each n there is a corresponding sequence $\{W_n(i),\, i \in \mathbb{N}\}$ given by

$$W_n(i) = \sum_{j=0}^{i} \beta_n^j \big\{\alpha_n(1-\alpha_n) \sum_{k=i+1}^{\infty} c_n(k)\alpha_n^{k-i-1} + c_n(i)(1-\alpha_n) - c_n(j)\beta_n\big\} + c_n(i)\beta_n^{i+1},\ i \in \mathbb{N}. \tag{30}$$

We now return to the Lagrangian relaxation in (11) and invoke the stationwise decomposition in (12) together with Theorem 4.1 to conclude the following:

Theorem 4.2 *For all $W \in \mathbb{R}$ the optimization problem in* (11) *is solved by a policy which operates as follows: When a generic customer arrives, if the system state is $\boldsymbol{x} = (x_1, x_2, \ldots, x_N)$ then the customer should be admitted at all stations in $A(\boldsymbol{x}, W)$ where*

$$A(\boldsymbol{x}, W) = \{n : W_n(x_n) \leq W\}. \tag{31}$$

To use the terminology of Whittle [50], the above analysis tells us that the system is *indexable* with $W_n : \mathbb{N} \to \mathbb{R}^+$ the index for station n, $1 \leq n \leq N$. Since it follows from the above analysis that when $W = W_n(i)$ for some i, n, both actions (accept, reject) are optimal for station n when its queue length is i then $W_n(i)$ may be thought of as a *fair charge* for rejecting a customer. A solution to the relaxation described at the beginning of this section (admit generic customers to one station on average) may then be obtained from the policies in Theorem 4.2 with an appropriate choice of the Lagrange multiplier. It is natural to follow Whittle [50] in proposing as a heuristic for the original problem the policy which routes all incoming generic customers to any station for which this fair charge is the smallest. Formally, when a generic customer arrives, if the system state is $\boldsymbol{x} = (x_1, x_2, \ldots, x_N)$, then the customer should be routed to any station n^* for which

$$W_{n^*}(x_n) = \min_{n \in \{1,2,\ldots,N\}} W_n(x_n).$$

5 Examples

In this section, we assume certain structures for the waiting cost function $C_n(\cdot)$ and give expressions for the routing indices developed in Sections 3 and 4. The first example considers an increasing non-continuous and non-convex cost function while the second example assumes an increasing, continuous, and strictly convex cost function. In Section 6, we test the performance of our heuristics along with two other heuristics from the literature under the cost structures assumed by these two examples.

5.1 Example 1: A non-continuous and non-convex cost function

In this example, we assume that a customer with a system time of t in station n will incur a cost

$$C_n(t) = h_n t + d_n I(t \geq \tau_n) + \hat{h}_n (t - \tau_n)^+, \tag{32}$$

where $(y)^+ = \max(0, y)$. Hence, for each time unit spent in the queue, the customer will incur h_n. Furthermore, should the customer spend more time in the system than τ_n then she will incur a one-time cost of d_n and subsequently an additional cost of $\hat{h}_n$ for each additional time unit above τ_n. Such cost functions as in (32) are relevant in many settings including service systems where customers are given delivery time guarantees and in emergency response settings, where there is typically a time window (e.g., "golden hour rule," see, e.g., [36]) for an acceptable level of response time for critically injured patients.

For the cost function given in (32), it can be shown that

$$c_n(i) = \frac{h_n(i+1)}{\mu_n} + d_n \sum_{j=0}^{i} \frac{(\mu_n \tau_n)^j}{j!} e^{-\mu_n \tau_n} + \hat{h}_n \sum_{j=0}^{i} (i+1-j)(1/\mu_n) \frac{(\mu_n \tau_n)^j}{j!} e^{-\mu_n \tau_n}.$$

Then, it is easy to see that this cost function trivially satisfies Assumptions 3.1 and 4.1. Moreover, after some tedious but straightforward algebraic manipulations, the policy improvement index (8) can be shown to be

$$\begin{aligned} D_n(i) = {} & \frac{h_n(i+1)}{\mu_n - \Lambda_n} + \frac{d_n e^{-\mu_n \tau_n}}{1-\rho_n} \left(e^{\Lambda_n \tau_n} \left(\frac{1-\rho_n^{i+1}}{\rho_n^i} \right) + \sum_{j=0}^{i} \frac{(\mu_n \tau_n)^j}{j!} - \frac{1}{\rho_n^i} \sum_{j=0}^{i} \frac{(\Lambda_n \tau_n)^j}{j!} \right) \\ & + \frac{\hat{h}_n e^{-\mu_n \tau_n}}{(1-\rho_n)^2 \mu_n} \left(\frac{e^{\Lambda_n \tau_n}(1-\rho_n^{i+1})}{\rho_n^i} + \frac{(1-\rho_n)(\mu_n \tau_n)^{i+1}}{i!} \right. \\ & \left. + ((1-\rho_n)(i - \mu_n \tau_n) + 1) \sum_{j=0}^{i} \frac{(\mu_n \tau_n)^j}{j!} - \frac{1}{\rho_n^i} \sum_{j=0}^{i} \frac{(\Lambda_n \tau_n)^j}{j!} \right). \end{aligned}$$

