



## Stochastics and Statistics

## A Bayesian approach to the triage problem with imperfect classification

Dong Li<sup>a,\*</sup>, Kevin D. Glazebrook<sup>b</sup><sup>a</sup> *Avis Europe, Avis House, Park Road, Bracknell, RG12 2EW, UK*<sup>b</sup> *Department of Management Science, Lancaster University, LA1 4YX, UK*

## ARTICLE INFO

## Article history:

Received 23 July 2010

Accepted 24 May 2011

Available online 30 May 2011

## Keywords:

Dynamic programming  
 Bayes sequential decision problem  
 Imperfect classification  
 Stochastic scheduling  
 Optimal service policy

## ABSTRACT

A collection of jobs (or customers, or patients) wait impatiently for service. Each has a random *lifetime* during which it is available for service. Should this lifetime expire before its service starts then it leaves unserved. Limited resources mean that it is only possible to serve one job at a time. We wish to schedule the jobs for service to maximise the total number served. In support of this objective all jobs are subject to an initial *triage*, namely an assessment of both their urgency and of their service requirement. This assessment is subject to error. We take a Bayesian approach to the uncertainty generated by error prone triage and discuss the design of heuristic policies for scheduling jobs for service to maximise the Bayes' return (mean number of jobs served). We identify problem features for which a high price is paid in number of services lost for poor initial triage and for which improvements in initial job assessment yield significant improvements in service outcomes. An analytical upper bound for the cost of imperfect classification is developed for exponentially distributed lifetime cases.

© 2011 Elsevier B.V. All rights reserved.

## 1. Introduction

A finite collection of jobs (or customers, or patients) await service. The wait is in all cases characterised by *impatience*, namely each job has a random *lifetime* during which it remains available for service. Should any job not receive service prior to the expiry of its lifetime, it leaves the system unserved. Limited service resources mean that it is only possible to serve one job at a time. A key goal of analysis concerns the development of policies for sequencing the jobs to maximise the number of jobs which receive service.

In discussing such problems, Argon et al. (2008) emphasise their relevance to a situation in which limited medical resources must be deployed optimally following a mass casualty incident (MCI). Immediately after such incidents, a significant number of injuries are caused and these may overwhelm local medical resources. To support efficient resource allocation, all patients are subject to an initial *triage*, namely an assessment of the severity of their condition, and are classified into distinct priority groups. This process must necessarily be undertaken speedily. However, it may well be the case that determination of the actual criticality level in a short timeframe is challenging, and that the initial triage is subject to significant levels of error. Indeed, in a review related to terrorist bombing events, Frykberg and Tepas (1988) found that on average 59% of those classified as critical were actually non-critical, with a 0.05% error rate for the reverse. In medical terminology,

these two triage errors are called *overtriage* and *undertriage*, respectively. See also Turégano-Fuentes et al. (2008). It has been shown by Frykberg and Tepas (1988) that the accuracy of triage can have a significant impact on casualty survival in MCIs.

In this paper we propose a simple analytical model which enables such issues to be explored. In our set up each job requiring service is subject to an initial triage to assess how much service resource it needs and how urgent it is. Each job is one of  $J$  types. Jobs of the same type have random lifetimes and service times which are identically distributed. Both of these key random quantities are independent for distinct jobs. A job will leave the system unserved unless its service begins before its lifetime expires. Effective sequencing of the jobs should prioritise those whose service times and/or lifetimes are small. Because the triage is subject to error, the decision-maker is uncertain regarding the true type of each job. We adopt a *Bayesian* approach to this uncertainty. Hence each job has a *posterior distribution* over job types which expresses the decision maker's beliefs about its true identity after triage but before service begins. As time passes, these beliefs are updated dynamically and *posteriors* are recomputed at every decision epoch using Bayes' Theorem. It is almost always true that these posteriors become increasingly informative about a job's true identity as time increases. The primary goal of analysis is the development of an approach to scheduling the jobs which maximises the expected number of jobs served or which comes close to doing so.

To our knowledge the above problem has been rarely studied in the literature. There are two contributions which address classification errors when controlling multiclass queueing systems. Van der Zee and Theil (1961) consider a single server queue with two

\* Corresponding author. Tel.: +44 7983 631166.

E-mail address: [devind@gmail.com](mailto:devind@gmail.com) (D. Li).

priority classes of jobs. Incoming jobs can be misclassified. The other contribution is due to Argon and Ziya (2009) who consider a set up in which each arriving job sends out a signal which gives partial information about its true identity. Both contributions focus on optimal priority assignment strategies only. Neither considers the problem of the subsequent scheduling of the jobs. In neither case does the model concerned incorporate impatience.

In Section 2 our service sequencing problem is formulated as a Bayes sequential decision problem and in Section 3 an approach to its solution via dynamic programming (DP) is described. Important contributions to the general theory of Bayesian DP are due to Bellman (1961), Furukawa (1970), Rieder (1975) and Kumar (1985).

An exact DP approach is computationally intractable for problems of reasonable size, even for the case in which triage is without error and job types are known with certainty. For this case, Glazebrook et al. (2004) and Argon et al. (2008) have proposed simple heuristics for job sequencing, neither of which have uniformly strong performance. The best solution approach to date for the perfect triage problem is due to Li and Glazebrook (2010) and utilises a fluid model of the system to underpin an approximate DP methodology. In Section 4 this approach is further developed so that it can yield effective solutions to our Bayesian model for the problem in which triage is subject to error. A numerical study in Section 5 testifies to the strong performance of the resulting heuristic service policy in 18,000 randomly generated problems.

In Sections 6 and 7 we explore the question of whether it is possible to identify problem features for which poor triage is particularly costly in terms of number of service completions lost. To this end, we introduce the (relative) cost of imperfect classification, denoted (R)CIC, as a natural measure of this. The quantity CIC is directly analogous to a measure called the value of perfect information developed by Berry and Kertz (1991) and applied in the context of multi-armed bandit problems. We are able to develop an analytical upper bound for (R)CIC for the case in which the random lifetimes are exponentially distributed. This bound tells us that in the exponential lifetime case, CIC is small whenever the product of mean service time and mean lifetime varies little across job types. In such cases the triage process is relatively unimportant for the scheduling problem. Numerical studies indicate that these insights extend beyond the exponential case.

Many earlier contributions to scheduling problems which have adopted a Bayesian approach to unknown system parameters have deployed the theory of Gittins indices for multi-armed bandits; see Gittins (1979, 1989). Application areas include job scheduling (Gittins and Glazebrook, 1977; Hamada and Glazebrook, 1993; Cai et al., 2009), experimental design (Glazebrook, 1978) and optimal exploration (Glazebrook and Boys, 1995). In a related theoretical contribution, Glazebrook and Owen (1995) assess the value of building adaptive policies for stochastic scheduling which are able to learn about unknown parameters. Application domains impacted by contributions unrelated to multi-armed bandits include the design of sequential screens (Boys et al., 1996), inventory control (Azoury, 1985) and optimal search (Zhao et al., 2008).

**2. A model for the triage problem with imperfect classification**

A clearing system has a single server and a collection of  $N$  jobs (or customers) awaiting service, which starts at time 0. Each job is one of  $J$  types, each type being identified by an integer  $i \in \{1, 2, \dots, J\}$ . Each type  $i$  job has associated with it two positive-valued random variables (r.v.s). One of these is its *lifetime*, namely the period during which the job is available for service, which is deemed to have the distribution of some r.v.  $X_i$  with distribution function  $F_i$ . A job will leave the system unserved if its lifetime expires before it is taken into service. The other is its *service time*,

which, for a type  $i$  job, is deemed to have the distribution of some r.v.  $Y_i$  with distribution function  $G_i$ . A job leaves the system when its service is complete. The lifetimes and service times of all jobs form a mutually independent collection. We assume that service is *nonpreemptive*, namely that a service, once begun, must be completed.

At time 0 the type of each job is observed with error. Should a type  $i$  job be assessed as type  $j$ , we shall say that it becomes a member of *class j*. Hence throughout, we shall use the term *class* to denote the *assessed type* of a job. Only job class is observed. We shall adopt the following simple probabilistic model of job (mis) classification. We write  $p_i$  for the unconditional probability that a job is of type  $i$  and  $\epsilon_{ij}$  for the conditional probability that a type  $i$  job is assessed as  $j$ . If  $\epsilon_{ii} < 1$  for any  $i$  then some misclassification is possible. By deployment of Bayes' Theorem we infer the conditional probability that, in advance of any service, a class  $j$  job is actually of type  $i$  to be

$$p_{ij}(0) \triangleq P(\text{type } i | \text{class } j) = \frac{\epsilon_{ij} p_i}{\sum_{k=1}^J \epsilon_{kj} p_k} \tag{1}$$

As time passes, jobs leave the system as services are completed and lifetimes expire. It is also true, that at some time  $t > 0$  our beliefs with regard to the (true) type of the remaining jobs need to be updated in light of their survival beyond  $t$ . Again applying Bayes' Theorem, we have

$$p_{ij}(t) \triangleq P(\text{type } i | \text{class } j, \text{ lifetime } > t) = \frac{\epsilon_{ij} p_i \{1 - F_i(t)\}}{\sum_{k=1}^J \epsilon_{kj} p_k \{1 - F_k(t)\}} \tag{2}$$

It is trivial to show that if all of the  $\epsilon_{\cdot}$ ,  $p_{\cdot}$  are strictly positive and, further, that there exists some type  $i^*$  which *outlasts the others* in the sense that

$$\lim_{t \rightarrow \infty} \frac{1 - F_i(t)}{1 - F_{i^*}(t)} = 0, \quad i \neq i^*, \tag{3}$$

then we must have

$$\lim_{t \rightarrow \infty} p_{i^*j}(t) = 1, \quad \forall j,$$

and consequently

$$\lim_{t \rightarrow \infty} p_{ij}(t) = 0, \quad i \neq i^*, \quad \forall j.$$

Hence such survival information can be very informative for (true) type identity.

