International Journal of the Physical Sciences Vol. 7(3), pp. 413 - 422, 16 January, 2012
Available online at http://www.academicjournals.org/IJPS
DOI: 10.5897/IJPS11.1350
ISSN 1992 - 1950 ©2012 Academic Journals

Full Length Research Paper

A selection approach for solving buffer allocation problem

Mohammad H. Almomani[1]*, Rosmanjawati Abdul Rahman[1], Adam Baharum[1] and Mahmoud H. Alrefaei[2]

[1]School of Mathematical Sciences, Universiti Sains Malaysia, Penang, Malaysia.
[2]Department of Mathematics, Statistics and Physics, Qatar University, Doha, Qatar.

Accepted 8 December, 2011

In this paper, we dealt with one problem in designing a production line, which is the problem of buffer allocation. A selection approach was developed and tested for selecting the best design for a huge number of alternatives set. The proposed selection approach is a combination between cardinal and ordinal optimization. The algorithm involves four procedures; ordinal optimization, optimal computing budget allocation, subset selection and indifference-zone. The purpose of this paper is to use the proposed selection approach to find the optimal allocation of buffers that maximizes the mean production rate (throughput) in short, unbalanced and reliable production lines. Numerical results are presented to demonstrate the efficiency of the selection algorithm in finding the best buffer profile where its mean production rate is at its maximum.

Key words: Buffer allocation problem, production lines, ranking and selection, ordinal optimization, optimal computing budget allocation, optimization.

INTRODUCTION

Buffer allocation is important in the optimization problem involved with a production line. In buffer allocation problem (BAP), the objective is to allocate Q buffer spaces (slots) amongst the q intermediate buffers between $q+1$ machines in a production line, in order to meet some specified purpose. There are different BAP and they depend on the chosen objective function. Particularly, the performance measures for a production line are the mean production rate and the average work in process. The concern is to allocate Q buffer spaces, over q buffers in order to maximize the mean production rate.

The BAP is a difficult optimization problem because it is difficult or impossible to calculate the exact value of the objective function for a given allocation. Thus, the objective function for this optimization problem needs to be estimated (Chaharsooghi and Nahavandi, 2003). At the same time, the BAP is involved with a huge number of feasible allocations with respect to the total number of buffer spaces and the number of stations in the production line. Assume that there are Q buffer spaces available that need to be allocated over the q buffers, then there are $\binom{Q+q-1}{Q}$ different alternative designs. Each alternative is called "buffer profile", where it represents a unique combination of storage allocation and with a potential to result in a different output level of the line. For example, if the number of buffer spaces $Q=18$ and the number of buffers $q=5$ then there are 7315 different buffer profiles. The objective would be to select from this huge set of alternatives the best buffer profile that has the maximum mean production rate.

In this paper, the selection approach as proposed by Almomani and Abdul Rahman (2012) is used to solve the BAP by selecting the best buffer profile where its mean production rate is at its maximum. Almomani and Abdul Rahman (2012) proposed a new selection approach that selects a good design for large scale problems by using

*Corresponding author. E-mail: mh_momani@yahoo.com.

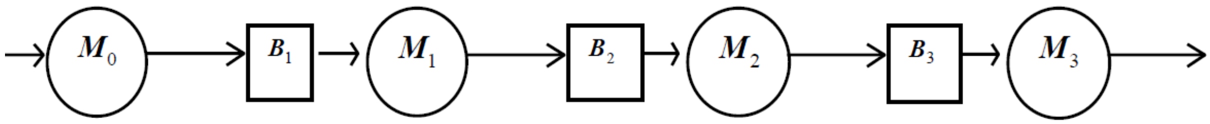

Figure 1. A production line of 4 machines and 3 buffers.

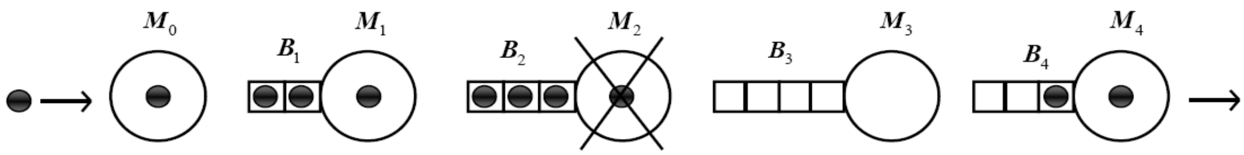

Figure 2. A production line with 5 machines, 12 buffer spaces and a buffer profile is (2 3 4 3).

the ordinal optimization and ranking and selection procedures. Their selection approach is a combination of four procedures; ordinal optimization (OO), optimal computing budget allocation (OCBA), subset selection (SS) and indifference-zone (IZ). The goal of their selection approach is to select a good design from a huge number of alternatives with high probability. The first step in this approach, involved with the OO to select a subset that overlaps with the set of the actual best $m\%$ design. Then, OCBA technique is used to allocate the available simulation samples in a way that maximize the probability of correct selection ($P(CS)$). This is followed by SS procedure to get a smaller subset that contains the best design among the subset that is selected before. Finally, the IZ procedure is used to select the best design among the survivor designs in the previous stage.

BUFFER ALLOCATION PROBLEM

Production designs are often organized with a queuing workstations or machines that are connected in a series and are separated by buffers. Figure 1 represents a production line of 4 machines, and this line is called a flow line, or a transfer line. Clearly in Figure 1, the circles represent machines and the squares represent buffers. The job moves in the direction of the arrows, from source inventory to the first machine for service, then to the first buffer where it waits until the second machine become available (empty) then it moves to the second machine, etc. until it finishes all the services (stations) in the queue and leaves the line.