Similarly, the Lagrangian relaxation index (30) can be shown to have the form:

$$W_n(i) = \frac{(1-\beta_n^{i+1})(1-\alpha_n)R_n(i)}{(1-\beta_n)\alpha_n^i} - S_n(i-1),$$

where

$$\begin{aligned} R_n(i) =& \frac{h_n}{\mu_n}\left(\frac{(i+1)\alpha_n^i - i\alpha^{i+1}}{(1-\alpha_n)^2}\right) + \frac{d_n e^{-\mu_n\tau_n}}{1-\alpha_n}\left(e^{\eta_n\tau_n} + \alpha_n^i\sum_{j=0}^{i-1}\frac{(\mu_n\tau_n)^j}{j!} - \sum_{j=0}^{i-1}\frac{(\eta_n\tau_n)^j}{j!}\right) \\ &+ \frac{\hat{h}_n e^{-\mu_n\tau_n}}{\mu_n(1-\alpha_n)}\Bigg(\alpha_n^i(i+(1/(1-\alpha_n)) - \mu_n\tau_n)\sum_{j=0}^{i-1}\frac{(\mu_n\tau_n)^j}{j!} \\ &- \frac{1}{1-\alpha_n}\sum_{j=0}^{i-1}\frac{(\eta_n\tau_n)^j}{j!} + \frac{(\eta_n\tau_n)^i}{(i-1)!} + \frac{e^{\eta_n\tau_n}}{1-\alpha_n}\Bigg) \end{aligned}$$

and

$$\begin{aligned} S_n(i) =& \frac{h_n\beta_n}{\mu_n(1-\beta_n)^2}\left(1-(i+2)\beta_n^{i+1} + (i+1)\beta_n^{i+2}\right) \\ &+ \frac{d_n e^{-\mu_n\tau_n}}{1-\beta_n}\left(\beta_n\sum_{j=0}^{i}\frac{(\theta_n\tau_n)^j}{j!} - \beta_n^{i+2}\sum_{j=0}^{i}\frac{(\mu_n\tau_n)^j}{j!}\right) \\ &+ \frac{\hat{h}_n\beta_n e^{-\mu_n\tau_n}}{\mu_n(1-\beta_n)^2}\Bigg(\sum_{j=0}^{i}\frac{(\theta_n\tau_n)^j}{j!} + \beta_n^i\left((i+1)\beta_n^2 - (i+2)\beta_n\right)\sum_{j=0}^{i-1}\frac{(\mu_n\tau_n)^j}{j!} \\ &+ \frac{(\theta_n\tau_n)^i}{i!}\left((i+1)\beta_n^2 - (i+2)\beta_n\right)\Bigg). \end{aligned}$$

5.2 Example 2: A continuous and convex cost function

For this example, we consider a cost function with structural properties that are quite different from those of the cost function of Example 1. To be more precise, we assume that a customer with a system time of t in station n will incur a cost

$$C_n(t) = t^2. \tag{33}$$

Note that this cost function is continuous, strictly increasing, and convex. Hence, it satisfies the conditions that are needed for asymptotic optimality of Stolyar's MinDrift rule [40].

For this cost function, it can be easily shown that

$$c_n(i) = \frac{(i+1)(i+2)}{\mu_n^2},$$

and hence Assumptions 3.1 and 4.1 are trivially satisfied. After some algebraic manipulations, the policy improvement index (8) can be shown to be

$$D_n(i) = \frac{(i+1)(i(1-\rho_n)+2)}{\mu_n^2(1-\rho_n)^2}.$$

On the other hand, the Lagrangian relaxation index (30) can be shown to have the form:

$$W_n(i) = \frac{(1-\beta_n^{i+1})(1-\alpha_n)}{(1-\beta_n)}\left(\frac{T_n(i)}{\alpha_n^i} + c_n(i)\right) - \sum_{j=0}^{i-1}\beta_n^{j+1}c_n(j),$$

where

$$T_n(i) = \frac{(i+3)(i+2)\alpha_n^{i+1}(1-\alpha_n)^2 + 2\alpha_n^{i+2}\Big((i+3)-(i+2)\alpha_n\Big)}{(1-\alpha_n)^3\mu_n^2}.$$

6 Numerical results

This section reports our findings from a numerical study that we conducted to assess the performance of the heuristics developed in Sections 3, 4, and 5, and to observe the effects of some system parameters on the performance of these heuristics. In our numerical study, we consider a system with two service stations ($N = 2$) and assume cost functions in the form of either (32) or (33). For such systems, it is possible (though expensive) to determine optimal routing policies numerically by dynamic programming, thus permitting a complete evaluation of the proposed heuristics.

Our preliminary numerical results suggested that the performance of the heuristics clearly depended on the system load, $(\lambda+\eta_1+\eta_2)/(\mu_1+\mu_2)$. Therefore, we compared the performance of the heuristics under three different traffic levels.

