Two Sequential Algorithms for Selecting one of the Best Simulated Systems

MAHMOUD H. ALREFAEI & HOSSAM ABDUL-RAHMAN

Department of Mathematics and Statistics
Jordan University of Science & Technology
P.O.Box 3030, Irbid, 22110, JORDAN

Abstract: We consider the problem of selecting one of the best simulated systems when the number of alternatives is large. We propose two sequential algorithms for selecting a good enough simulated system based on the idea of ordinal optimization that focuses on ordinal rather the cardinal of the competent systems. In the first algorithm, we use the idea of ordinal optimization together with the Ranking and Selection procedure. In the second algorithm, we use the ordinal optimization with the optimal computing budget allocation algorithm. Numerical experiments for comparing these algorithms are presented.

Key-Words:- Ordinal Optimization, Ranking and Selection, Optimal Computing Budget Allocation, Stochastic Optimization

1. Introduction

In this paper, we consider the problem of selecting the best simulated system (the expected performance of some stochastic system). This problem is described as follows:

$$\min_{\theta \in \Theta} J(\theta) = E(L(\theta, \xi)) \tag{1}$$

where Θ, the search space, is an arbitrary, huge, has no structure but finite set, θ a design alternative, J, the performance criterion which is the expectation of L, the sample performance, as a functional of θ and ξ, the randomness of the systems.

We are interested in finding a good enough design, so a good approach is to use the idea of Ordinal Optimization (OO) which allows us to reduce sampling, because OO focuses on finding good enough designs rather than trying to find the best design, see Ho et. al. [3] and Ho [4]. When the search space is small, less than 20, say, then Ranking and Selection (R&S) can be applied for selecting the best design among k different designs with a prespecified significance level see for example Rinott [7] for a two stage R&S procedure. R&S cannot be applied for large scale problems because it needs a huge computational time. Recently, the Optimal Computing Budget Allocation (OCBA) has been proposed by Chen et. al. [2] to optimally distribute the available computational budget among the competent designs to maximize the probability of correct selection. In this paper, we propose two sequential algorithms for selecting a design from the good enough designs. The first algorithm uses the idea of OO and the R&S procedure. The second one uses the idea of OO with OCBA.

The rest of the paper is organized as follows, in Section 2 we review the basic ideas of OO, in Section 3 we review the R&S procedure for selecting the best design among a small set; next, in section 4 we review the OCBA. In Section 5 we present the proposed algorithms. In Section 6 we give some numerical experiments about the convergence of both algorithms, and we make comparisons between the two proposed algorithms. Finally, in Section 7, we give some concluding remarks.

2. Ordinal optimization procedure

The Ordinal Optimization (OO) has emerged an efficient technique for simulation and optimization. We consider the problem (1), if our goal is to find a good enough design rather than to find the best design then it is advantageous to use the OO. OO could significantly reduce the computational cost for discrete event stochastic systems simulation. The basic idea is to select a subset of the search space and the Correct Selection (CS) here is to select a subset that contains one of the top $m\%$ best designs rather than doing extensive simulation to select the best design, see Ho et. al. [3]. Suppose we randomly select a subset of g designs from the search space Θ, then the probability that this subset contains at least one of the top $m\%$ best designs is given by

$$P(CS) \approx \left(1 - (1 - \frac{m}{100})^g\right).$$

It is clear that this $P(CS)$ increases when the sample size g increases. In the literature $P(CS)$ is denoted as the alignment probability.

3. Ranking and selection procedures

Ranking and Selection (R&S) procedures are used to make comparisons between several alternative systems. There are three types of R&S procedures. The first procedure is called the indifference zone approach and can be used for selecting a design that is indifferent from the actual best design by less than a pre specified value d^* called the indifference zone with a pre specified significance level. The second procedure can be used for selecting a subset of m systems from Θ so that this selected subset contains the best system with high probability. The third procedure is used for selecting a subset that contains the best m systems from Θ with a specified probability for more details see Law and Kelton [5].

In this paper, we are interested in the indifference zone approach that can be applied for selecting the best from several systems. The objective of this procedure is to select the best design of k alternatives. For $i = 1, 2, \ldots, k$ let $\xi_{i1}, \xi_{i2}, \ldots, \xi_{in_i}$ be a sample of n_i i.i.d observations from system i, let $\mu_i = E(L(\theta_i, \xi_{ij}))$, and let μ_{i_l} be the l^{th} smallest of $\mu_i^{'}s$, so that $\mu_{i_1} \leq \mu_{i_2} \leq \ldots \leq \mu_{i_k}$. The correct selection (CS) here is to select a system with the smallest expected mean μ_{i1} (if we want the largest mean μ_{ik}, the signs of the X_{ij}'s and μ_i's can simply be reversed). Our aim is to maximize the probability of correct selection i.e., $P(CS) \geq p^*$ provided that $\mu_{i2} - \mu_{i1} \geq d^*$ where the indifference zone $d^* > 0$ (i.e., selecting a design that is within d^* from the the actual best design).