The importance of buffers in the production designs comes from the fact that job flow may be affected by machine failures or by other variables during processing. Therefore, buffers are inserted between machines to limit the spread of disruptions, lessen congestion, decrease the negative effect of blocking and of course all these will cause an increase in the efficiency of the line. However, inserting buffers in the production line requires additional capital investment and floor space, which is expensive. Therefore, the challenge here is how to determine the buffer size or how to allocate the buffer spaces in a way that will achieve the desired performance. Note that, the buffers cannot be too large because; an increase in the buffer size usually will increase the total of work in progress, time to the customer, inventory and capital. The work-in-process inventory and capital costs incurred will outweigh the benefit of increasing productivity, so the buffers may be inadequate. On the other hand, the buffers cannot be too small because the machines will be untapped to meet demand.

There are two types of BAP; short and long lines as presented in Papadopoulos et al. (2009). The short line is a production line with up to 6 machines with a maximum of up to 20 buffer spaces, whereas the larger lines is otherwise. Furthermore, the BAP can be defined as balanced and unbalanced line, where the balanced line is a line with equal mean service time at each of the $q+1$ machine. Production line also, can be defined as a reliable or unreliable line, where in reliable line; each machine of the line cannot fail. An illustration of the definitions is given in Figure 2.

Figure 2 shows a short production line with $q+1=5$ machines represented by M_0, M_1, M_2, M_3, M_4, the intermediate buffers are represented by B_1, B_2, B_3, B_4 with $q=4$, with the total buffer spaces $Q=12$ and the current buffer allocation (buffer profile) is (2 3 4 3). Clearly, the first machine (M_0) has unlimited supply of jobs, and the last machine (M_4) has unlimited space, which means that (M_4) is never been blocked. Furthermore, the second machine (M_1) is currently been blocked, (M_2) has failed and is under repair and (M_3) is starved. Note that, each job enter the design from the first machine, passes in order through all machines and the intermediate buffer allocations and exits the line from the last machine. If the machine completed its service and the next buffer has space available, then the job will be passed on, and the machine will receive a new job from its input buffer. If the buffer has no jobs then the machine will remain empty until a new job is placed in the buffer. This type of production line is subject to manufacturing blocking and starving (blocking after service).

In this paper, it is assume that the production line as short, reliable and is unbalanced with unlimited supply of jobs in the first machine (the machine will never been starved) and unlimited space after the final machine (the machine will never be blocked). Jobs received service at each machine with the service times being independently random variables following the exponential distribution with rate μ_i, for $i = 0,1,\ldots,q$. With the model given previously, the objective is to maximize the mean production rate, subject to a given total buffer spaces (slots). In mathematical notation, this BAP can be stated as follows:

$$\max P(\underline{B})$$

$$s.t. \sum_{i=1}^{q} B_i = Q$$

$$B_i \geq 0 \ i = 1,2,\ldots,q$$

where, $P(\underline{B})$ is the production rate of the $q+1$ machine production line as a function of the buffer sizes vector (buffer profile); $\underline{B} = (B_1 \ B_2 \ \ldots \ B_q)$ is the buffer vector (to represent buffer profile), where B_i as integer for all $i = 1,2,\ldots,q$, and Q is a fixed nonnegative integer represents the available buffer spaces (slots) in the production line.

Other works related to the BAP, among others are Lutz et al. (1998) has solved the BAP by determining buffer location and size in production lines using a simulation-search heuristic procedure, which is based on tabu search, combined with simulation. Spinellis and Papadopoulos (2000) described a simulated annealing procedure for solving the BAP in reliable production lines, with the objective of maximizing the mean production line. Chaharsooghi and Nahavandi (2003) presented a heuristic algorithm to find the optimal allocation of buffers that maximizes mean production rate. Alon et al. (2005) presented a stochastic algorithm for solving the BAP, based on the cross-entropy method. Gershwin and Schor (2000) described the efficient approaches for determining how buffer spaces should be allocated in a flow line, with two types of problems; the primal and dual. Yuzukirmizia and Smithb (2008) proposed a new procedure to get a sub-optimal buffer profile for closed queuing networks with multiple servers and finite buffers. For more details about BAP js given in Alrefaei and Andradóttir (2005), Daskalaki and Smith (2004), Foley and Park (2002), Kim et al. (2002), Huang et al. (2002), Diamantidis and Papadopoloulos (2004), Roser et al. (2003), Malekian and Abdullah (2011) and Hedayati et al. (2011).

SELECTION APPROACHES WITH ORDINAL OPTIMIZATION

Statistical selection approaches are used to identify the best of simulated design from a finite set of simulation alternatives, with the best simulated design is defined in terms of the maximum (minimum) expected value of each alternative. This paper considers the problem of selecting the best design from a finite and large set of alternatives, where the expected value of each alternative can be inferred by simulation. This problem has been considered by several authors; a comprehensive review is given in Fu et al. (2005) who have considered the case where the objective function values cannot be evaluated exactly, but has to be estimated using simulation and the problem is then called a simulation optimization problem. Such problem is described as

$\max_{\theta \in \Theta} J(\theta)$ (1)

where $J(\theta) = E[L(\theta, Y)]$, with J, the expected performance measure of some complex stochastic design, Θ is an arbitrary feasible solution set, that finite and has no structure, θ is a vector representing the system design parameters, Y represents all the random effect of the design and L is a deterministic function that depends on θ and Y. When the feasible solution set Θ is small, then Ranking and Selection (R&S) procedures can be used for ranking the designs and selecting a subset that contains the best designs with a pre specified significance level. Unfortunately, R&S procedures are not applicable for large scale problems because it needs a huge computational time. Therefore, to reduce the computational effort, the idea of ordinal optimization (OO) proposed by Ho et al. (1992) will be used. However, with some changes in the objectives were instead of looking for the best design, the focus will be on finding a good enough design.