(i) Light traffic: System load ranges from 0 to 0.6.

(ii) Medium traffic: System load ranges from 0.6 to 0.8.

(iii) Heavy traffic: System load ranges from 0.8 to 1.0.

Under each setting, we generated scenarios by choosing η_1, η_2, and λ uniformly over the interval $[0, 1]$, and μ_1 and μ_2 over the interval $[0, 1.5]$, and discarding the cases with system loads that did not fall into the desired range. For each scenario, we computed the performance of the optimal policy, the policy improvement heuristic (PIH), the Lagrangian relaxation heuristic (LRH), the MinDrift heuristic (MDH), the greedy heuristic (GH), and the optimal static routing heuristic (SH). Sections 3, 4, and 5 provide detailed descriptions of how PIH and LRH work. We next explain the MDH, GH, and SH policies.

The MinDrift heuristic, which is due to Stolyar [40], is an index heuristic similar to PIH and LRH. It simply routes each incoming customer to station n for which the corresponding index $M_n(i)$ is the smallest when there are i customers at station n. The index $M_n(i)$ is obtained by using the derivative of the cost function and the service rates, ignoring any jumps in the cost function. For Example 1, the MinDrift index (for station n when there are i customers) is given by

$$M_n(i) = \begin{cases} h_n/\mu_n, & i/\mu_n < \tau_n, \\ (h_n + \hat{h}_n)/\mu_n, & i/\mu_n \geq \tau_n, \end{cases}$$

while the MinDrift index for Example 2 is

$$M_n(i) = i/\mu_n^2.$$

Stolyar [40] proves that the MinDrift heuristic is asymptotically optimal under heavy traffic for strictly convex and continuous waiting cost functions. Therefore, there is no reason

to expect that this rule will perform well for Example 1, for which the cost function is neither continuous nor convex. However, both conditions are satisfied in Example 2, and consequently stronger performance of MDH is expected there, especially so in heavy traffic instances.

The greedy heuristic is the standard proposal for routing problems. It routes each incoming generic customer to the queue for which the expected cost for that particular customer is the smallest. (Note that GH is equivalent to SDR when cost parameters are the same for all stations.) Finally, the optimal static routing heuristic is the optimal policy among all static routing policies (see Equation (2)) and is the input policy for the second stage of PIH. Note that this heuristic is guaranteed to be inferior to the policy improvement heuristic. However, we still report the performance of SH in order to quantify the improvements gained by state-dependent policies over this static policy.

The remainder of this section is organized as follows. In Section 6.1, we compare the performance of our heuristic policies LRH and PIH with the performances of MDH, SH, and GH. In Section 6.2, we discuss the effects of two system parameters on the performance of the heuristics under consideration.

6.1 Comparison of the heuristic policies

In this section, we compare the performance of the heuristics for Examples 1 and 2, which are described in Sections 5.1 and 5.2, respectively.

6.1.1 Performance of the heuristics: Example 1

In this part of the study, we assumed a cost function in the form of (32) with the choices $h_1 = h_2 = \hat{h}_1 = \hat{h}_2 = 1$, $d_1 = d_2 = 8$, and $\tau_1 = \tau_2 = 5$. For each traffic setting, we randomly

generated 5,000 scenarios, and under each scenario, for each heuristic, we computed the percentage deviation of the performance from that of the optimal policy and constructed a 99% confidence interval on the percentage deviation. For each heuristic, we also determined the median, the lower and upper quartiles for the percentage deviation of each policy from the optimal cost, and the number of occasions upon which that heuristic provided the smallest cost among all heuristics considered. We report our findings in Tables 1, 2, and 3 under light, medium, and heavy traffic settings, respectively. Note that the last columns in the tables add up to a value larger than 5,000 due to ties among heuristics.

Heuristic	**Lower quartile**	**Median**	**Upper quartile**	**99% C. I. for the mean**	**Best heuristic in**
LRH	0.000	0.000	0.011	0.034 ± 0.005	4030 scenarios
PIH	0.000	0.014	0.557	0.668 ± 0.053	1927 scenarios
MDH	0.420	8.747	42.754	31.108 ± 1.705	223 scenarios
GH	0.018	0.327	1.597	1.224 ± 0.070	932 scenarios
SH	0.369	5.338	23.877	13.761 ± 0.628	0 scenarios

Table 1: Performance of the heuristics for Example 1 under light traffic (numbers except those in the last column are for percentage deviation from the optimal)

Heuristic	**Lower quartile**	**Median**	**Upper quartile**	**99% C. I. for the mean**	**Best heuristic in**
LRH	0.015	0.086	0.315	0.603 ± 0.421	3543 scenarios
PIH	0.054	0.712	2.782	2.148 ± 0.693	1213 scenarios
MDH	3.513	22.638	64.147	43.790 ± 2.028	73 scenarios
GH	0.416	1.653	4.018	2.685 ± 0.108	555 scenarios
SH	3.669	25.618	50.253	28.816 ± 0.900	0 scenarios

Table 2: Performance of the heuristics for Example 1 under medium traffic (numbers except those in the last column are for percentage deviation from the optimal)