Usually, the indifference zone approach consists of two stages. In the first stage, all designs are sampled using n_0 simulation runs to get an initial estimate of $f(\theta), \forall \theta \in \Theta$ and their variances. Depending on the information obtained in the first stage, we compute how many more samples are needed in the second stage for each design that guarantee the probability of correct selection is within a pre specified significance level. This huge amount of computational time restrict the size of the problems that these approaches can solve, in fact they can solve a small size problems, say less than 20. Nelson et. al. [6] and Alrefaei and Alawneh [1] propose using subset selection in the first stage to reduce the competent designs. These designs are carried out to the second stage in which the indifference zone approach is used where the information in the first stage are used.

4. The optimal computing budget allocation idea

Consider the optimization problem given by equation (1) and suppose that we are allowed to make T samples from all the design space. Our interest here is where to allocate these designs in order to maximize the probability of correct selection, this adds a new constraint to the optimization problem (1) and it becomes:

$\min J(\theta_i) = E(L(\theta_i, \xi_i))$

such that $\sum_{i=1}^{k} N_i \leq T$

Suppose we can make i.i.d samples $\xi_{ij}, j = 1, \ldots, N_i$ from each design i. Let the estimate of $J(\theta_i)$ be defined as $\overline{J}_i = \frac{1}{N_i} \sum_{j=1}^{N_i} L(\theta_i, \xi_{ij})$. Let b be the index of the design that minimizes the approximated function $\overline{J}$. Then the following theorem by Chen [2] gives the relation between N_i and N_j for $i \neq j \neq b, i, j = 1, \ldots, k$ and between N_b and N_i that guarantees the maximization of $P(CS)$.

Theorem 1 *Given a total number of simulation samples T to be allocated on k computing designs whose performance is depicted by random variables with means $J(\theta_1), J(\theta_2), ..., J(\theta_k)$, and finite variances $\sigma_1^2, \sigma_2^2, ..., \sigma_k^2$ respectively, as $T \rightarrow \infty$ the Approximate Probability of Correct Selection (APCS) can be asymptotically maximized when*

1. $\frac{N_i}{N_j} = (\frac{\sigma_i/\delta_{bi}}{\sigma_j/\delta_{bi}})^2, \quad i, j \in \{1, 2, ..., k\}$, *and* $i \neq j \neq b$
2. $N_b = \sigma_b \sqrt{\sum_{i=1, i\neq b}^{k} \frac{N_i^2}{\sigma_i^2}}$

where N_i is the number of samples allocated to design i, $\delta_{bi} = \overline{J}_b - \overline{J}_i$ and $\overline{J}_b \leq \min \overline{J}_i$

5. Two sequential algorithms

Consider the optimization problem given by equation (1) and suppose that the search space Θ is huge and our goal is to select a design that belongs to the top $m\%$ best designs, then an efficient approach is to use the idea of OO. We propose two sequential algorithms; the first one combines the idea of OO with the R&S, and the second one combines the idea of OO with OCBA. Each iteration of these two algorithms is based on a two stage procedure. In the first stage, we randomly select a subset G, where $|G| = g$ from the search space Θ. In the second stage of the first algorithm, R&S is used to select the best design in the selected subset G with a specified significance level. In the second algorithm the idea of OCBA is used in the second stage to select the best design θ_b among the designs selected in the first stage. θ_b will be carried out to the next iteration and the other $g - 1$ designs will be replaced by newly sampled designs from Θ. The two algorithms continue in this manner until a specified number of generated samples is achieved. Now, we present the first Algorithm

Algorithm 1 (OO with R&S)

Parameters $G, g, n_0, d^*, P^* = 1 - \alpha, h_1$.

step 0 Select a subset G randomly from Θ, such that $|G| = g$

Step1 Generate n_0 i.i.d samples $\xi_{ij}, i = 1, 2, ..., g, j = 1, 2, ..., n_0$, for all $i = 1, \ldots, g$, compute

$$\overline{J}_i^{(1)}(n_0) = \frac{\sum_{j=1}^{n_0} L(\theta_i, \xi_{ij})}{n_0},$$

$$S_i^2(n_0) = \frac{\sum_{j=1}^{n_0} [L(\theta_i, \xi_{ij}) - \overline{J}_i^{(1)}(n_0)]^2}{n_0 - 1},$$

$$\text{and } N_i = \max\left\{ n_0 + 1, \left\lceil \frac{h_1^2 S_i^2(n_0)}{(d^*)^2} \right\rceil \right\},$$

where $\lceil u \rceil$ is the smallest integer greater than or equal u, d^* is the indifference zone, and h_1 depends on g, n_0, and $(1 - \alpha)$, and can be obtained from tables of Wilcox [8] or Law and Kelton [5]