Ranking and selection

Selecting a design with the largest or smallest expected performance is one of the major problems that arise during a simulation. When the number of alternatives n is small, R&S procedures can be used to select the best design or a subset that contains the best design. There are two different R&S approaches; indifference-zone (IZ) and subset selection (SS). Suppose there are n alternative designs that are normally distributed with unknown means $\mu_1, \mu_2, \ldots, \mu_n$ and suppose that these means are ordered as $\mu_{[1]}, \mu_{[2]}, \ldots, \mu_{[n]}$ and the objective is to locate the design that has the best maximum mean $\mu_{[n]}$. In IZ, the correct selection (CS) is achieved by selecting a design that is within δ^* from the best, where δ^* is the indifference zone, with a pre-specified significance level P^*, P^* is the probability for the CS. In mathematical notation, it is to select an alternative i^* such that $\mu_{i^*} \in [\mu_{[n]} - \delta^*, \mu_{[n]}]$, and requires that $P(CS) \geq P^*$ where $1/n \leq P^* \leq 1$. For this problem, Rinott (1978) has proposed a two stage approach when the variances are completely unknown, and to use the IZ procedure, the number of alternatives n should be less than or equal 20.

However, when the number of alternatives is relatively large, SS procedures can be used. In this procedure, a small subset of design that contains the actual best design will be selected, and it is required that $P(CS) \geq P^*$ where $1/n \leq P^* \leq 1$. The SS approach dates back to Gupta (1965), who presented a single stage approach for producing a subset containing the best design with a specified probability. More details on the R&S approaches can be found in Bechhofer et al. (1995) and Kim and Nelson (2006a, b, 2007).

In real world problems, the number of alternatives is very huge (e.g. engineering applications), so the R&S approaches cannot be used to select the best designs, due to a huge computational time. Nelson et al. (2001) proposed a combination between SS and IZ procedures to obtain a computationally and statistically efficient approach for selecting the best design when the number of alternatives is large with the unknown variances. This procedure consists of two stages; in the first stage, SS procedure screened out and eliminated alternatives that are not competitive. This is followed by selecting the best system from the competitive alternative systems in the second stage (IZ procedure). Kim and Nelson (2001) proposed a fully sequential procedure, to select the best design when the number of alternative designs is large. They showed that their procedure works well for up to n = 500 design and it required unequal variances for all designs. Their procedure objective is to eliminate, at an early stage, those stochastic designs that are apparently inferior, in order to reduce the overall computational effort that required selecting the best system. Alrefaei and Almomani (2007) proposed two sequential algorithms for selecting a subset of k designs that is contained in the set of the top s designs.

Ordinal optimization and optimal computing budget allocation

The OO's goal is to isolate a subset of good designs with high probability and to reduce the required simulation time for discrete event simulation. The aim of this procedure, as proposed by Ho et al. (1992) is to find good designs, rather than to estimate the performance value of these designs accurately. Therefore, the OO procedure is used to select a subset that overlaps with the set of the actual best $m\%$ designs with high probability.

Suppose the correct selection is to select a subset G of g designs from the feasible solution set Θ that contains at least one of the top $m\%$ best designs. Since Θ is very huge, then the probability of correct selection is given by $P(CS) \approx \left(1 - \left(1 - \frac{m}{100}\right)^g\right)$. Furthermore, suppose that the correct selection is to select a subset G of g designs that contains at least r of the best s designs. If S is assumed to be the subset that contains the actual best s designs, then the probability of correct selection can be obtained using the hyper geometric distribution as $P(CS) = P(|G \cap S| \geq r) = \sum_{i=r}^{g} \frac{\binom{s}{i}\binom{n-s}{g-i}}{\binom{n}{g}}$. Since the number of alternatives is very large then the $P(CS)$ can be approximated by the binomial random variable, as $P(CS) \approx \sum_{i=r}^{g} \binom{g}{i} \left(\frac{m}{100}\right)^i \left(1 - \frac{m}{100}\right)^{g-i}$. More details of OO can be found in Deng and Ho (1999), Lee et al. (1999), Li et al. (2002), Zhao et al. (2005) and Ho et al. (2007).

Meanwhile, the OCBA technique was proposed to improve the performance of OO procedure by determining the optimal numbers of simulation samples for each design, instead of equally simulating all designs. The technique is used to determine the best simulation lengths for all simulation designs to reduce the total computation time. The goal of OCBA is to allocate the total simulation samples from all designs in a way that maximizes the probability of selecting the best design within a given computing budget (Chen et al., 2000, 1999; Banks, 1998). Let B be the total sample that is available for solving the optimization problem given in (1). The target is to allocate these computed simulating samples to maximize the $P(CS)$, written in mathematical notation as follows:

$$\max_{T_1,\ldots,T_n} P(CS)$$
$$s.t. \sum_{i=1}^{n} T_i = B$$
$$T_i \in \mathbb{N} \; i = 1,2,\ldots,n$$

where, $\mathbb{N}$ is the set of non-negative integers, T_i is the number of samples allocated to design i, and $\sum_{i=1}^{n} T_i$ denotes the total computational samples, and assume that the simulation times for different designs are roughly the same. To solve this problem, Chen et al. (2000) proposed the following theorem.

Theorem 1

Given a total number of simulated samples B to be allocated to n competing designs whose performance is depicted by random variables with means $\mathrm{J}(\theta_1), \mathrm{J}(\theta_2), \ldots, \mathrm{J}(\theta_n)$, and finite variances $\sigma_1^2, \sigma_2^2, \ldots, \sigma_n^2$

respectively, as $B \rightarrow \infty$, the approximate probability of correct selection can be asymptotically maximized when

1. $\frac{T_i}{T_j} = \left(\frac{\sigma_i/\delta_{b,i}}{\sigma_j/\delta_{b,j}}\right)$; *where* $i, j \in \{1,2,\dots,n\}$ *and* $i \neq j \neq b$.
2. $T_b = \sigma_b \sqrt{\sum_{i=1, i\neq b}^{n} \frac{{T_i}^2}{{\sigma_i}^2}}$

where $\delta_{b,i}$ is the estimated difference between the performance of the two designs ($\delta_{b,i} = \overline{J_b} - \overline{J_i}$), and $\overline{J_b} \leq \min_i \overline{J_i}$ for all i. Here $\overline{J_i} = \frac{1}{T_i}\sum_{j=1}^{T_i} Y_{ij}$, where Y_{ij} is a sample from Y_i for $j = 1,\dots,T_i$.