**Example 1.** Consider an example with  $J = 2$  and with the following parameters:

- $p_1 = 0.3, p_2 = 0.7,$
- $\epsilon_{12} = 0.3, \epsilon_{21} = 0.4,$
- $X_1 \sim \text{Weibull}(1.68, 1.62), X_2 \sim \text{Weibull}(1.16, 13.43).$

Note that here and elsewhere, we shall use Weibull  $(\alpha, \beta)$  to denote the distribution function

$$F(t) = 1 - \exp \left\{ - \left( \frac{t}{\beta} \right)^\alpha \right\}, \quad t \in \mathbb{R}^+.$$

At time zero, the probability that a class 1(2) job is indeed of type 1(2) is 0.43(0.82) from Eq. (1). As time passes these conditional probabilities are updated according to Eq. (2). Their values at times 2 and 5 are given below:

$$p_{11}(2) = 0.17, \quad p_{22}(2) = 0.95; \\ p_{11}(5) = 0.00, \quad p_{22}(5) = 1.00.$$

Since type 2 jobs have much longer mean lifetimes (with mean 12.75) than type 1 (mean 1.45), it is unsurprising that we have

$i^* = 2$  here and that type 2 jobs outlast type 1. As time passes, all surviving jobs are increasingly likely to be of type 2.

The goal of analysis is the development of an approach to the allocation of service to surviving jobs (i.e., those still in the system) to maximise the expected number of jobs which leave the system served. Equivalently, we seek to minimise the expected number of jobs which leave the system on the occasion of the expiry of their lifetimes. Decision epochs occur whenever a service completion occurs. Suppose that  $t$  is such an epoch and that at time  $t$ ,  $n_j$  class  $j$  jobs survive. We write  $(\mathbf{n}, t)$  for the corresponding system state, where  $\mathbf{n} \triangleq (n_1, n_2, \dots, n_J)$  is the vector summarising the class membership of surviving jobs. Write

$$\Delta(\mathbf{n}) \triangleq \{j; n_j > 0\}$$

for the set of admissible actions in state  $(\mathbf{n}, t)$ . A service policy  $\pi$  maps each system state  $(\mathbf{n}, t)$  to the action set  $\Delta(\mathbf{n})$ . In applying any such service policy  $\pi$ , specification of the class of job to be served next is all that is required. In terms of the Bayes' return achieved (see (6) below), it is immaterial which job within the chosen class is selected. For definiteness, we may assume that job selection is effected by the application of a uniform randomization over the remaining members of the chosen class. We shall conventionally use  $(\mathbf{L}, 0)$  for the initial system state, where  $\mathbf{L} = \{L_j, 1 \leq j \leq J\}$ . Hence at time 0,  $L_j$  jobs are placed in class  $j$ ,  $1 \leq j \leq J$ .

We give each job present at time 0 a numerical identifier  $m \in \{1, 2, \dots, N\}$ . We shall suppose that this numbering of the jobs is such that the  $L_j$  members of class  $j$  which are initially present have the identifiers  $K_{j-1} + 1 \leq m \leq K_j$ , where  $K_j = \sum_{k=1}^j L_k$ ,  $1 \leq j \leq J$ . We shall use  $\mathbf{I} = \{I_m; 1 \leq m \leq N\}$  for the unobservable true initial state, regarded as a random variable, where  $I_m$  is the true type of job  $m$ . Observation of initial system state  $(\mathbf{L}, 0)$  modifies the unconditional distribution for  $\mathbf{I}$  to the conditional posterior distribution given by

$$P[\mathbf{I} = \mathbf{i} | (\mathbf{L}, 0)] = \prod_{j=1}^J \prod_{m=K_{j-1}+1}^{K_j} p_{i_m j}(0). \tag{4}$$

We shall use  $E_{\mathbf{I} | (\mathbf{L}, 0)}$  as a notational shorthand for an expectation taken with respect to the distribution in (4). Denoting the lifetime of job  $m$  by  $X_m \sim F_m$ , we use  $T_m(\pi)$  to denote the time at which service policy  $\pi$  begins to process job  $m$ . Conditional upon the true initial state  $\mathbf{I}$  we write  $N(\pi | \mathbf{I})$  for the number of jobs to be served to completion under service policy  $\pi$ . We have

$$N(\pi | \mathbf{I}) = \sum_{m=1}^N \mathbb{1}\{T_m(\pi) \leq X_m\}, \tag{5}$$

where  $\mathbb{1}$  is an indicator. For any initial system state  $(\mathbf{L}, 0)$ , the goal of analysis is the determination of a service policy  $\pi$  to maximise the quantity

$$V_{\pi}^e(\mathbf{L}, 0) = E_{\mathbf{I} | (\mathbf{L}, 0)} E\{N(\pi | \mathbf{I})\}. \tag{6}$$

In (6) the inner expectation is taken over realisations of the system, evolving under service policy  $\pi$  from true initial state  $\mathbf{I}$ . The outer expectation is taken with respect to the posterior distribution in (4). We call the quantity in (6) the Bayes' return generated by service policy  $\pi$  from initial state  $(\mathbf{L}, 0)$ . Please note that in (6) and hereafter we shall use a superscript  $e$  to denote the fact that we are dealing with an object associated with a classification which is prone to error (and hence uncertainty).

### 3. Formulation of the Bayes sequential decision problem as a dynamic programme (DP)

We now formulate the Bayes sequential decision problem

$$V^e(\mathbf{L}, 0) \equiv \sup_{\pi} V_{\pi}^e(\mathbf{L}, 0) \tag{7}$$

as a dynamic programme (DP). Any service policy achieving the supremum in (7) is a Bayes' policy. In order to formulate the associated DP we require additional notation. We shall use  $X_j^e$  to denote the random lifetime of a job classified as  $j$  at time 0. Using (1) above, the associated distribution function is given by

$$F_j^e(s) = \sum_{i=1}^J p_{ij}(0) F_i(s), \quad s \in \mathbb{R}^+, \quad 1 \leq j \leq J. \tag{8}$$

Please note that in Eq. (8) and in what follows we shall use  $s$  and  $t$  as time variables. Our practice will be to use  $t$  for decision epochs when required. We shall assume that each  $F_i$  has an associated absolutely continuous density  $f_i$ . Hence  $X_j^e$  has the density

$$f_j^e(s) = \sum_{i=1}^J p_{ij}(0) f_i(s), \quad s \in \mathbb{R}^+, \quad 1 \leq j \leq J.$$

The expression in (8) is a predictive description of the distribution of the random lifetime concerned at time 0. Suppose now that a job classified as  $j$  at time 0 is still alive at time  $t$ . The appropriate conditional distribution for the job's lifetime is then given by

$$\begin{aligned} P(X_j^e \leq t + s | X_j^e > t) &= \frac{F_j^e(t + s) - F_j^e(t)}{1 - F_j^e(t)} \\ &= \sum_{i=1}^J p_{ij}(t) \left( \frac{F_i(t + s) - F_i(t)}{1 - F_i(t)} \right) \end{aligned}$$

from which we can deduce the continuous hazard

$$\theta_j^e(s) = f_j^e(s) \{1 - F_j^e(s)\}^{-1} = \sum_{i=1}^J p_{ij}(s) \theta_i(s), \quad s \in \mathbb{R}^+, \tag{9}$$

$$1 \leq j \leq J,$$

where in (9),  $\theta_i = f_i \{1 - F_i\}^{-1}$  denotes the continuous hazard associated with the type  $i$  lifetime  $X_i$ . Should type  $i^*$  outlast the others in the sense of (3) above then we will have

$$\lim_{s \rightarrow \infty} \{\theta_j^e(s) - \theta_{i^*}^e(s)\} = 0, \quad 1 \leq j \leq J.$$

Concerning job service times, consider a situation in which a job, originally classified as  $j$ , is still in the system and is scheduled for service at time  $t$ . We shall use  $Y_{j,t}^e$  for the corresponding service time, whose distribution function is given by

$$G_{j,t}^e(s) = \sum_{i=1}^J p_{ij}(t) G_i(s), \quad s \in \mathbb{R}^+, \quad 1 \leq j \leq J.$$

In the important special case that the true service times  $Y_i$  are deterministic and distinct (in which case we use  $S_i$  for the type  $i$  service time), we have

$$P\{Y_{j,t}^e = S_i\} = p_{ij}(t), \quad t \in \mathbb{R}^+, \quad 1 \leq i, \quad j \leq J. \tag{10}$$

Now let  $t$  be a decision epoch for the problem and let  $(\mathbf{n}, t)$  be the system state then. If action  $j \in \Delta(\mathbf{n})$  is taken and results in a service time (realised value of  $Y_{j,t}^e$ ) equal to  $s$  then the system state at the next decision epoch will be  $(\mathbf{n}', t + s)$  with probability  $p^e(\mathbf{n}' | \mathbf{n}, t, j, s)$  given by

$$\begin{aligned} p^e(\mathbf{n}' | \mathbf{n}, t, j, s) &= \prod_{m=1}^J \binom{n_m - \delta_{mj}}{n'_m} \{P[X_m^e \geq t + s | X_m^e > t]\}^{n'_m} \\ &\quad \{P[X_m^e < t + s | X_m^e > t]\}^{n_m - \delta_{mj} - n'_m} \\ &= \prod_{m=1}^J \binom{n_m - \delta_{mj}}{n'_m} \left\{ \frac{1 - F_m^e(t + s)}{1 - F_m^e(t)} \right\}^{n'_m} \\ &\quad \left\{ \frac{F_m^e(t + s) - F_m^e(t)}{1 - F_m^e(t)} \right\}^{n_m - \delta_{mj} - n'_m}, \\ &0 \leq n'_m \leq n_m - \delta_{mj}, \quad 1 \leq m \leq J. \end{aligned} \tag{11}$$