While LRH, PIH, and GH all perform reasonably well, the numerical results presented in Tables 1, 2, and 3 demonstrate the superior performance of LRH under each traffic setting. Much the most important aspect of these tables is the clear evidence it provides that LRH

Heuristic	Lower quartile	Median	Upper quartile	99% C. I. for the mean	Best heuristic in
LRH	0.735	1.728	4.799	4.230 ± 0.418	3851 scenarios
PIH	2.433	6.130	9.912	7.880 ± 0.361	720 scenarios
MDH	5.695	14.040	32.616	25.603 ± 1.235	414 scenarios
GH	2.450	4.898	8.361	5.782 ± 0.150	528 scenarios
SH	13.885	29.998	43.946	29.514 ± 0.689	0 scenarios

Table 3: Performance of the heuristics for Example 1 under heavy traffic (numbers except those in the last column are for percentage deviation from the optimal)

outperforms the standard proposal GH for the problems studied. One can also observe from these tables that the performance of all heuristics except for MDH deteriorates as the system load increases. Note however that the mean performance of LRH is around 4.23% with median 1.728% even under heavy traffic. Also, comparison of the median and quartiles with the mean performance indicates that the distribution of the performances are skewed to the right for all the heuristics. The performance of MDH is clearly poor; however, this is not surprising since MDH was developed under assumptions that are not satisfied for the cost function considered for this example. One observation is that MDH seems to provide its best performance under heavy traffic and its worst performance under medium traffic.

Although LRH clearly seems to be the best heuristic, comparison of PIH and GH is more complicated. PIH seems to be the heuristic that is affected most by the increase in the traffic. Its performance gets significantly worse as we move from medium traffic to heavy traffic. Under heavy traffic, the mean performance of GH is better than that of PIH. However, PIH is the best heuristic in more scenarios than is GH. At other traffic levels, PIH outperforms GH in all aspects.

Finally, Tables 1, 2, and 3 clearly show that state-dependent policies provide substantial improvements over the "best" static policy, SH.

6.1.2 Performance of the heuristics: Example 2

In this section, we assume that the waiting cost function is in the form of (33). Note that this cost function satisfies all the assumptions needed for the asymptotic optimality of MDH (see Stolyar [40]), and therefore MDH is expected to perform well at least under heavy traffic.

For each traffic setting, we randomly generated 5,000 scenarios, and computed the performance of each heuristic under each scenario. Tables 4, 5, and 6 present our findings under light, medium, and heavy traffic, respectively. Similar to our results for Example 1, LRH has the best performance among all the heuristics considered under this convex and continuous function. Except for MDH, the performance of all the heuristics worsen as the traffic intensity increases. The effect of the traffic load on MDH is not clear, since one would reach different conclusions depending on whether the mean and the confidence interval or median and the quartiles are chosen to be the deciding performance measure. However, if compared with the performance of all the other heuristics at all traffic levels, one can observe that as expected, the relative performance of MDH is the best when under heavy traffic. However, LRH still seems to perform slightly better than MDH even under heavy traffic (as defined in this paper). PIH performs relatively well under light and medium traffic, but under heavy traffic, its overall performance is poor. On the other hand, GH is not as good as PIH under light and medium traffic, but its performance under heavy traffic is much better than that of PIH on the average. However, even under heavy traffic, PIH is the best heuristic in more scenarios than GH is. Finally, as in Example 1, the performance of SH is clearly poor, demonstrating the superiority of state-dependent policies.

Heuristic	Lower quartile	Median	Upper quartile	99% C. I. for the mean	Best heuristic in
LRH	0.000	0.000	0.011	0.046 ± 0.007	4038 scenarios
PIH	0.000	0.017	0.706	0.915 ± 0.074	1821 scenarios
MDH	0.698	3.048	7.809	8.116 ± 0.874	192 scenarios
GH	0.028	0.505	2.260	1.700 ± 0.095	836 scenarios
SH	0.394	6.543	35.792	20.142 ± 0.943	0 scenarios

Table 4: Performance of the heuristics for Example 2 under light traffic (numbers except those in the last column are for percentage deviation from the optimal)

Heuristic	Lower quartile	Median	Upper quartile	99% C. I. for the mean	Best heuristic in
LRH	0.017	0.130	0.442	0.339 ± 0.018	3660 scenarios
PIH	0.051	1.000	4.094	2.827 ± 0.284	1134 scenarios
MDH	1.770	4.503	8.599	6.408 ± 0.360	149 scenarios
GH	0.679	2.525	5.778	3.832 ± 0.145	472 scenarios
SH	3.735	39.943	83.889	47.099 ± 1.567	0 scenarios

Table 5: Performance of the heuristics for Example 2 under medium traffic (numbers except those in the last column are for percentage deviation from the optimal)

Heuristic	Lower quartile	Median	Upper quartile	99% C. I. for the mean	Best heuristic in
LRH	1.302	2.588	6.597	6.004 ± 0.328	2849 scenarios
PIH	3.315	9.248	15.469	18.871 ± 1.845	982 scenarios
MDH	2.634	5.388	10.173	7.276 ± 0.225	1193 scenarios
GH	4.016	7.131	11.053	7.951 ± 0.188	224 scenarios
SH	22.539	58.985	92.626	61.173 ± 1.759	0 scenarios

Table 6: Performance of the heuristics for Example 2 under heavy traffic (numbers except those in the last column are for percentage deviation from the optimal)

6.2 Effects of system parameters on the performance of the heuristics

In this section, we focus our attention on Example 1, assuming a cost function in the form of (32) with parameter values set as in Section 6.1.1, and investigate the effects of two system parameters on the performances of LRH, PIH, and GH, the three best heuristics for this example.