ALGORITHM OF THE SELECTION APPROACH

The selection approach as proposed by Almomani and Abdul Rahman (2012) includes a combination of OO, OCBA, SS and IZ procedures. Initially, using OO procedure, a subset G is randomly selected from a feasible solution set that overlaps with the set that contains the actual best $m\%$ designs with high probability $(1-\alpha_1)$. Then OCBA procedure is used to allocate the available computing budget. This is followed with SS procedure to get a smaller subset I with probability as high as $(1-\alpha_2)$, that contains the best design among the previous selected subset. Finally, IZ procedure is applied to select the best design from that set I with high probability $(1-\alpha_3)$. The algorithm of the selection approach is as follows:-

1. Setup: Specify g and k where $|G| = g$, and $|G'| = k$. Consider the number of initial simulation samples $t_0 \geq 2$ with the indifference zone δ^*, and $t = t_{(1-\alpha_2/2)^{1/g-1}, t_0-1}$ from the t-distribution. Let $T_1^l = T_2^l = \cdots = T_g^l = t_0$, and determine the total computing budget B. Note that; G is the selected subset from Θ, that satisfies $P(G$ contains at least one of the best $m\%$ designs) $\geq 1-\alpha_1$, whereas G' is the selected subset from G, where $g \geq k$. Here the iteration number is represented as l.

Select a subset G of size g randomly from Θ, and also take a random sample t_0 of observations y_{ij} $(j = 1,\dots,t_0)$ for each design i in G, where $i = 1,\dots,g$.

2. Initialization: Calculate the sample mean ${\bar{y}_i}^{(1)}$ and variances ${s_i}^2$, where ${\bar{y}_i}^{(1)} = \frac{\sum_{j=1}^{T_i^l} y_{ij}}{T_i^l}$ and ${s_i}^2 = \frac{\sum_{j=1}^{T_i^l}\left(y_{ij} - {\bar{y}_i}^{(1)}\right)^2}{T_i^l - 1}$, for all $i = 1,\dots,g$. Order the designs in G according to their sample averages; ${\bar{y}_{[1]}}^{(1)} \leq {\bar{y}_{[2]}}^{(1)} \leq \cdots \leq {\overline{y}_{[g]}}^{(1)}$. Then select the best k designs (with the largest mean) from the set G, and represent this subset as G'.

3. Stopping rule: If $\sum_{i=1}^{g} {T_i}^l \geq B$, then stop. Otherwise, randomly select a subset G'' of the $g-k$ alternatives from $\Theta - G'$, let $(G = G' \cup G'')$.

4. Simulation budget allocation: Increase the computing budget by Δ and compute the new budget allocation, $T_1^{l+1}, T_2^{l+1}, \dots, T_g^{l+1}$ by using Theorem 1. Perform additional $\max\{0, T_i^{l+1} - T_i^l\}$ simulations for each design i, $i = 1,\dots,g$, let $l \leftarrow l+1$. Go to Initialization.

5. Screening: Set $I = \{i: 1 \leq i \leq k$ and ${\bar{y}_i}^{(1)} \geq {\bar{y}_j}^{(1)} - [W_{ij} - \delta^*]^+, \forall\, i \neq j\}$ where $W_{ij} = t\left(\frac{{s_i}^2}{T_i} + \frac{{s_j}^2}{T_j}\right)^{1/2}$ for all $i \neq j$, and $[x]^+ = x$ if $x > 0$ and $[x]^+ = 0$ otherwise. If I contains a single index, then this design is the best design. Otherwise, for all $i \in I$, compute the second sample size $N_i = \max\left\{T_i, \left\lceil\left(\frac{hs_i}{\delta^*}\right)^2\right\rceil\right\}$ where $h = h(1-\alpha_3, t_0, |I|)$ be the Rinott (1978) constant and can be obtained from tables of Wilcox (1984). Take additional $N_i - T_i$ random samples of y_{ij} for each design $i \in I$, and compute the overall sample means for $i \in I$; ${\bar{y}_i}^{(2)} = \frac{\sum_{j=1}^{N_i} y_{ij}}{N_i}$. Select design $i \in I$ with the largest ${\bar{y}_i}^{(2)}$ as the best.

Nelson et al. (2001) have shown that with probability at least $1-(\alpha_2+\alpha_3)$, the selection approach will select the best design from the subset G. Therefore, if G contains at least one of the top $m\%$ designs, then the selection approach selects a good design with probability $1-(\alpha_2+\alpha_3)$. On the other hand, from the OO procedure, it shows that the selected set G contains at least one of the best $m\%$ designs with probability of $(1-\alpha_1) = 1-\left(1-\frac{m}{100}\right)^g$. Therefore, P(the selected design from the selection approach is in the top m% designs) $\geq \Big(1-\Big(1-$ *m100g1−(α2+α3)≥1−1−m100g+ α2+α3.*

NUMERICAL EXAMPLE

Here, the specified BAP is solved by using Almomani and Abdul Rahman (2012) approach. A production line involving $q+1$ machines $M_0, M_1, \dots, M_q$, modeled as a single server queuing stations, with q intermediate buffers $B_1, B_2, \dots, B_q$ have been considered as shown in Figure 3. Assume that there are unlimited supply of jobs in front of machine M_0 and unlimited space after machine M_q. Jobs receive service at each machine with the service times at machine M_i are being independent and exponentially distributed with rate μ_i, for all $i = 1,2,\dots,q$.