In (11),  $\delta_{mj}$  is the Kronecker delta which is equal to one when  $m = j$  and is zero otherwise. We can re-express the transition probability in (11) using the hazard in (9). We have

$$p^e(\mathbf{n}'|\mathbf{n}, t, j, s) = \prod_{m=1}^J \binom{n_m - \delta_{mj}}{n'_m} \left\{ \exp\left(-\int_t^{t+s} \theta_m^e(\tau) d\tau\right) \right\}^{n'_m} \left\{ 1 - \exp\left(-\int_t^{t+s} \theta_m^e(\tau) d\tau\right) \right\}^{n_m - \delta_{mj} - n'_m}, \quad 0 \leq n'_m \leq n_m - \delta_{mj}, \quad 1 \leq m \leq J. \quad (12)$$

In order to formulate the optimality equations for our Bayes sequential decision problem, we require the value function  $V^e: \Omega \rightarrow [0, N]$ , where  $\Omega$  is the system's state space, given by

$$\Omega = \{(\mathbf{n}, t); 0 \leq n_j \leq L_j, t \in \mathbb{R}^+\}.$$

The quantity  $V^e(\mathbf{n}, t)$  is the maximal expected number of service completions which can be delivered from system state  $(\mathbf{n}, t)$ . Note that the quantity  $V_\pi^e(\mathbf{L}, 0)$  is developed in (4)–(6) above. To obtain  $V_\pi^e(\mathbf{n}, t)$ , we modify that development as follows: Write  $N(t) = \sum_{j=1}^J n_j$  for the number of jobs remaining at  $t$  and represented in the system state  $(\mathbf{n}, t)$  and assume that these are reallocated numerical identifiers in the set  $\{1, 2, \dots, N(t)\}$  in ascending order of class identifier such that the class  $j$  jobs present at  $t$  are numbered  $K_{j-1}(t) + 1 \leq m \leq K_j(t)$ , where  $K_j(t) = \sum_{k=1}^j n_k, 1 \leq j \leq J$ . To obtain the appropriate posterior for the true unobservable state we now replace (4) by

$$P[\mathbf{I} = \mathbf{i} = \{i_m; 1 \leq m \leq N(t)\} | (\mathbf{n}, t)] = \prod_{j=1}^J \prod_{m=K_{j-1}(t)+1}^{K_j(t)} p_{imj}(t), \quad (13)$$

where by slight abuse of notation we now use  $\mathbf{I}$  for the true but unobservable state at time  $t$ . We now obtain  $V_\pi^e(\mathbf{n}, t)$  by replacing the (6) by one with respect to the posterior (13). We then have

$$V^e(\mathbf{n}, t) = \sup_{\pi} V_{\pi}^e(\mathbf{n}, t).$$

With the above in place we may write the DP optimality equations as

$$V^e(\mathbf{n}, t) = 1 + \max_{j \in \Delta(\mathbf{n})} \left\{ \int_0^\infty \sum_{\mathbf{n}'} p^e(\mathbf{n}'|\mathbf{n}, t, j, s) V^e(\mathbf{n}', t + s) dG_{j,t}^e(s) \right\}, \quad \mathbf{n} \neq 0 \quad (14)$$

and

$$V^e(\mathbf{0}, t) = 0.$$

In the special case of deterministic service times considered in (10), we can specialise (14) to

$$V^e(\mathbf{n}, t) = 1 + \max_{j \in \Delta(\mathbf{n})} \left\{ \sum_{i=1}^J \sum_{\mathbf{n}'} p_{ij}(t) p^e(\mathbf{n}'|\mathbf{n}, t, j, S_i) V^e(\mathbf{n}', t + S_i) \right\}, \quad \mathbf{n} \neq 0. \quad (15)$$

We further note that DP value iteration is available to us to compute the Bayes' returns associated with any specified service policy  $\pi$ . In this event we use the recursion

$$V_{\pi}^e(\mathbf{n}, t) = 1 + \left\{ \int_0^\infty \sum_{\mathbf{n}'} p^e(\mathbf{n}'|\mathbf{n}, t, \pi(\mathbf{n}, t), s) V_{\pi}^e(\mathbf{n}', t + s) dG_{\pi(\mathbf{n}, t), t}^e(s) \right\}, \quad \mathbf{n} \neq 0, \quad (16)$$

which in the case of deterministic service times becomes

$$V_{\pi}^e(\mathbf{n}, t) = 1 + \left\{ \sum_{i=1}^J \sum_{\mathbf{n}'} p_{i\pi(\mathbf{n}, t)} p^e(\mathbf{n}'|\mathbf{n}, t, \pi(\mathbf{n}, t), S_i) V_{\pi}^e(\mathbf{n}', t + S_i) \right\}, \quad \mathbf{n} \neq 0. \quad (17)$$

In the special case of classification without error (in which job class and type are identified), the simple static service policy which serves jobs in increasing order of the product  $E(X_j)E(Y_i)$  is known to work well when loss rates are low. See Glazebrook et al. (2004). We adapt this policy to the current case in which classification is error prone as follows: list the job classes in increasing order of the quantity  $E(X_j^e)E(Y_{j,0}^e)$ , i.e. such that

$$E(X_1^e)E(Y_{1,0}^e) \leq E(X_2^e)E(Y_{2,0}^e) \leq \dots \leq E(X_J^e)E(Y_{J,0}^e). \quad (18)$$

In any state  $(\mathbf{n}, t)$ , the proposed policy  $\pi^S$  chooses action  $\pi^S(\mathbf{n}, t)$  where

$$\pi^S(\mathbf{n}, t) = \min\{j; n_j \geq 1\}.$$

In what follows, the policy  $\pi^S$  will be both assessed as a policy in its own right and also used as a building block in the construction of strongly performing heuristic policies for our Bayes sequential decision problems. How we do this is described in the next section.

#### 4. On the development of effective heuristic service policies

For problems of realistic size, the utilisation of full DP to develop optimal service policies via suitable deployment of optimality Eq. (14) is computationally intractable. One possible route to the development of effective policies would be to apply a single DP policy improvement step to the simple static proposal  $\pi^S$ . The result would be a new dynamic policy  $\pi^{SPI}$  determined as follows:

$$\pi^{SPI}(\mathbf{n}, t) = \operatorname{argmax}_{j \in \Delta(\mathbf{n})} \left\{ \int_0^\infty \sum_{\mathbf{n}'} p^e(\mathbf{n}'|\mathbf{n}, t, j, s) V_{\pi^S}^e(\mathbf{n}', t + s) dG_{j,t}^e(s) \right\}, \quad \mathbf{n} \neq 0. \quad (19)$$

Policy  $\pi^{SPI}$  is constructed with the objective of choosing an optimal action in the current state, given that all future actions will be determined according to the simple static proposal  $\pi^S$ . Equivalently, the inaccessible value function  $V^e$ , required for the implementation of full DP in (14), is replaced by the approximation  $V_{\pi^S}^e$ . Sadly, the computation of this latter approximation is in many cases also intractable. Hence, we propose the deployment of a fluid approximation to the stochastic service system in order to develop  $V_{\pi^S}^{app}$ , an approximation to  $V_{\pi^S}^e$ . This approach, which extends that of the authors for the simpler perfect classification case, is now described.

In our fluid approximation, we fix  $j \in \{1, 2, \dots, J\}$  and represent the class  $j$  situation when its processing begins under static policy  $\pi^S$  by the pair  $(m_j, t_j)$ . In this representation,  $t_j$  is the time at which class  $j$  service begins under  $\pi^S$  and  $m_j$  is an amount of fluid representing the number of class  $j$  jobs surviving then. The nature of policy  $\pi^S$  means that class  $j$  will be served continually from  $(m_j, t_j)$  until all of the class  $j$  jobs are completed, namely until all of the corresponding fluid is drained in the approximating model. The process of draining class  $j$  fluid is as follows: if  $m_j \geq 1$  a single unit of fluid is removed instantaneously at time  $t_{j0} = t_j$  and signifies the guaranteed completion of a single job's service. Loss of fluid is thereafter experienced at rate  $\theta_j^e(s)$ , the hazard rate associated with class  $j$  jobs as given in (9), during the period of this initial class  $j$  service, which in the fluid model occupies the time interval  $[t_{j0}, t_{j1})$ , where  $t_{j1} = t_{j0} + E(Y_{j,t_{j0}}^e)$ . Should the amount of class  $j$  fluid remaining at  $t_{j1}$  exceed one then a further single unit of fluid is removed instantaneously then and signifies the guaranteed completion of a second class  $j$  service, and so on. In what follows,  $t_{jk}$  denotes the time of the  $k$ th class  $j$  service completion while  $Fl_j(s)$  denotes the amount of class  $j$  fluid remaining at time  $s$ . Class  $j$  fluid draining as service is offered to class  $j$  continuously from  $(m_j, t_j)$  is modelled as follows:

$$t_{j0} = t_j; t_{jk+1} = t_{jk} + E(Y_{j,t_{jk}}^e), \quad k \in \mathbb{N}, \quad (20)$$

$$Fl_j(t_{j0}) = m_j; Fl_j'(s) = -\theta_j^e(s)Fl_j(s), \quad s \notin \{t_{jk}, k \in \mathbb{N}\}; Fl_j(t_{jk}^+) = \{Fl_j(t_{jk}) - 1\}^+, \quad k \in \mathbb{N}.$$

We now introduce the quantities

$$K_j(m_j, t_j) = \min \{k; Fl_j(t_{jk}^+) = 0\} \quad (21)$$

and

$$N_j(m_j, t_j) = K_j(m_j, t_j) + Fl_j(t_{jK_j(m_j, t_j)}). \quad (22)$$

From (21),  $K_j(m_j, t_j)$  is the (integer) number of fully completed class  $j$  jobs under the fluid model while  $Fl_j(t_{jK_j(m_j, t_j)})$  is the fractional amount of class  $j$  fluid remaining at the conclusion of the class  $j$  processing and which is deemed to yield a further fractional completion within the approximating fluid model. In the fluid model, we take the total processing time of class  $j$  to be

$$T_j(m_j, t_j) = \sum_{k=0}^{K_j(m_j, t_j)-1} E(Y_{j,t_{jk}}^e) + Fl_j(t_{jK_j(m_j, t_j)})E(Y_{j,t_{jK_j(m_j, t_j)}}^e). \quad (23)$$