- **Effects of dedicated load heterogeneity:** In the system under consideration, each

service station has its own dedicated customers, and therefore has its own capacity that can be allocated to the service of generic customers. If these residual capacities for the two stations are close to each other, we say that the dedicated loads are homogeneous, otherwise we say that they are heterogeneous. To observe the effects of the heterogeneity of the stations on the performance of the heuristics, we first define a heterogeneity index, Φ as

$$\Phi = \left| 1 - \frac{\eta_1/\mu_1}{\eta_2/\mu_2} \right|.$$

After computing the heterogeneity index for each of the 5,000 scenarios, we ordered all scenarios according to this index from the smallest to the largest. Then, we computed the moving average (of order 2,000) of the mean performance across the 5,000 scenarios (see Figure 1 for plots of the moving averages). In all cases, under each traffic setting, we observed a worsening performance of the heuristics with an increase in Φ. Hence, the more balanced the system load is across the two stations, the better is the average performance of all the heuristics. However, the deterioration in performance is more significant for GH, less so for LRH, and even less so for PIH. Furthermore, for LRH and PIH, under heavy traffic, the average performance does not seem to change drastically as we go from very low levels of heterogeneity to mid levels whereas the performance starts to deteriorate significantly as we go from mid levels to high levels. On the other hand, the performance of GH worsens at a seemingly constant level as we go from very low levels of heterogeneity to very high levels.

- **Effects of dedicated arrival rates:** We have also studied the numerical results to identify cases where the heuristic algorithms performed poorly. We observed that the

dedicated arrival rates have an interesting effect on the performance of the heuristics. To observe this effect, we ordered the scenarios with respect to their total dedicated arrival rates, $\eta_1 + \eta_2$, and carried out an analysis similar to that for the load heterogeneity (see Figure 2). We observed that the smaller the total dedicated arrival rate, the worse is the performance of the LRH on the average. (Note, however, that LRH still has the best performance among all the heuristics even over the region where its performance is poor). Its performance is significantly worse for very low levels of the dedicated arrival rate. This is surprising since lower dedicated arrival rates indicate low traffic intensity, which overall seems to have a positive impact on the performance of all heuristics.

We observed a similar effect on the performance of PIH although the effect does not seem to be as strong as for LRH. On the other hand, the effect on GH seems to depend on the total traffic load on the system. Under medium traffic, very low and very high levels of dedicated traffic rate have a positive effect on the performance while moderate levels yield worse performance on the average. Under light and heavy traffic, the performance worsens as the total dedicated arrival rate increases.