The blocking scheme have been considered as follows: When a job j receives service at machine M_{i-1}, job j attempts to enter buffer B_i. If the buffer B_i is full, then job j will be forced to stay at machine M_{i-1} until it finds a space in the buffer B_i. During this period of time,

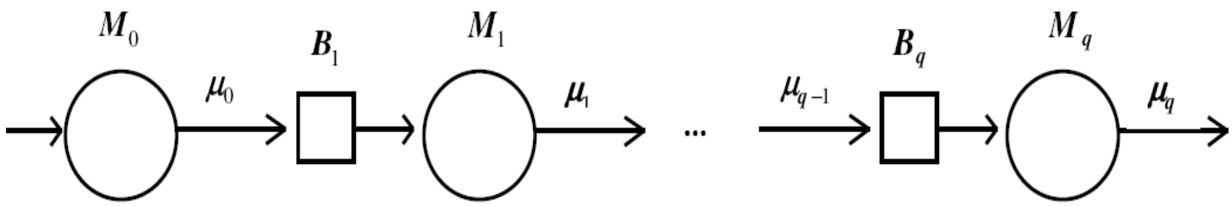

Figure 3. A production line with q+1 machines, q buffers, unlimited supply jobs in front of machine M_0, and unlimited room for all jobs departing from machine M_q.

Table 1. The implementation of the algorithm given the parameters $n = 3876, g = 50, k = 10, \Delta = 30, m\% = 5\%,\ t_0 = 10, B = 5000$.

Replication	$\sum_{i=1}^{g} T_i$	$\sum_{i \in I} N_i$	Best	Bufered profile	P(x)	E(OC)
1	95687	229462	3290	(6 2 4 2 1)	3.81454	0.12760
2	98889	232659	3037	(5 2 6 1 1)	3.82115	0.12099
3	107524	241294	2834	(4 6 2 3 0)	3.87345	0.06869
4	106755	240530	3389	(7 0 0 7 1)	3.80967	0.13247
5	117284	251055	3489	(7 2 5 1 0)	3.90003	0.04211
6	96639	230410	3140	(5 6 3 1 0)	3.86599	0.07615
7	108598	242366	2330	(3 3 4 4 1)	3.77622	0.16592
8	110278	244041	3126	(5 5 5 0 0)	3.90624	0.03590
9	103583	237350	3359	(6 5 3 0 1)	3.86138	0.08076
10	101249	235020	1966	(2 6 4 3 0)	3.89896	0.04318

machine M_{i-1} is blocked and cannot start processing another job until job j moves to the next machine. Assume that there are Q buffer spaces (slots) available that need to be allocated over the q buffers in order to maximize the mean production rate. It can be shown that there are $\binom{Q+q-1}{Q}$ different designs for such systems. Here, buffers are allowed to have zero size. The interest would be in selecting a design that gives a maximum mean production rate. In other words, is tried to solve the following maximization problem:

$$\max_{x \in \Theta} P(x)$$

where $P(x)$ is the production rate of the design, given that design x is being used and Θ is the set of all $\binom{Q+q-1}{Q}$ possible designs (alternatives). Here, assume that the production line is a reliable line.

The aforementioned BAP will be solved by using Almomani and Abdul Rahman (2012) selection approach with two different parameter settings. In the first setting, assume that there are $Q = 15$ buffer spaces to be allocated over $q = 5$ buffers. Thus, there are 6 machines and Θ contains 3876 different designs ($|\Theta| = n$ =3876), and assume that $\mu_0 = \mu_1 = \mu_2 = \mu_3 = 5$ and $\mu_4 = \mu_5 = 10$, which means that, the production line in this example is unbalanced line. Furthermore, size of set G is $g = 50$, size of set G' is $k = 10$, number of initial simulation samples $t_0 = 10$, total computing budget $B = 5000$, indifference zone $\delta^* = 0.05$, increment in simulation samples $\Delta = 30$ and $\alpha_2 = \alpha_2 = 0.005$. Suppose that the goal is selecting the design from the best *5%* designs in the set Θ. Therefore, the correct selection here will be the selected design i^* that belongs to set $\{x_1, x_2, \ldots, x_{193}\}$, where x_i, $i = 1,2, \ldots, 193$ representing the top designs that have the maximum mean production rate in the set Θ. The analytical probability of correct selection is P(CS) $\geq$ $1 - \left((1-0.05)^{50} + 0.005 + 0.005\right) \geq 0.91$. Table 1 contains the results of this experiment with 10 replications, where $\sum_{i=1}^{g} T_i$ is the total sample size used in stopping rule step in the algorithm of Almomani and Abdul Rahman (2012), $\sum_{i \in I} N_i$ is the total sample size used in screening step, "Best" means the index of the chosen design that is being considered as the best design, $P(x)$ is the production rate for each design and $E(OC)$ represents the expected opportunity cost of a potential incorrect selection (He et al., 2007). The $E(OC)$

Table 2. The implementation of the algorithm given the parameters $n = 10626, g = 100, k = 20, \Delta = 50, m\% = 3\%, t_0 = 10, B = 10000$.

Replications	$\sum_{i=1}^{g} T_i$	$\sum_{i \in I} N_i$	Best	Bufered profile	P(x)	E(OC)
1	513313	813192	5487	(3 6 8 1 2)	4.04343	0.06410
2	471731	771610	8795	(7 10 2 1 0)	4.02004	0.08749
3	511493	811376	6515	(4 6 5 5 0)	4.03954	0.06799
4	623511	923376	5553	(3 7 7 1 2)	4.03508	0.07245
5	574317	874196	7506	(5 9 4 1 1)	3.99606	0.11147
6	610636	910506	8628	(7 4 5 1 3)	4.00918	0.09835
7	533611	833489	8119	(6 6 5 2 1)	4.06380	0.04373
8	500361	800240	8064	(6 5 4 3 2)	4.06569	0.04184
9	544412	844286	6353	(4 4 5 2 5)	4.02684	0.08069
10	500882	800755	9171	(8 5 4 3 0)	4.03886	0.06867

values are defined as the absolute difference between the mean production rate achieved by Almomani and Abdul Rahman (2012) approach and the maximum mean production rate. So, $E(OC) = |P(x_b) - P(x_{i^*})|$, where x_b is the best design and x_{i^*} is the design selected by the Almomani and Abdul Rahman (2012) approach.