We now fix system state  $(\mathbf{n}, t)$  and use the quantities developed in (20)–(23) to develop  $V_{\pi^S}^{app}(\mathbf{n}, t)$ , the estimate of the expected number of job completions secured under static policy  $\pi^S$  from state  $(\mathbf{n}, t)$  obtained from our approximating fluid model. We define the quantities  $\zeta_j(\mathbf{n}, t)$ ,  $\psi_j(\mathbf{n}, t)$ ,  $1 \leq j \leq J$ , inductively as follows:

$$\zeta_1(\mathbf{n}, t) = N_1(n_1, t), \quad \psi_1(\mathbf{n}, t) = T_1(n_1, t)$$

and

$$\zeta_j(\mathbf{n}, t) = N_j \left( n_j \exp \left\{ - \int_0^{t_j} \sum_{k=1}^{j-1} \psi_k(\mathbf{n}, t) \theta_k^e(t+v) dv \right\}, t + \sum_{k=1}^{j-1} \psi_k(\mathbf{n}, t) \right), \quad 1 \leq j \leq J, \quad (24)$$

$$\psi_j(\mathbf{n}, t) = T_j \left( n_j \exp \left\{ - \int_0^{t_j} \sum_{k=1}^{j-1} \psi_k(\mathbf{n}, t) \theta_k^e(t+v) dv \right\}, t + \sum_{k=1}^{j-1} \psi_k(\mathbf{n}, t) \right), \quad 1 \leq j \leq J. \quad (25)$$

We now obtain the approximating value function as

$$V_{\pi^S}^{app}(\mathbf{n}, t) = \sum_{j=1}^J \zeta_j(\mathbf{n}, t). \quad (26)$$

Our proposed sequential decision rule  $\pi^{SF}$  is obtained by deploying the approximate value function from (26) within (19). We write,

$$\pi^{SF}(\mathbf{n}, t) = \arg \max_{j \in \Delta(\mathbf{n})} \left\{ \int_0^\infty \sum_{\mathbf{n}'} p^e(\mathbf{n}' | \mathbf{n}, t, j, s) V_{\pi^S}^{app}(\mathbf{n}', t+s) dG_{j,t}^e(s) \right\}, \quad \mathbf{n} \neq 0. \quad (27)$$

**Example 2.** For the problem with the  $p_i$ ,  $\epsilon_{ij}$  and lifetime distributions given in Example 1, and the following additional problem parameters

- $L_1 = L_2 = 5$ ,
- $S_1 = 1.41$ ,  $S_2 = 4.76$ ,

we computed both the exact value functions  $V_{\pi^S}$  and their fluid approximations  $V_{\pi^S}^{app}$ . In Fig. 1 we plot both values at a set of states at four time points which are chosen to be representative of likely decision epochs. It is shown that the performance of the approximation is outstanding, with consistently small errors across all the states examined. We have also calculated the percentage approximation errors,  $\Delta(V_{\pi^S}, V_{\pi^S}^{app}) = 100|1 - V_{\pi^S}^{app}/V_{\pi^S}|%$ , over the effective state space. Such a space excludes all the states which will never be visited given the initial conditions. Of 2267 points in the effective state space, the median error is only 1.48% and the worst

case is just 4.55%. The computational time to calculate the fluid approximation for each of these states does not vary much. The average is about 0.012 seconds. In sharp contrast, the time to calculate the exact value functions varies significantly with states, of which the maximum is 26 seconds. Please note that the computation of  $V_{\pi^S}$  grows exponentially in  $J$ , the number of job classes, while the growth of the computation of  $V_{\pi^S}^{app}$  is linear in  $J$ . Hence for larger problems we would expect the differences in computation time to be much greater.

### 5. Performance quality of the heuristic service policy $\pi^{SF}$

In what follows we compare the performance of the heuristic  $\pi^{SF}$  developed in the preceding section with that of the static proposal  $\pi^S$  described around (18). We shall also consider  $\pi^M$ , an adaptation of a myopic policy proposed by Argon et al. (2008) for the perfect classification case. In state  $(\mathbf{n}, t)$ ,  $\pi^M$  chooses the action from  $\Delta(\mathbf{n})$  to be the non-empty class  $j$  with the smallest associated value of

$$E(Y_{j,t}^e) \left[ \sum_{k=1}^J (n_k - \delta_{kj}) \{E(X_k^e - t | X_k^e > t)\}^{-1} \right].$$

The Bayes' returns for all of  $\pi^{SF}$ ,  $\pi^S$  and  $\pi^M$  are compared to the optimum for 18,000 randomly generated problems in each of which job lifetimes are Weibull and service times are deterministic. Recall that we use  $S_i$  for the (deterministic) service time of a type  $i$  job while for type  $i$  lifetimes we assume the following form of the Weibull hazard, namely

$$\theta_i(s) = \alpha_i \beta_i^{-\alpha_i} s^{\alpha_i-1}, \quad s \in \mathbb{R}^+.$$

All problems are chosen such that optimal policies can be developed via (15), notwithstanding the fact that exact value iteration is very expensive of computing time. The problems are generated at random under four different sets of assumptions (represented by categories A,B,C and D) regarding the relative lengths of service times and lifetimes of individual jobs and under three different sets of assumptions (poor, medium and good) regarding the quality of the initial job classification. Some problems involve just two job types ( $J = 2$ ) while for others there are four ( $J = 4$ ). The key problem features  $S_i$ ,  $\alpha_i$ ,  $\beta_i$  and  $N$  are sampled/chosen as follows:

$$S_i \sim U[1, 10] \quad (J = 2 \text{ cases}); \quad (28a)$$

$$S_i \sim U[1, 5] \quad (J = 4 \text{ cases}); \quad (28b)$$

$$\alpha_i \sim U[1, 2]; \quad (28c)$$

$$\beta_i \Gamma(1 + \alpha_i^{-1}) S_i^{-1} | S_i, \alpha_i \sim U[0.1, 0.5] \quad (\text{very short lifetimes, A}); \quad (28d)$$

$$\beta_i \Gamma(1 + \alpha_i^{-1}) S_i^{-1} | S_i, \alpha_i \sim U[0.5, 2] \quad (\text{short lifetimes, B}); \quad (28e)$$

$$\beta_i \Gamma(1 + \alpha_i^{-1}) S_i^{-1} | S_i, \alpha_i \sim U[2, 10] \quad (\text{moderate lifetimes, C}); \quad (28f)$$

$$\beta_i \Gamma(1 + \alpha_i^{-1}) S_i^{-1} | S_i, \alpha_i \sim U[10, 100] \quad (\text{long lifetimes, D}); \quad (28g)$$

$$N = 20 \quad (J = 2 \text{ cases}); \quad (28h)$$

$$N = 10 \quad (J = 4 \text{ cases}). \quad (28i)$$

Further, for each problem the  $p_i$  are obtained by first sampling independently from  $U[0.1, 0.9]$  and then normalising. The (mis) classification probabilities  $\epsilon_{ij}$  are obtained as follows: first obtain the probabilities of correct classification  $\epsilon_{ii}$  by sampling as follows:

$$\epsilon_{ii} \sim U[0.5, 0.65] \quad (\text{poor classification}); \quad (28j)$$

$$\epsilon_{ii} \sim U[0.65, 0.85] \quad (\text{medium classification}); \quad (28k)$$

$$\epsilon_{ii} \sim U[0.85, 1] \quad (\text{good classification}). \quad (28l)$$