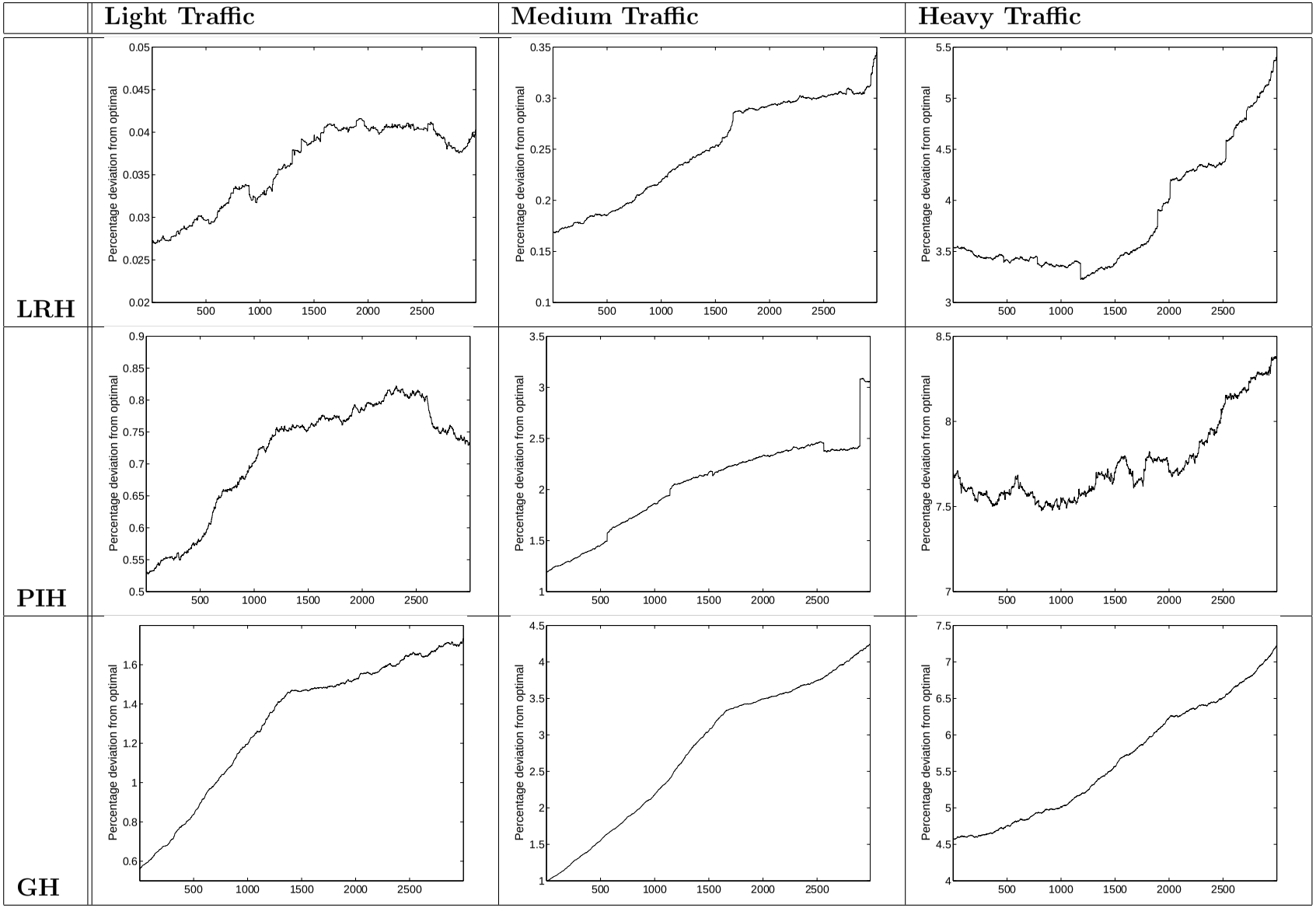

Figure 1: Moving average plots of percentage deviations of heuristics from the optimal with respect to the number of scenarios that are ordered according to the increasing heterogeneity index of dedicated load Φ