From Table 1, note that, in the first replication, Almomani and Abdul Rahman (2012) algorithm selected the design numbered 3290 with buffer profile (6 2 4 2 1) and the estimated production rate is 3.81454. It means in this replication, the maximum production rate was achieved when the buffer spaces are allocated on the buffers as follows; the buffer spaces in B_1 is 6, the buffer spaces in B_2 is 2, the buffer spaces in B_3 is 4, the buffer spaces in B_4 is 2 and the buffer spaces in B_5 is 1. For a comparison, we have simulated all the 3876 designs for a long simulation run and found that the best design is 2816 with buffer profile (4 5 4 2 0) and the mean production rate is 3.94214. Clearly, the production rate for the selected design is very closed to the production rate for the best design. Note also that the $E(OC)$ value in this replication is 0.12760 which is too small. Since, the selected design in the first replication belongs to the best *5%* designs from the set of 3876 designs, so it is consider as a correct selection.

Using the same algorithm (Almomani and Abdul Rahman, 2012), the second settings have considered buffer spaces $Q = 20$, that is to be allocated over $q = 5$ buffers. Here, the set of all alternatives Θ contains 10626 different designs. The algorithm is applied with the other following parameters such as; $n = 10626, g = 100, k = 20, t_0 = 10, B = 10000, \delta^* = 0.05, \Delta = 50$ and $\alpha_2 = \alpha_2 = 0.005$. Now, the goal is selecting the design from the best *3%* designs from set Θ. Therefore, the correct selection here is selecting design i^* that belongs to the set $\{x_1, x_2, \dots, x_{318}\}$, where x_i, $i = 1,2,\dots,318$ represents the top designs in the set Θ. The analytical probability of correct selection is $P(CS) \geq 1 - \left((1 - 0.03)^{100} + 0.005 + 0.005\right) \geq 0.94$. The results of the first 10 replications of this experiment are recorded in Table 2. All the 10626 designs have been simulated for a long simulation runs and found that the best design is numbered 7394 with a buffer profile (5 6 6 2 1) with mean production rate been 4.10753.

The first replication in Table 2 shows that, Almomani and Abdul Rahman (2012) algorithm has selected the design numbered 5487 with buffer profile (3 6 8 1 2) with the estimated production rate been 4.04343. It means that the maximum production rate in this replication was achieved according to the following buffer spaces allocation: The buffer spaces in B_1 is 3, the buffer spaces in B_2 is 6, the buffer spaces in B_3 is 8, the buffer spaces in B_4 is 1 and the buffer spaces in B_5 is 2. Clearly, it shows that the production rate for the selected design is closed to the production rate for the best design, and the $E(OC)$ value in this replication is 0.06410 which is small. However, the design that has been selected in the first replication does not belong to the best *3%* designs from the set of 10626 designs, so it was an incorrect selection.

This two experiments are then repeated for 100 replications, and the results are summarized in Table 3, where $\overline{\sum_{i=1}^{g} T_i}$ represents the average number of the total sample size in the "stopping rule", $\overline{\sum_{i \in I} N_i}$ is the average number of the total sample size in "screening" and the $\overline{E(OC)}$ is the average number of the $E(OC)$. Clearly, Almomani and Abdul Rahman (2012) approach selected the best buffer profile with high $P(CS)$ and its value is close to the analytical values. In the same time, the number of simulation samples used are relatively small, and the $E(OC)$ is small too.

Figures 4 and 5 show the $E(OC)$ for the selection approach when $n = 3876$ and $n = 10626$ respectively, for

Table 3. The performance of the selection algorithm over 100 replications.

Number of buffers	Buffer spaces	n	$\overline{\sum_{i=1}^{g} T_i}$	$\overline{\sum_{i \in I} N_i}$	$\overline{E(OC)}$	Approach P(CS)	Analytical P(CS)
5	15	3876	104845	238618	0.102443	0.86	0.91
5	20	10626	533543	833271	0.0910697	0.81	0.94