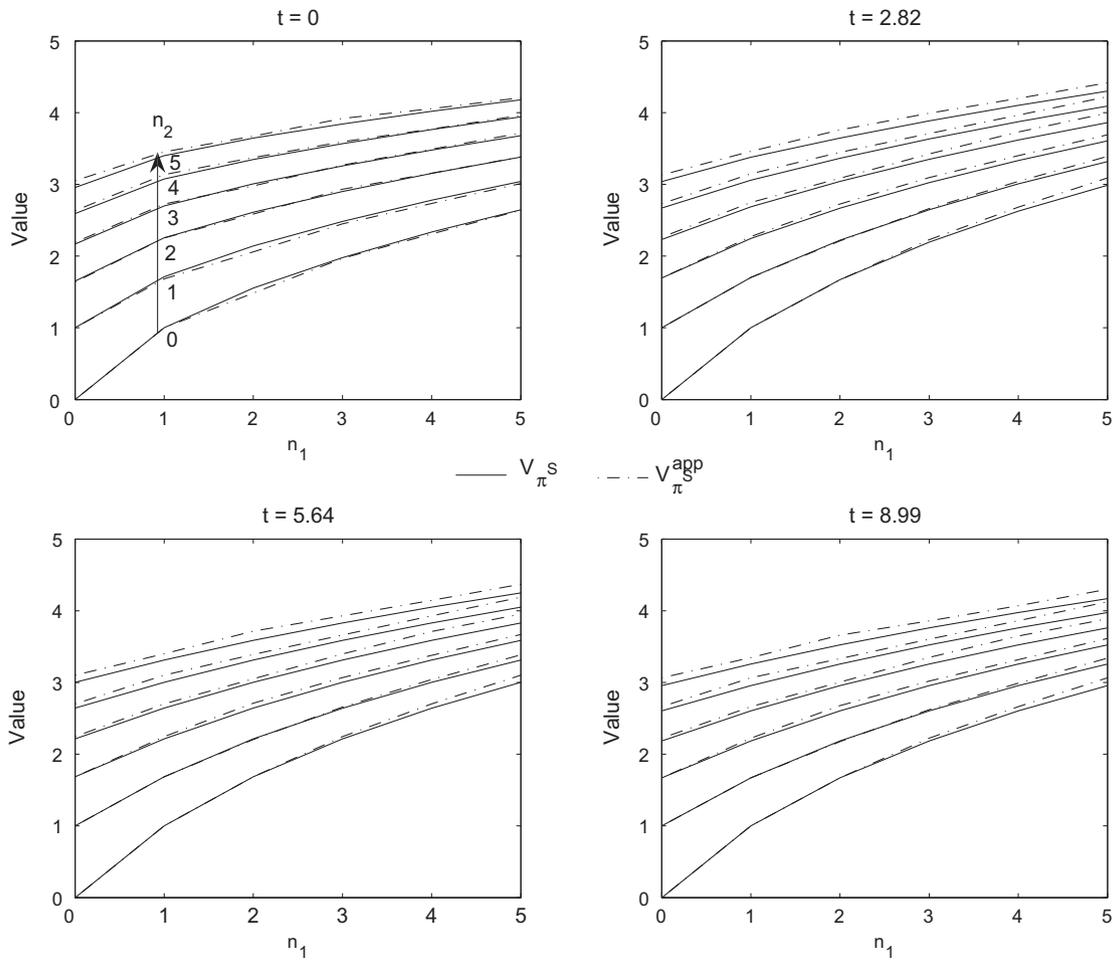


Fig. 1. Plots of the exact value function  $V_{\pi^S}$  and of its fluid approximation  $V_{\pi^S}^{app}$  for states  $(n_1, n_2, t)$  where  $0 \leq n_1, n_2 \leq 5$  and  $t = 0, 2.82, 5.64, 8.99$ .

Then obtain the  $\epsilon_{ij}, i \neq j$ , by sampling independently from  $U[0, 1]$  and normalising suitably.

Each sampled instance from the above is called a *profile*. For each profile, we generate a range of problems with different initial states  $(\mathbf{L}, 0)$ . This is done in two steps as follows: first, the number  $K_i$  of jobs of type  $i$  is obtained by sampling from the multinomial distribution

$$\{K_1, K_2, \dots, K_J\} \sim \text{Multinomial}(N; p_1, p_2, \dots, p_J). \tag{29}$$

Second, for each fixed  $i$ , the  $L_{ij}, 1 \leq j \leq J$ , namely the number of type  $i$  jobs classified initially as  $j$  is then obtained by sampling from the multinomial distribution

$$\{L_{i1}, L_{i2}, \dots, L_{ij}\} \sim \text{Multinomial}(K_i; \epsilon_{i1}, \epsilon_{i2}, \dots, \epsilon_{ij}). \tag{30}$$

These samples are drawn independently for distinct  $i$ . We obtain the components of initial state  $\mathbf{L}$  by setting  $L_j$ , the total number of jobs initially classified as  $j$ , equal to  $\sum_{i=1}^J L_{ij}$ . Please note that our sampling scheme is such that the  $L_j$  will tend to be smaller for the  $J=4$  cases studied than for the  $J=2$  ones. The job service times are also sampled from a smaller range. These choices are dictated by the computational requirements of the value iteration scheme needed for the determination of the maximal Bayes' return and the associated optimal policies. Please note that our heuristic policies can themselves be computed easily for much larger problems.

For  $J=2$ , 200 profiles were generated according to the above sampling scheme in (28a)–(28l) for each of the 12 combinations of problem category (A, B, C, D) and classification quality (poor, medium, good). For each profile, 5 problems (i.e., initial states)

were generated according to (29) and (30). Thus the total number of problems generated for each problem category/classification quality combination was 1,000, making 12,000 problems overall. For each problem, the quantities  $V_{\pi^{SF}}^e(\mathbf{L}, 0), V_{\pi^S}^e(\mathbf{L}, 0)$  and  $V_{\pi^M}^e(\mathbf{L}, 0)$ , the Bayes' returns respectively for the three heuristics  $\pi^{SF}, \pi^S$  and  $\pi^M$  were computed along with the maximal return  $V^e(\mathbf{L}, 0)$ . All computations used an appropriate form of DP value iteration from (15) and (17). For each heuristic  $\pi = \pi^{SF}, \pi^S, \pi^M$  and each problem  $(\mathbf{L}, 0)$  generated, the percentage suboptimality

$$\Delta_{\pi}(\mathbf{L}, 0) \equiv 100 \{V^e(\mathbf{L}, 0) - V_{\pi}^e(\mathbf{L}, 0)\} \{V^e(\mathbf{L}, 0)\}^{-1}$$

was computed. Further, for each subcollection of 1,000 problems corresponding to a problem category/classification quality combination, the minimum, mean and maximum values of  $\Delta_{\pi}(\mathbf{L}, 0)$  were computed for each heuristic. These values may be found in Table 1.

We observe that the policy  $\pi^{SF}$  developed by utilising an approximating fluid model within a one-step DP policy improvement performs robustly well throughout. Its mean percentage suboptimality never exceeds 0.03% with a worst case, among all 12,000 problems of just 1.21% suboptimal. It comfortably outperforms  $\pi^S$  and  $\pi^M$  excepting only the category B/poor classification case where it is marginally outperformed by  $\pi^M$ . Serious suboptimality is observed for the policies  $\pi^S$  and  $\pi^M$ , especially for problem configurations for which they were not designed.

Note that, within problem categories, percentage suboptimality tends to increase with classification quality. To understand this, consider the worst case for classification in which  $\epsilon_{ij} = J^{-1} \forall i, j$  in

**Table 1**  
Percentage deviation from optimal performance for Weibull lifetimes and deterministic service times when  $J = 2$ .

Category	Classification	Poor			Medium			Good		
		$\pi^{SF}$	$\pi^S$	$\pi^M$	$\pi^{SF}$	$\pi^S$	$\pi^M$	$\pi^{SF}$	$\pi^S$	$\pi^M$
A	Min	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	Mean	0.00	0.32	0.01	0.00	1.15	0.02	0.00	1.95	0.05
	Max	0.00	4.39	0.43	0.00	20.65	2.32	0.00	26.08	5.35
B	Min	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	Mean	0.03	0.12	0.01	0.02	0.89	0.05	0.01	1.53	0.12
	Max	0.76	4.15	0.27	0.91	11.36	2.06	0.69	20.93	7.31
C	Min	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	Mean	0.02	0.11	0.17	0.02	0.21	0.53	0.02	0.54	0.89
	Max	0.58	1.64	3.46	0.85	4.59	8.82	1.21	11.35	15.74
D	Min	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	Mean	0.00	0.04	0.18	0.00	0.12	0.62	0.01	0.25	1.24
	Max	0.07	0.73	3.77	0.12	1.64	7.88	0.99	3.24	13.24

which the classification process randomly allocates jobs to classes. In this case, the assessment process fails to offer useful information on job type and posterior probabilities (of true type) are independent of class. All jobs are effectively members of a single undifferentiated class and service policies which make use of class information are indistinguishable. As the classification improves from this worst case, the classes become more distinct, information on class membership more informative, policies more distinct and hence the choice of policy more important. It is thus unsurprising that the differences between the heuristic service policies is most pronounced when classification is reasonable or good. A cautionary note is that the small suboptimality when classification is poor do not necessarily indicate satisfactory service outcomes. It may well be that even the optimal policy cannot achieve a high Bayes' return when the classification errors are significant.

The study for problems with  $J = 4$  was conducted in the same way, except that its computational demands were such that only 100 profiles were generated for each of the problem category/classification quality combinations. Hence in this part of the study a further 6,000 problems were studied. The results are qualitatively very similar to those for the  $J = 2$  cases and may be found in Table 2.

**6. The cost of imperfect classification – analytical insight**

In the previous sections, we have described tools to develop  $\pi^{SF}$ , an effective and easily computed policy for the triage problem with imperfect classification which has been seen to achieve a Bayes' return close to the optimum in a large number of problems. We now explore the complementary question of the price paid in reduced

service completions for our inability to classify perfectly. We shall describe how we develop a measure of this, called the *cost of imperfect classification (CIC)*. It will assist the forthcoming discussion if (7) by  $\pi^B$ .

Consider initial system state  $(\mathbf{L}, 0)$ . Conditional upon  $(\mathbf{L}, 0)$ , the true state of the system (4). Condition upon the event  $\{\mathbf{I} = \mathbf{i}\}$ , namely that the true state of the jobs in the system is actually given by the vector  $\mathbf{i}$  of job types. Consider a decision-maker with access to the true state  $\mathbf{i}$ . We write  $v$  for a scheduling policy which is able to exploit that access and  $N(v|\mathbf{i})$  for the number of jobs served to completion under  $v$ . This decision-maker is able to achieve an expected number of job completions given by

$$\max_v E\{N(v|\mathbf{i})\},$$

where the expectation is taken with respect to realisations of the system evolving from  $\mathbf{i}$  under policy  $v$  and the maximum is over all policies which have access to the true state of the system throughout. A decision-maker who does not have access to the true state  $\mathbf{i}$  can do no better than use the Bayes' policy  $\pi^B$  which, conditional upon  $\mathbf{i}$ , yields an expected return

$$E\{N(\pi^B|\mathbf{i})\}.$$

Hence, conditional upon  $\mathbf{i}$ , the reduction in the expected number of jobs completed because of our failure to classify perfectly is given by

$$\max_v E\{N(v|\mathbf{i})\} - E\{N(\pi^B|\mathbf{i})\}. \tag{31}$$

When the initial system state is  $(\mathbf{L}, 0)$ , the measure of job completions lost we seek is obtained by taking an expectation of the (31)

**Table 2**  
Percentage deviation from optimal performance for Weibull lifetimes and deterministic service times when  $J = 4$ .