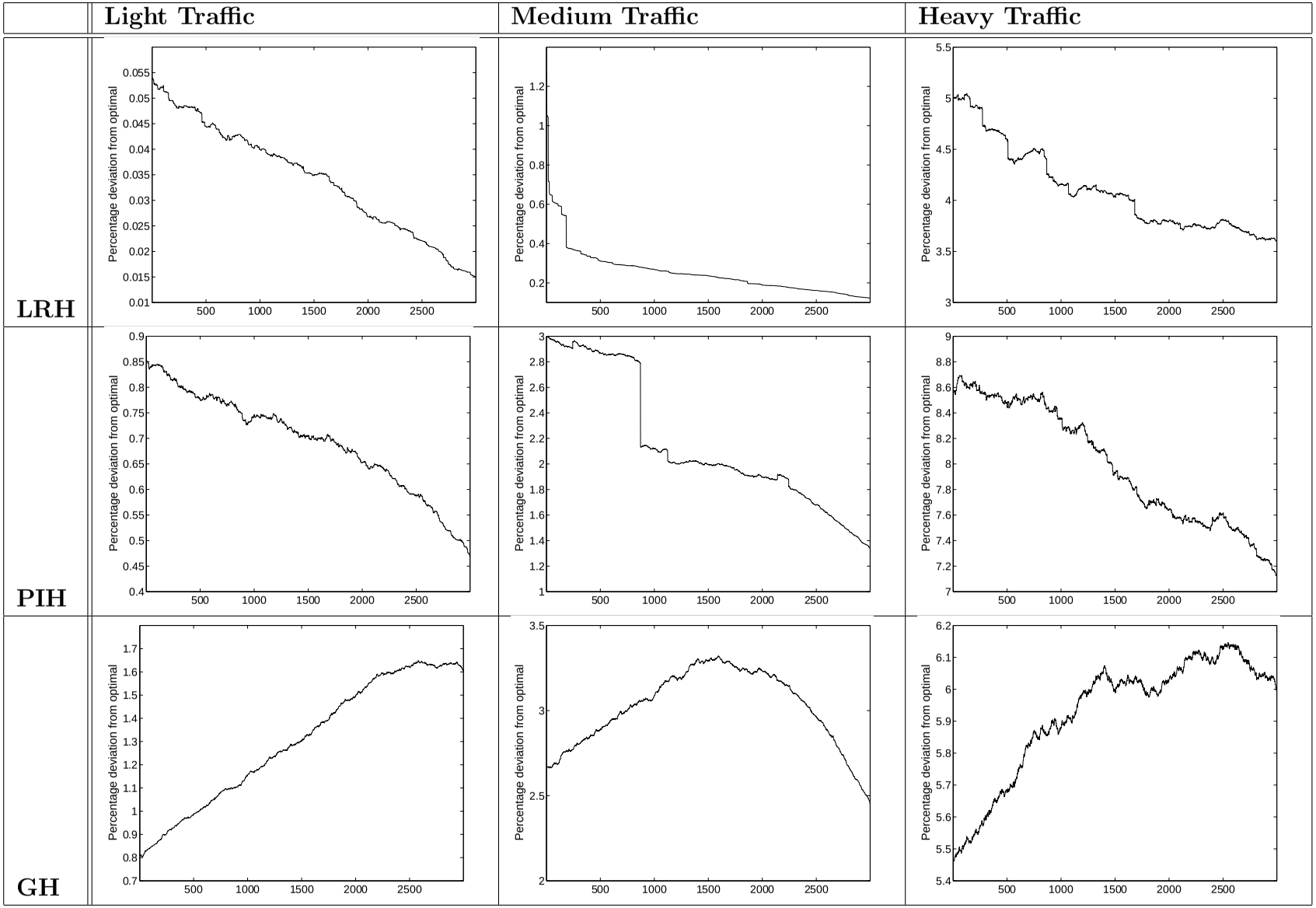

Figure 2: Moving average plots of percentage deviations of heuristics from the optimal with respect to the number of scenarios that are ordered according to the increasing total dedicated arrival rates $\eta_1 + \eta_2$

Acknowledgement: The second author acknowledges support given by Edinburgh University and the British Council during her doctoral studies. The third author acknowledges EPSRC support made available through grant GR/S45188/01.

References

[1] P. S. Ansell, K. D. Glazebrook, and C. Kirkbride, Generalised "join the shortest queue" policies for the dynamic routing of jobs to multi-class queues, Journal of Operational Research Society 54 (2003) 379–389.

[2] P. S. Ansell, K. D. Glazebrook, J. Niño-Mora, and M. O'Keeffe, Whittle's index policy for a multi-class queueing system with convex holding costs, Mathematical Methods in Operations Research 57 (2003) 21–39.

[3] M. Armony, Dynamic routing in large-scale service systems with heterogeneous servers, Queueing Systems 51 (2005) 287–329.

[4] M. Armony and C. Maglaras, On customer contact centers with a call-back option: Customer decisions, routing rules, and system design, Operations Research 52 (2004) 271–292.

[5] M. Armony and C. Maglaras, Contact centers with a call-back option and real time delay information, Operations Research 52 (2004) 527–545.

[6] A. Bassamboo, J. M. Harrison, and A. Zeevi, Dynamic routing and admission control in high-volume service systems: Asymptotic analysis via multi-scale fluid limits, Queueing Systems 51 (2005) 249–285.

[7] K. J. Becker, D. P. Gaver, K. D. Glazebrook, P. A. Jacobs, and S. Lawphongpanich, Allocation of tasks to specialised processors: a planning approach, European Journal of Operational Research 126 (2000) 80–88.

[8] S. Bhulai, Dynamic routing policies for multi-skill call centers, Technical Report 2004-11, Vrije University, Amsterdam, The Netherlands (2004).

[9] S. Bhulai and G. Koole, On the structure of value functions for threshold policies in queueing models, Journal of Applied Probability 40 (2003) 613–622.

[10] K. Charnsirisakskul, P. M. Griffin, and P. Keskinocak, Order selection and scheduling with leadtime flexibility, IIE Transactions 36 (2004) 697–707.

[11] R. D. Foley and D. R. McDonald, Join the shortest queue: stability and exact asymptotics, Annals of Applied Probability 11 (2001) 569–607.

[12] G. J. Foschini and J. Salz, A basic dynamic routing problem and diffusion, IEEE Transactions on Communications 26 (1978) 320–327.

[13] N. Gans and Y. P. Zhou, A call routing problem with service level constraints, Operations Research 51 (2003) 255–271.

[14] K. D. Glazebrook, R. R. Lumley, and P. S. Ansell, Index heuristics for multiclass M/G/1 systems with nonpreemptive service and convex holding costs, Queueing Systems 45 (2003) 81–111.

[15] K. D. Glazebrook, J. Niño-Mora, and P. S. Ansell, Index policies for a class of discounted restless bandits, Advances in Applied Probability 34 (2002) 754–774.

[16] K. D. Glazebrook, H. M. Mitchell, and P. S. Ansell, Index policies for the maintenance of a collection of machines by a set of repairmen, European Journal of Operational Research 165 (2005) 267–284.

[17] K. D. Glazebrook, C. Kirkbride, and J. Ouenniche, Index policies for the routing of impatient customers to heterogeneous service stations, submitted for publication, (2005).

[18] D. Gross and C. M. Harris, Fundamentals of Queueing Theory, Wiley, 1998.

[19] J. M. Harrison and A. Zeevi, Dynamic scheduling of a multi-class queue in the Halfin-Whitt heavy traffic regime, Operations Research 52 (2004) 243–257.

[20] W. J. Hopp and M. L. R. Sturgis, Quoting manufacturing due dates subject to a service level constraint, IIE Transactions 32 (2000) 771–784.