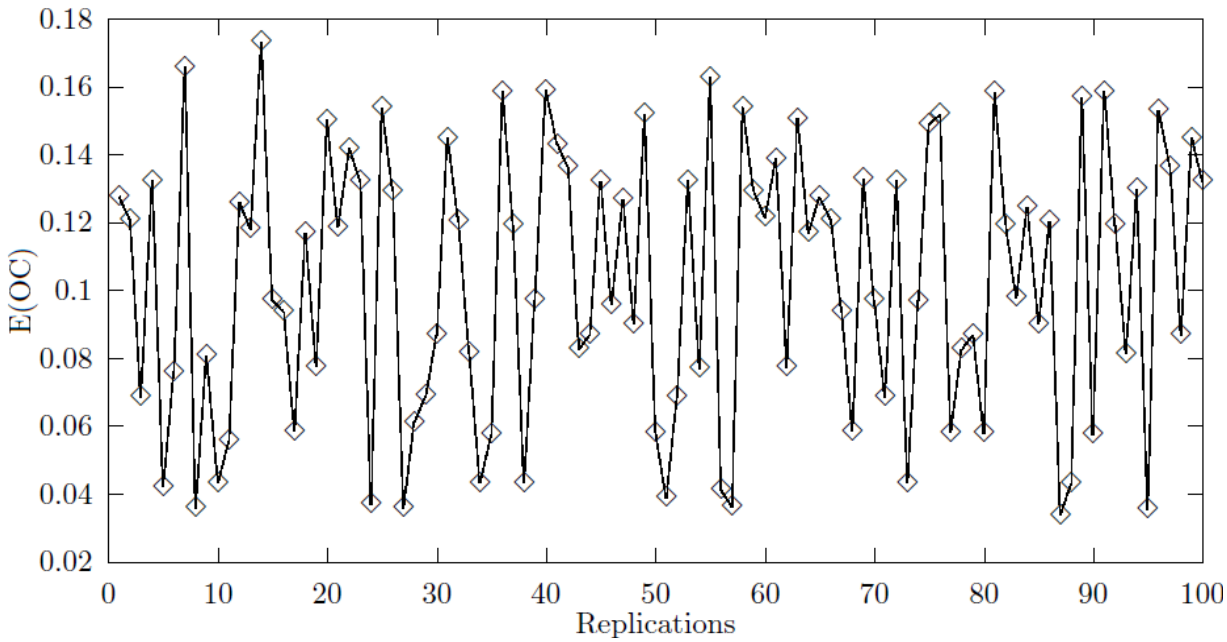

Figure 4. The $E(OC)$ for the selection approach when $n = 3876, g = 50, m\% = 5\%$ over 100 replications.

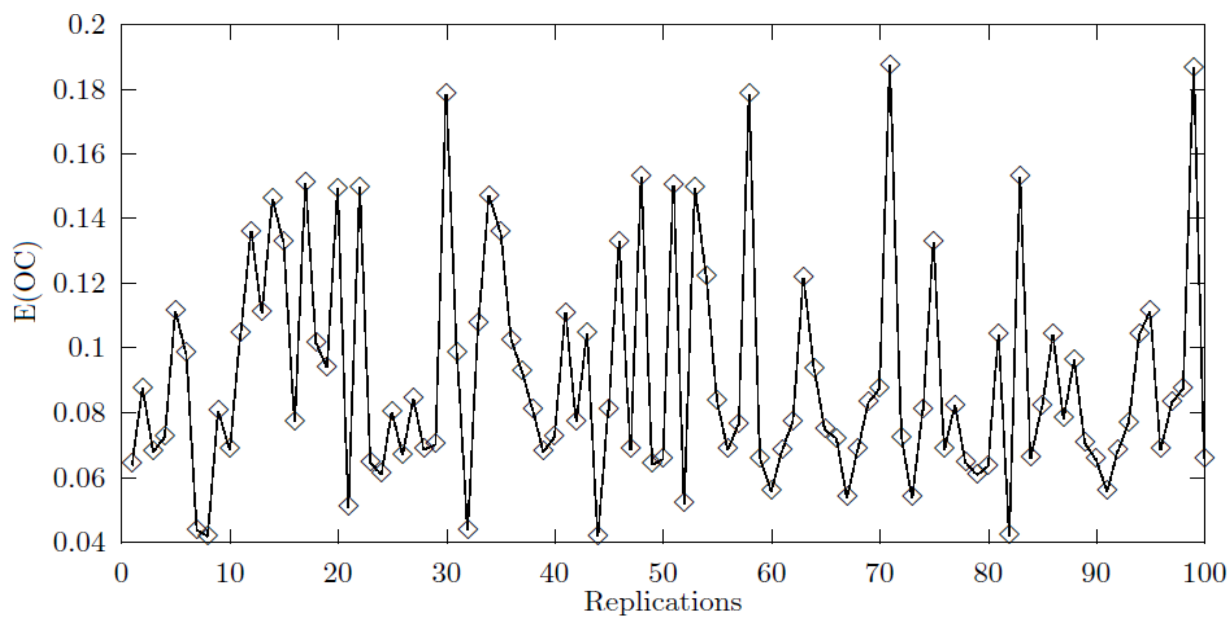

Figure 5. The $E(OC)$ for the selection approach when $n = 10626, g = 100, m\% = 3\%$ over 100 replications.

a 100 replication. It is clear that for both buffer spaces, the $E(OC)$ values for the selection approach are small, showing that the estimated production rate for the selected designs by using the Almomani and Abdul Rahman (2012) approach is closed to the production rate for the best design.

Conclusion

In this paper, one of the most difficult problems in the performance modeling, which is the buffer allocation problem, was solved for a finite production line. The difficulty of this problem is due to the difficulty to find the exact value of the objective function and it involves a large number of alternatives. Using a selection approach proposed by Almomani and Abdul Rahman (2012), the problem of buffer allocation was solved. The main idea was to decrease the number of the competing alternatives by using the ordinal optimization method to make it appropriate for the cardinal optimization methods. The advantage of the selection approach is that, it can be used to select the best buffer profile from a huge number of alternatives. It consists of four stages. Initially, the OO procedure was used to select randomly a subset G from a feasible solution set that overlaps with the set that contains the actual best $m\%$ designs with high probability. Then the OCBA procedure was used to allocate the available computing budget. This was followed by the SS procedure to get a smaller subset I with high probability, which contains the best design among the previous selected subset. Finally, the IZ procedure was applied to select the best design from that set I. Numerical illustrations have demonstrated that the selection algorithm finds optimal (near optimal) allocations in a short, unbalanced and reliable production lines. It also shown that the algorithm was capable to allocate a design with maximum mean production rate by using a relatively small simulation samples, and at the same time with a minimum expected opportunity cost and high probability of correct selection.

ACKNOWLEDGMENT

The researchers would like to thank School of Mathematical Sciences, USM and the USM fellowship scheme for financial support. This work is partly sponsored by USM PRGS Grant 1001/PMATHS/ 844038 and USM Short Term Grant 304/PMATHS/639045.

REFERENCES

Almomani MH, Abdul Rahman R (2012). Selecting a good stochastic system for the large number of alternatives. Commun. Stat.-Simul. Comput., 41(2): 222-237.

Alon G, Kroese DP, Raviv T, Rubinstein RY (2005). Application of the cross-entropy method to the buffer allocation problem in a simulation-based environment. Ann. Oper. Res., 134: 137-151.