Category	Classification	Poor			Medium			Good		
		$\pi^{SF}$	$\pi^S$	$\pi^M$	$\pi^{SF}$	$\pi^S$	$\pi^M$	$\pi^{SF}$	$\pi^S$	$\pi^M$
A	Min	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	Mean	0.00	0.92	0.28	0.00	1.17	0.23	0.00	1.25	0.18
	Max	0.00	11.83	5.97	0.33	15.45	6.91	0.00	19.34	9.64
B	Min	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	Mean	0.06	1.00	0.08	0.07	1.49	0.10	0.06	1.63	0.14
	Max	2.01	9.82	1.87	3.14	10.83	3.66	2.02	17.06	6.26
C	Min	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	Mean	0.05	0.30	0.44	0.03	0.27	0.60	0.04	0.24	0.60
	Max	1.13	5.11	3.67	0.69	3.82	5.89	0.84	2.88	6.05
D	Min	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	Mean	0.03	0.23	0.42	0.01	0.29	0.40	0.02	0.36	0.69
	Max	0.96	2.74	3.19	0.41	2.87	3.26	0.60	2.63	4.87

with respect to the posterior distribution (4)). We thus obtain the quantity

$$E_{\mathbf{I}(\mathbf{L},0)} \left[ \max_v E\{N(v|\mathbf{I})\} - E\{N(\pi^B|\mathbf{I})\} \right] = E_{\mathbf{I}(\mathbf{L},0)} \left[ \max_v E\{N(v|\mathbf{I})\} \right] - V^e(\mathbf{L},0),$$

where the latter equation follows from (6) and (7). We now give a formal definition.

**Definition 1.** The cost of imperfect classification for initial state  $(\mathbf{L},0)$  is given by

$$CIC(\mathbf{L},0) \equiv E_{\mathbf{I}(\mathbf{L},0)} \left[ \max_v E\{N(v|\mathbf{I})\} \right] - V^e(\mathbf{L},0). \tag{32}$$

The relative cost of imperfect classification is given by,

$$RCIC(\mathbf{L},0) \equiv \frac{CIC(\mathbf{L},0)}{V^e(\mathbf{L},0)}. \tag{33}$$

**Remark 1.** Berry and Kertz (1991) developed a directly equivalent measure to CIC (which they called ‘the value of perfect information’) in the context of multi-armed bandit problems. For such problems they compute the value of perfect information by first considering the difference between the expected reward achieved by a player with perfect information regarding the success probabilities of each arm of the bandit (and who thus pulls the arm with the highest success probability on every occasion) and a player who has imperfect information of the success probabilities mediated through a prior distribution and who can do no better than use a Bayes’ policy for the problem. This difference is then averaged with respect to the prior concerned. This is exactly the approach taken here and which yields  $CIC(\mathbf{L},0)$ .

**Example 3.** For the problem with the  $p_i, L_j$  and lifetime and service time distributions given in Examples 1 and 2, we computed the relative cost  $RCIC$  with respect to different misclassification probabilities in the range 0.0–0.5. Note that  $\epsilon_{ij} = 0.5, i, j = 1, 2$  is the worst case when there are only two types of jobs. Fig. 2 below plots  $RCIC$ .

In Fig. 2 we see that the relative cost  $RCIC$  is continuous, increasing and concave componentwise in the misclassification probabilities  $\epsilon_{12}$  and  $\epsilon_{21}$  and is much more sensitive to the latter. A major factor here is the fact that type 1 jobs have much shorter lifetimes (with mean 1.45) than type 2 (mean 12.75), and are thus lost from the system much more quickly in any event. Should type 2 jobs be classified correctly and scheduled for processing appropriately, they are much more likely to be served to completion and contribute to the system’s return. Mathematically, the quantity  $E_{\mathbf{I}(\mathbf{L},0)} [\max_v E\{N(v|\mathbf{I})\}]$  in the expression for  $CIC(\mathbf{L},0)$  increases with  $\epsilon_{21}$  but decreases with  $\epsilon_{12}$ . For any initial state  $(\mathbf{L},0)$ , the higher the value of  $\epsilon_{21}$ , the more jobs are likely to actually be of type 2, and the higher are likely to be the number of successful completions. The quantity  $V^e(\mathbf{L},0)$  decreases with both misclassification probabilities.

In the special case of exponentially distributed lifetimes studied by Glazebrook et al. (2004), it is possible to gain analytical insight into system characteristics which impact the cost of imperfect classification  $CIC$  and which will inform our upcoming numerical study. We first state a simple result which will be of use in the analysis.

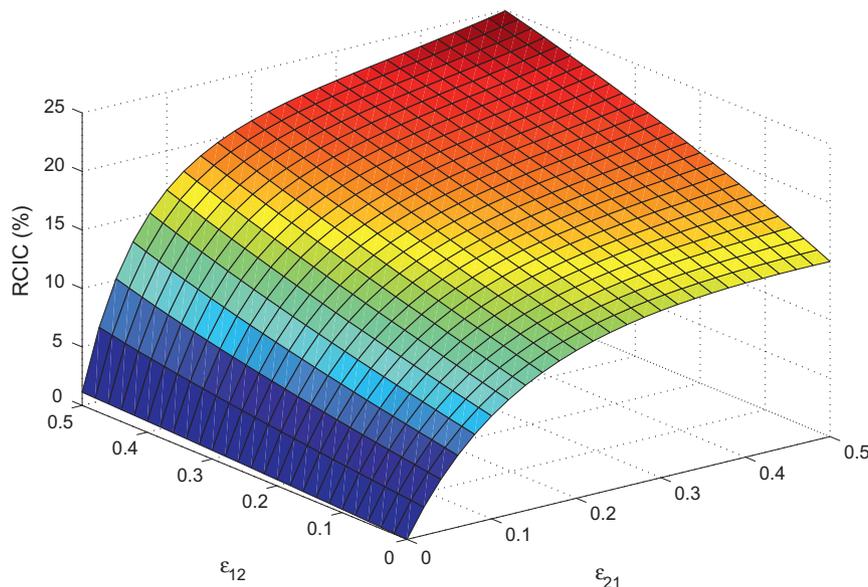
**Lemma 1**

$$E_{\mathbf{I}(\mathbf{L},0)} \left[ \max_v E\{N(v|\mathbf{I})\} \right] \geq V^e(\mathbf{L},0) \geq E_{\mathbf{I}(\mathbf{L},0)} \left[ \min_v E\{N(v|\mathbf{I})\} \right]. \tag{34}$$

In order to state our main result, we need some additional notation. In the exponential lifetime case we shall write the distribution of type  $i$  lifetimes as  $X_i \sim \exp(\rho\theta_i), 1 \leq i \leq J$ , where the  $\theta_i$  are taken to be fixed, and we shall be interested in the ‘no loss’ limit  $\rho \rightarrow 0$ . Recall that, for true but unobservable state  $\mathbf{I}$ , we write  $L_{ij}$  for the number of type  $i$  jobs classified as  $j$ . Hence for initial system state  $(\mathbf{L},0)$  we must have  $\sum_i L_{ij} = L_j, 1 \leq j \leq J$ . We then use the triple  $ijk$  for the  $k$ th type  $i$  job to be classified as  $j$  and  $Y_{ijk}$  for its service time. The random variable  $Y_{ijk}$  has the distribution function  $G_i$ . Hence we have

$$\mathbf{Y}(\mathbf{I}) = \sum_{i=1}^J \sum_{j=1}^J \sum_{k=1}^{L_{ij}} Y_{ijk}$$

for the total service requirement associated with true state  $\mathbf{I}$ .



**Fig. 2.** The relative costs  $RCIC$  for a problem with  $J = 2$ , Weibull lifetimes and deterministic service times.

Finally, we need the system parameter  $\delta$ , defined by

$$\delta \equiv \max_i \frac{\theta_i}{E(Y_i)} - \min_i \frac{\theta_i}{E(Y_i)}$$

$$= \rho^{-1} \left[ \max_i \{E(X_i)E(Y_i)\}^{-1} - \min_i \{E(X_i)E(Y_i)\}^{-1} \right].$$

**Theorem 1.** For the case of exponential lifetimes, we have

$$CIC(\mathbf{L}, 0) \leq \frac{1}{2} \delta \rho E_{\mathbf{I}(\mathbf{L}, 0)} [E\{\mathbf{Y}(\mathbf{I})\}^2] + O(\rho^2)$$

and

$$RCIC(\mathbf{L}, 0) \leq \frac{1}{2N} \delta \rho E_{\mathbf{I}(\mathbf{L}, 0)} [E\{\mathbf{Y}(\mathbf{I})\}^2] + O(\rho^2)$$

**Proof.** It follows from the above Lemma and from the definition of CIC that

$$CIC(\mathbf{L}, 0) \leq E_{\mathbf{I}(\mathbf{L}, 0)} \left[ \max_{\mathbf{v}} E\{N(\mathbf{v}|\mathbf{I})\} - \min_{\mathbf{v}} E\{N(\mathbf{v}|\mathbf{I})\} \right]. \quad (35)$$

It further follows from the analysis of Glazebrook et al. (2004), that when the true state is observable and has initial realised value  $\mathbf{I} = \mathbf{i}$  then the static service policy  $v^{SI}$  which serves the jobs in increasing order of the quantity  $E(X_i)E(Y_i)$  secures a return which is within an  $O(\rho^2)$  quantity of the maximum, namely

$$\max_{\mathbf{v}} E\{N(\mathbf{v}|\mathbf{i})\} - E\{N(v^{SI}|\mathbf{i})\} \leq O(\rho^2). \quad (36)$$

Similarly, it can easily be established that the static policy  $v^{SD}$  which serves the jobs in decreasing order of the quantity  $E(X_i)E(Y_i)$  secures a return which is within an  $O(\rho^2)$  quantity of the minimum, namely

$$E\{N(v^{SD}|\mathbf{i})\} - \min_{\mathbf{v}} E\{N(\mathbf{v}|\mathbf{i})\} \leq O(\rho^2). \quad (37)$$