[21] A. Hordijk and G. Koole, On the optimality of the generalized shortest queue policy, Probability in the Engineering and Information Sciences 4 (1990) 477–487.

[22] D. J. Houck, Comparison of policies for routing customers to parallel queueing systems, Operations Research 35 (1987) 306–310.

[23] P. K. Johri, Optimality of the shortest line discipline with state-dependent service rates, European Journal of Operational Research 41 (1989) 157–161.

[24] D. Jones, Some hospital ERs begin guaranteeing quick service, USA Today, December 3, 2004.

[25] F. P. Kelly and C. N. Laws, Dynamic routing in open queueing networks: Brownian models, cut constraints and resource pooling, Queueing Systems 13 (1993) 47–86.

[26] Y. Kogan, Y. Levy, and R. A. Milito, Call routing to distributed queues: Is FIFO really better than MED?, Telecommunication Systems 7 (1997) 299–312.

[27] G. Koole, P. D. Sparaggis, and D. Towsley, Minimizing response times and queue lengths in systems of parallel queues, Journal of Applied Probability 36 (1999) 1185–1193.

[28] K. R. Krishnan, Joining the right queue: a Markov decision rule, Proceedings of the 26th IEEE Conf. Decision Control (1987) 1863–1868.

[29] K. R. Krishnan, Joining the right queue: A state-dependent decision rule, IEEE Transactions on Automatic Control 35 (1990) 104–108.

[30] K. R. Krishnan and T. J. Ott, State-dependent routing for telephone traffic: Theory and results, Proceedings of the 25th IEEE Conf. Decision Control (1986) 2124–2128.

[31] C. N. Laws, Resource pooling in queueing networks with dynamic routing, Advances in Applied Probability 24 (1992) 699–726.

[32] E. J. MacKenzie, D. B. Hoyt, J. C. Sacra, G. J. Jurkovich, A. R. Carlini, S. D. Taitelbaum, and H. Teter, Jr., National Inventory of Hospital Trauma Centers, Journal of American Medical Association 289 (2003) 1515–1522.

[33] A. Mandelbaum and A. L. Stolyar, Scheduling flexible servers with convex delay costs: Heavy traffic optimality of generalized $c\mu$-rule, Operations Research 52 (2004) 836–855.

[34] B. Menich and R. Serfozo, Optimality of routing and servicing in dependent parallel processing systems, Queueing Systems 9 (1991) 403–418.

[35] J. Niño-Mora, Dynamic allocation indices for restless projects and queueing admission control, Mathematical Programming 93 (2002) 361–413.

[36] J. J. Osterwalder, Can the "golden hour of shock" be safely extended in blunt polytrauma patients, Prehospital and Disaster Medicine 17 (2002) 75–80.

[37] K. W. Ross and D. D. Yao, Optimal load balancing and scheduling in a distributed computer system, Journal of the Association for Computing Machinery 38 (1991) 676–690.

[38] S. A. E. Sassen, H. C. Tijms, and R. D. Nobel, A heuristic rule for routing customers to parallel servers, Statistica Neerlandica 51 (1997) 107–121.

[39] P. D. Sparaggis, D. Towsley, and C. G. Cassandras, Extremal properties of the SNQ and the LNQ policies in finite capacity systems with state-dependent service rates, Journal of Applied Probability 30 (1993) 223–236.

[40] A. L. Stolyar, Optimal routing in output-queued flexible server systems, Probability in the Engineering and Informational Sciences 19 (2005) 141–189.

[41] Y. Teh and A. Ward, Critical thresholds for dynamic routing in queueing networks, Queueing Systems 42 (2002) 297–316.

[42] T. Tezcan, Optimal control of distributed parallel server systems under the Halfin and Whitt regime, School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, GA, U.S.A. (2005).

[43] H. C. Tijms, Stochastic Models - An Algorithmic Approach, Wiley, 1994.

[44] D. Towsley, P. D. Sparaggis, and C. G. Cassandras, Optimal routing and buffer allocation for a class of finite capacity queueing systems, IEEE Transactions on Automatic Control 37 (1993) 1446–1451.

[45] J. A. Van Mieghem, Dynamic scheduling with convex delay costs: The generalized $c\mu$ rule, The Annals of Applied Probability 5 (1995) 809–833.

[46] R. B. Wallace and W. Whitt A staffing algorithm for call centers with skill-based routing, Manufacturing and Service Operations Management 7 (2005) 276–294.

[47] R. Weber, On the optimal assignment of customers to parallel servers, Journal of Applied Probability 15 (1978) 406–413.

[48] L. M. Wein, Brownian networks with discretionary routing, Operations Research 39 (1991) 322–340.

[49] W. Whitt, Deciding which queue to join: Some counterexamples, Operations Research 34 (1986) 55–62.

[50] P. Whittle, Restless bandits: activity allocation in a changing world, Journal of Applied Probability A25 (1988) 287–298.

[51] P. Whittle, Optimal Control: Basics and Beyond, Wiley, 1996.

[52] W. Winston, Optimality of the shortest-processing-time discipline, Journal of Applied Probability 14 (1977) 181–189.