Alrefaei MH, Almomani MH (2007). Subset selection of best simulated systems. J. Frankl. Inst.-Eng. Appl. Math., 344: 495-506.

Alrefaei M, Andradóttir S (2005). Discrete stochastic optimization using variants of the stochastic ruler method. Nav. Res. Logist., 52(4): 344-360.

Banks J (1998). *Handbook of simulation*. John Wiley, pp. 335-393.

Bechhofer RE, Santner TJ, Goldsman DM (1995). *Design and analysis of experiments for statistical selection, screening, and multiple comparisons*. Wiley: New York, xii+325 pp. ISBN: 0-471-57427-9

Chaharsooghi SK, Nahavandi N (2003). Buffer allocation problem, a heuristic approach. Sci. Iran. 10(4): 401-409.

Chen CH, Wu CD, Dai L (1999). Ordinal comparison of heuristic algorithms using stochastic optimization. *IEEE Trans. Robotics Automation,* 15(1): 44-56.

Chen CH, Yücesan E, Chick SE (2000). Simulation budget allocation for further enhancing the efficiency of ordinal optimization. Discret. Event Dyn. Syst.-Theory Appl., 10(3): 251-270.

Daskalaki S, Smith JM (2004). Combining routing and buffer allocation problems in series-parallel queuing networks. Ann. Oper. Res., 125: 47-68.

Deng M, Ho YC (1999). An ordinal optimization approach to optimal control problems. Automatica, 35: 331-338.

Diamantidis AC, Papadopoloulos CT (2004). A dynamic programming algorithm for the buffer allocation problem in homogeneous asymptotically reliable serial production lines. Math. Probl. Eng., 3: 209-223.

Foley RD, Park BC (2002). Optimal allocation of buffers and customers in a two-node cyclic network with multiple servers. Oper. Res. Lett., 30: 19-24.

Fu MC, Glover FW, April J (2005). Simulation optimization: A review, new developments, and applications. Proceedings of the 2005 winter simulation conference, Institute of Electrical and Electronics Engineers, Piscataway, New Jersey, pp. 83-95.

Gershwin SB, Schor JE (2000). Efficient algorithms for buffer space allocation. Ann. Oper. Res., 93: 117-144.

Gupta SS (1965). On some multiple decision (selection and ranking) rules. Technometrics, 7(2): 225-245.

He D, Chick SE, Chen CH (2007). Opportunity cost and OCBA selection procedures in ordinal optimization for a fixed number of alternative systems. *IEEE Trans. Syst.,* 37: 951-961.

Hedayati A, Feshaaraki MN, Badie K, Aghazarian V (2011). Bahar: a new hybrid of GA and auction method for dynamic bandwidth allocation based on EPON networks. Int. J. Phys. Sci., 6(6): 1342-135.

Ho YC, Sreenivas RS, Vakili P (1992). Ordinal optimization of DEDS. Discret. Event. Dyn. Syst.-Theory Appl., 2: 61-88.

Ho YC, Zhao QC, Jia QS (2007). *Ordinal optimization: soft optimization for hard problems*. Springer, p. 317.

Huang MG, Chang PL, Chou YC (2002). Buffer allocation in flow-shop-type productions systems with general arrival and service patterns. Comput. Oper. Res., 29: 103-121.

Kim DS, Kulkarni DM, Lin F (2002). An upper bound for carriers in a three-workstation closed serial production system operating under production blocking. IEEE Trans. Autom. Control, 47(7): 1134-1138.

Kim SH, Nelson BL (2001). A fully sequential procedure for indifference-zone selection in simulation. ACM Trans. Model. Comput. Simul., 11: 251-273.

Kim SH, Nelson BL (2006a). Selecting the best system. Handbooks in Operations Research and Management Science. Chapter, 17: 501-534.

Kim SH, Nelson BL (2006b). On the asymptotic validity of fully sequential selection procedures for steady-state simulation. Oper. Res., 54(3): 475-488.

Kim SH, Nelson BL (2007). Recent advances in ranking and selection. *Proceedings of the 2007 Winter Simulation Conference, pp.* 162-172.

Lee LH, Lau TWE, Ho YC (1999). Explanation of goal softening in ordinal optimization. IEEE Trans. Autom. Control, 44: 94-99.

Li D, Lee LH, Ho YC (2002). Constraint ordinal optimization. Inf. Sci., 148: 201-220.

Lutz CM, Davis KR, Sun M (1998). Determining buffer location and size in production lines using tabu search. Eur. J. Oper. Res., 106: 301-316.
Malekian R, Abdullah AH (2011). A mathematical model to determine the maximum end-to-end delay bound on label switched path for real time applications over mobile IPv6. Int. J. Phys. Sci., 6(12): 2958-2964.
Nelson BL, Swann J, Goldsman D, Song W (2001). Simple procedures for selecting the best simulated system when the number of alternatives is large. Oper. Res., 49(6): 950-963.
Papadopoulos CT, O'kelly MEJ, Vidalis MJ, Spinellis D (2009). Analysis and design of discrete part production lines. Springer, p. 279.
Rinott Y (1978). On two-stage selection procedures and related probability-inequalities. Commun. Stat.-Theory Methods, A7: 799-811.
Roser C, Nakano M, Tanaka M (2003). Buffer allocation model based on a single simulation. Proceedings *of the 2003 winter simulation conference:* 1238-1246.
Spinellis DD, Papadopoulos CT (2000). A simulated annealing approach for buffer allocation in reliable production lines. Ann. Oper. Res., 93: 373-384.
Wilcox RR (1984). A table for Rinott's selection procedure. J. Qual. Technol., 16(2): 97-100.
Yuzukirmizi M, Smith JM (2008). Optimal buffer allocation in finite closed networks with multiple servers. Comput. Oper. Res., 35: 2579-2598.
Zhao QC, Ho YC, Jia QS (2005). Vector ordinal optimization. J. Optim. Theory Appl., 125: 259-274.