To simplify the argument at this point, we relabel the jobs 1 to  $N$  such that, conditional upon  $\mathbf{I} = \mathbf{i}$ , the static policy  $v^{SI}$  is identified with the permutation  $(1, 2, \dots, N)$  and the static policy  $v^{SD}$  with  $(N, N-1, \dots, 1)$ . Now, the permutation  $(1, 2, \dots, N)$  can be obtained from  $(N, N-1, \dots, 1)$  by a series of  $\binom{N}{2}$  pairwise interchanges. At each stage, a permutation of the form  $(\dots, l, m, \dots)$ ,  $l > m$ , is transformed via a single interchange to  $(\dots, m, l, \dots)$ . From Glazebrook et al. (2004), the static service policies corresponding to these permutations have associated returns such that

$$E\{N(\dots, m, l, \dots|\mathbf{i})\} - E\{N(\dots, l, m, \dots|\mathbf{i})\}$$

$$= \left\{ \frac{\rho \theta_m}{E(Y_m)} - \frac{\rho \theta_l}{E(Y_l)} \right\} E(Y_l)E(Y_m) + O(\rho^2)$$

$$\leq \delta \rho E(Y_l)E(Y_m) + O(\rho^2). \quad (38)$$

Note that, in (38), by slight abuse of our notation, we have used subscripts  $l, m$  to identify quantities (hazard rates, service times) identified with particular jobs. If we now aggregate the impact on returns from all  $\binom{N}{2}$  pairwise interchanges, we infer from (38) that

$$E\{N(v^{SI}|\mathbf{i})\} - E\{N(v^{SD}|\mathbf{i})\} \leq \delta \rho \sum_{l > m} E(Y_l)E(Y_m) + O(\rho^2)$$

$$\leq \frac{1}{2} \delta \rho [E\{\mathbf{Y}(\mathbf{i})\}^2] + O(\rho^2). \quad (39)$$

We now infer from (36), (37) and (39) that

$$\max_{\mathbf{v}} E\{N(\mathbf{v}|\mathbf{I})\} - \min_{\mathbf{v}} E\{N(\mathbf{v}|\mathbf{I})\} \leq \frac{1}{2} \delta \rho [E\{\mathbf{Y}(\mathbf{I})\}^2] + O(\rho^2) \quad (40)$$

and the bound for  $CIC(\mathbf{L}, 0)$  now follows from (35) by applying the expectation  $E_{\mathbf{I}(\mathbf{L}, 0)}$  to both sides of (40). The bound for  $RCIC(\mathbf{L}, 0)$  uses that for  $CIC(\mathbf{L}, 0)$  together with the fact that

$$V^e(\mathbf{L}, 0) = N + O(\rho).$$

This concludes the proof.  $\square$

**Remark 2.** If we suppose that service time  $Y_i$  has mean and variance  $\mu_i$  and  $\sigma_i^2$  respectively then it is straightforward to show that the key quantity in the bounds given in Theorem 1 is given by

$$E_{\mathbf{I}(\mathbf{L}, 0)} [E\{\mathbf{Y}(\mathbf{I})\}^2] = \sum_{j=1}^J \sum_{i=1}^J L_j p_{ij}(0) [\sigma_i^2 + \mu_i^2 \{1 - p_{ij}(0)\}]$$

$$- \sum_{j=1}^J \sum_{i \neq k} L_j p_{ij}(0) p_{kj}(0) \mu_i \mu_k.$$

It follows from the above theoretical results that when job lifetimes are exponentially distributed and, moreover, long, the cost of imperfect classification will be small for problems in which the key quantity  $E(X_i)E(Y_i)$  varies little across distinct job types. We shall see in the upcoming numerical study that the insight afforded by these cases has much broader application. Note, for example, the following development of Example 3 above. We first introduce the quantity

$$R_{i_1 i_2} \equiv \frac{E(X_{i_1})E(Y_{i_1})}{E(X_{i_2})E(Y_{i_2})}$$

as the index ratio between job types  $i_1$  and  $i_2$ .

**Example 4.** In Example 3, should the service time for types of job 2 be decreased to 0.16 then  $R_{12} = R_{21} = 1$  and the relative cost of imperfect classification is drastically reduced. See Fig. 3, where the maximum of  $RCIC$  over the displayed range for the adjusted problem is now just 2.32%.

We now consider further the insights afforded by the above material in the next section.

## 7. The cost of imperfect classification in the worst case – a numerical study

We shall now explore the role of the above index ratio by numerically investigating the behaviour of the key quantities  $CIC$  and  $RCIC$  in the worst case, namely when  $\epsilon_{ij} = J^{-1} \forall i, j$ . We shall thus adopt a conservative viewpoint and ask questions about how much damage is done by a failure of the classification process to achieve anything better than random allocation of jobs to classes. The reader is referred back to the above definitions of  $CIC(\mathbf{L}, 0)$  and  $RCIC(\mathbf{L}, 0)$  around (32). Observe that in this worst case, the quantity  $V^e(\mathbf{L}, 0)$  which plays a key role in (32), can be obtained by computing the Bayes' return for any service policy. In the computations below, we shall in fact use the static proposal  $\pi^S$  for this purpose.

**Remark 3.** One approach to the computation of the key quantity  $E_{\mathbf{I}(\mathbf{L}, 0)} [\max_{\mathbf{v}} E\{N(\mathbf{v}|\mathbf{I})\}]$  is to estimate it via repeated sampling from the conditional distribution for  $\mathbf{I}(\mathbf{L}, 0)$ . We opted instead for an exact approach which computed  $[\max_{\mathbf{v}} E\{N(\mathbf{v}|\mathbf{I})\}]$  for each  $\mathbf{I}$  in the support of this distribution and then computed the exterior expectation. The reader should note that even though the quantity  $[\max_{\mathbf{v}} E\{N(\mathbf{v}|\mathbf{I})\}]$  is available in the numerical study reported here, it is challenging to obtain in general. Instead we can approximate it by computing returns for our heuristic policy  $\pi^{SF}$ , whose design is described in Section 4, to cases with perfect classification and whose initial state is summarised by  $\mathbf{I}$ . We know from Li and

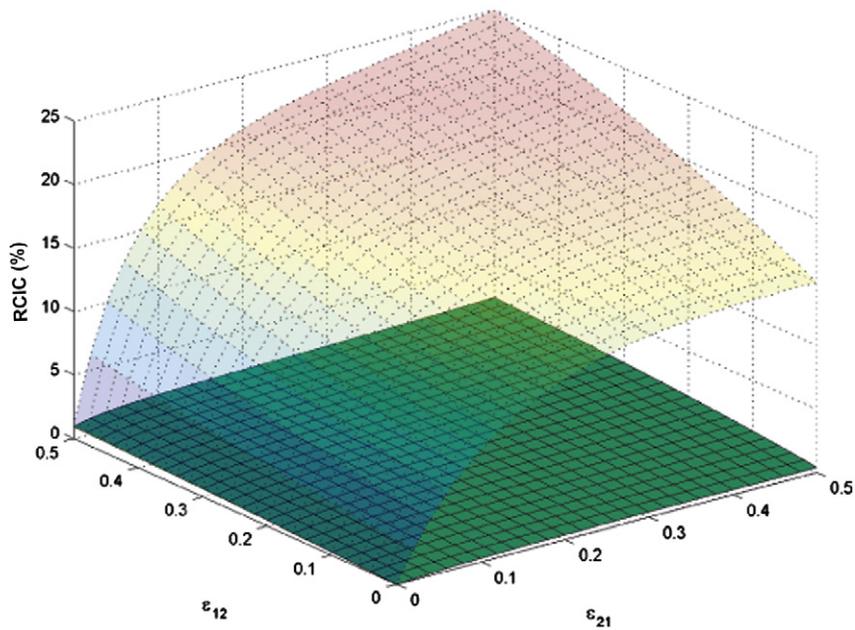


Fig. 3. The relative costs  $RCIC$  for a problem with  $J = 2$ , Weibull lifetimes and deterministic service times, and two different service rates of type 2 jobs.

Glazebrook (2010) that the performance of  $\pi^{SF}$  is very close to optimal in such cases and that any underestimate of the quantities  $CIC(L,0)$  and  $RCIC(L,0)$  which results will be small.

Our numerical study will consider problems with Weibull lifetimes and deterministic lifetimes and  $J = 2, 4$ . Problems will be created as in (28a)–(28i) above except for the sampling of  $\beta_j$ , which is conducted as follows. Firstly we sample  $\beta_1$  for the type 1 from one of

(28d)–(28g), according to the specified category. The remaining  $\beta_j$ 's are determined such that, conditioned on the sampled values of the  $\alpha_i, S_i$ , and  $\beta_1$  the index ratios  $R_{i+1}$  are all equal to some  $R$ , say, where without loss of generality we take  $R > 1$ . We then check the relative length of service times and the lifetimes for every type other than the first one. If all of them are in the same category as of type 1, one problem instance is created. Otherwise the entire process is repeated. We take four cases for the setting of an  $R$ -value which are

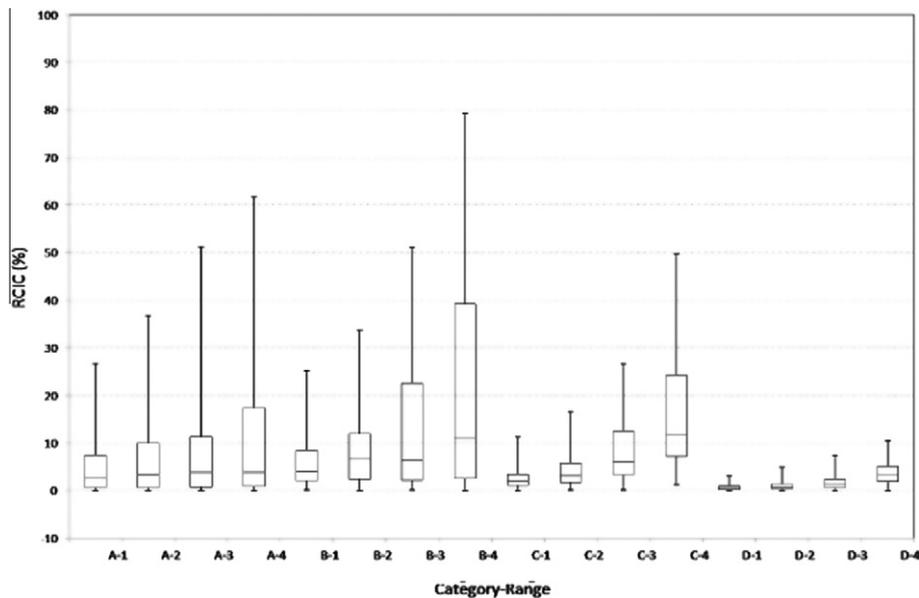


Fig. 4. Boxplot of the worst case relative costs  $RCIC$  for Weibull lifetimes and deterministic service times when  $J = 2$ .

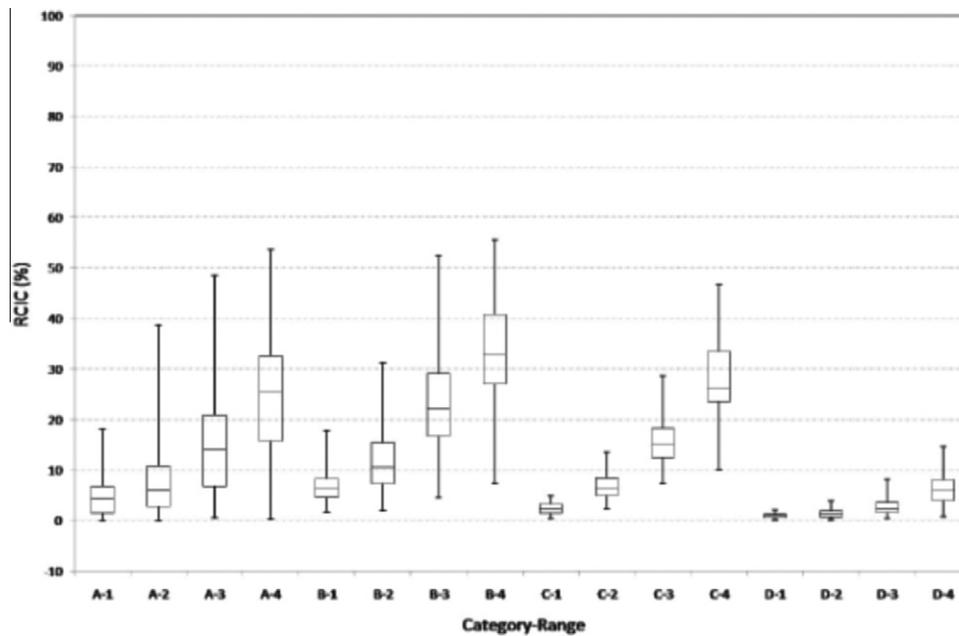


Fig. 5. Boxplot of the worst case relative costs  $RCIC$  for Weibull lifetimes and deterministic service times when  $J = 4$ .

$R \sim U[1, 1.1]$  (Range 1),  
 $R \sim U[1.1, 1.5]$  (Range 2),  
 $R \sim U[1.5, 2]$  (Range 3),  
 $R \sim U[2, 4]$  (Range 4).

Please note that Range 1 (respectively 4) allows the quantity  $E(X_4)E(Y_4)$  to be between 1 (respectively 8) and 1.331 (respectively 64) times as large as  $E(X_1)E(Y_1)$ . We thus investigate a wide range of cases, including some in which the values of  $E(X_i)E(Y_i)$  are nearly equal  $\forall i$  to others in which there can be very large differences. As above, the  $p_i$  will be obtained by first sampling independently from  $U[0.1, 0.9]$  and then normalising. However, we replace (28j)–(28l) by the choice  $\epsilon_{ij} = J^{-1} \forall i, j$ .

The computational effort needed here is very considerably greater than for the studies reported in Section 5. Happily, for the worst classification case we are considering the classification outcomes  $\mathbf{L}$  are immaterial. Nonetheless, for each imperfect classification problem we need to solve to obtain  $V^c(\mathbf{L}, 0)$ , there are  $\binom{N+J-1}{N}$  corresponding perfect classification problems whose solutions contribute to the quantity  $E_{\mathbf{I}(\mathbf{L}, 0)}[\max_v E\{N(v|\mathbf{I})\}]$ . This number increases very rapidly with  $N$  in the  $J = 4$  cases.

For  $J = 2$ , and each category (A, B, C, D)/range (1, 2, 3, 4) combination, 500 profiles were generated according to the above scheme, making 8,000 problems in total. Values of the relative costs  $RCIC$  are presented in Fig. 4 in the form of a boxplot for each of the 16 category/range combinations.

Computational demands are such that for the  $J = 4$  study, the number of profiles for each category/range combination was reduced to 100. Hence 1,600 problems with  $J = 4$  are investigated. The results are given in Fig. 5.

It is very clear from both figures that the values of  $RCIC$  do indeed increase markedly with  $R$ , this increase being rather more dramatic for the  $J = 4$  cases. This is as to be expected since the latter cases accommodate much greater variability in the type-specific values  $E(X_i)E(Y_i)$ . Another clear feature is the tendency of  $RCIC$  to decrease as lifetimes grow. To see why this might be expected, observe that in the limit in which no jobs are lost (lifetimes are infinite), all jobs will ultimately be served and no costs incurred

by any misclassification. The median values for  $RCIC$  are considerably larger for the  $J = 4$  cases reflecting the fact that misclassification has more impact when the number of job types is greater. That the maximum values are nevertheless reduced for  $J = 4$  is almost certainly due to the fact that the initial number of jobs in these problems is rather smaller (10 rather than 20), depressing the variability of the outcomes.

## Acknowledgement

The first author would like to thank Lancaster University Management School for a doctoral studentship to support his Ph.D. studies. Both authors would like to thank Mr. Kevin Martin for his valuable advice concerning some of the computations and two referees for their careful reading of the paper and helpful comments which led to significant improvements.

## References

- Argon, N.T., Ziya, S., 2009. Priority assignment under imperfect information on customer type identities. *Manufacturing and Service Operations Management* 11, 674–693.
- Argon, N.T., Ziya, S., Richter, R., 2008. Scheduling impatient jobs in a clearing system with insights on patients triage in mass casualty incidents. *Probability in the Engineering and Informational Sciences* 22, 301–332.
- Azoury, K.S., 1985. Bayes solution to dynamic inventory models under unknown demand distribution. *Management Science* 31, 1150–1160.
- Bellman, R., 1961. *Adaptive Control Processes: A Guided Tour*. Princeton University Press, Princeton, NJ.
- Berry, D.A., Kertz, R.P., 1991. Worth of perfect information in Bernoulli bandits. *Advances in Applied Probability* 23, 1–23.
- Boys, R.J., Glazebrook, K.D., Laws, D.J., 1996. A class of Bayes-optimal two-stage screens. *Naval Research Logistics* 43, 1109–1125.
- Cai, X., Wu, X., Zhou, X., 2009. Stochastic scheduling subject to preemptive-repeat breakdowns with incomplete information. *Operations Research* 57, 1236–1249.
- Frykberg, E.R., Tepas, J.J., 1988. Terrorist bombings Lessons learned from Belfast to Beirut. *Annals of Surgery* 208, 569–576.
- Furukawa, N., 1970. Fundamental theorems in a Bayes controlled process. *Bulletin of Mathematical Statistics* 14, 103–110.
- Gittins, J.C., 1979. Bandit processes and dynamic allocation indices (with discussion). *Journal of the Royal Statistical Society, Series B (Methodological)* 41, 148–177.
- Gittins, J.C., 1989. *Multi-armed Bandit Allocation Indices*. John Wiley and Sons, Chichester.

- Gittins, J.C., Glazebrook, K.D., 1977. Bayesian models in stochastic scheduling. *Journal of Applied Probability* 14, 556–565.
- Glazebrook, K.D., 1978. On the optimal allocation of two or more treatments in a controlled clinical trial. *Biometrika* 65, 335–340.
- Glazebrook, K.D., Ansell, P.S., Dunn, R.T., Lumley, R.R., 2004. On the optimal allocation of service to impatient tasks. *Journal of Applied Probability* 41, 51–72.
- Glazebrook, K.D., Boys, R.J., 1995. A class of Bayesian models for optimal exploration. *Journal of the Royal Statistical Society, Series B (Methodological)* 57, 705–720.
- Glazebrook, K.D., Owen, R.W., 1995. On the value of adaptive solutions to stochastic scheduling problems. *Mathematics of Operations Research* 20, 65–89.
- Hamada, T., Glazebrook, K.D., 1993. A Bayesian sequential single-machine scheduling problem to minimize the expected weighted sum of flowtimes of jobs with exponential processing times. *Operations Research* 41, 924–934.
- Kumar, P.R., 1985. A survey of some results in stochastic adaptive control. *SIAM Journal on Control and Optimization* 23, 329–380.
- Li, D., Glazebrook, K.D., 2010. An approximate dynamic programming approach to the development of heuristics for the scheduling of impatient jobs in a clearing system. *Naval Research Logistics* 57, 225–236.
- Rieder, U., 1975. Bayesian dynamic programming. *Advances in Applied Probability* 7, 330–348.
- Turégano-Fuentes, F., Pérez-Díaz, D., Sanz-Sánchez, M., Ortiz Alonso, J., 2008. Overall assessment of the response to terrorist bombings in trains, Madrid, 11 March 2004. *European Journal of Trauma and Emergency Surgery* 34, 433–441.
- Van der Zee, S.P., Theil, H., 1961. Priority assignment in waiting-line problems under conditions of misclassification. *Operations Research* 9, 875–885.
- Zhao, Y.J., Patek, S.D., Beling, P.A., 2008. Decentralized Bayesian search using approximate dynamic programming methods. *IEEE Transactions on Systems Man and Cybernetics Part B-Cybernetics* 38, 970–975.