Step 2 For $i = 1, \ldots, g$ compute

$$\overline{J}_i^{(2)}(N_i - n_0) = \frac{\sum_{j=n_0+1}^{N_i} L(\theta_i, \xi_{ij})}{N_i - n_0}$$

$$\widetilde{J}_i(N_i) = W_{i1}\overline{J}_i^{(1)}(n_0) + W_{i2}\overline{J}_i^{(2)}(N_i - n_0)$$

where the weights W_{i1} and W_{i2} are given by:

$$W_{i1} = \frac{n_0}{N_i}\left[1 + SQ\right]$$

where

$$SQ = \sqrt{1 - \frac{N_i}{n_0}\left(1 - \frac{(N_i - n_0)(d^*)^2}{h_1^2 S_i^2(n_0)}\right)}$$

and $W_{i2} = 1 - W_{i1}$. (See Chapter 10 of Law and Kelton [5]).

Step 3 Find

$$\theta_b = \text{argmin}_{i=1,2,\ldots,g}\widetilde{J}_i(N_i)$$

If a stopping criterion is not met then select a set G' randomly from $\Theta - \theta_b$, where $|G'| = g - 1$ and let $G = G' \cup \{\theta_b\}$ and go to Step 1. Otherwise, stop and return θ_b as the best design.

Note that in the first stage, the probability that the randomly selected subset G contains one of the best $m\%$ designs can be evaluated as follows $P(G$ contains one of the best $m\%$ designs$) \approx 1 - (1 - \frac{m}{100})^g$. Let b_i be the the selected best design in iteration i of the algorithm, and let G'_i be the selected subset in iteration i, where $|G'_i| = g - 1$. Let $P^* = 1 - \alpha$ be the significance level used in the R&S procedure (the second stage of the algorithm). Then the probability of selecting one of the top $m\%$ best designs in the first iteration is

$$P_1 \geq (1 - \alpha)\left(1 - (1 - \frac{m}{100})^g\right)$$

and the probability of correct selection in the j^{th} iteration is given by $P_j = P(b_{j-1}$ is in the top $m\%$ best designs or G'_j contains one of the top $m\%$ best designs) $\times(1 - \alpha)$. Which equals to $P_j = (P(b_{j-1}$ is in the top $m\%$ best designs) $+\ P(G'_j$ contains one of the top $m\%$ best designs) $-P(b_{j-1}$ is in the top $m\%$ best designs) $\times\ P(G_j$ contains one of the top $m\%$ best designs) $)\times(1 - \alpha)$.

As an example, consider $m\% = 10\%$ and $1-\alpha = 0.95$ then $P_1 \geq 0.619, P_2 \geq 0.808, P_3 \geq 0.925$. This means that in three iterations, we can select a design that belongs to the top 10% best designs with probability greater than or equal to 0.925 regardless the size of the optimization problem given in equation (1).

We know that OCBA maximizes $P(CS)$ if the computing budget is distributed as in Theorem 1. Therefore, if we replace the R&S procedure in the second stage with the OCBA procedure, then we expect that the new algorithm gives better performance as we will see in the numerical example below. The modified algorithm is described in Algorithm 2 below. Each iteration of the algorithm consists of sub iterations, these sub iterations will be repeated until the designated subtotal computing budget is used. Note that the total computing budget is TT. TT will be divided into m iterations. Let T be the total computing budget in each iteration, and let G be the selected subset from the search space Θ. Let n_0 be the number of initial number of samples for each design $\theta \in G$ and Δ be the number of additional samples for each sub iteration that need to be allocated into members of G then the modified algorithm will be described as follows:

Algorithm 2 (OO with OCBA)

Parameters T, TT, n_0, Δ, g

step 0 Select a subset G randomly from Θ, such that $|G| = g$. Let $TS = 0$.

Step 1 For each $i = 1, 2, ..., g$, generate i.i.d samples $\xi_{ij}, j = 1, 2, ..., n_0$. Let l = 1 and let $N_1^l = N_2^l = \ldots = N_g^l = n_0$.

Step 3 Compute

$$\overline{J}_i = \frac{1}{N_i^l} \sum_{j=1}^{N_i^l} L(\theta_i, \xi_{ij}), i = 1, \ldots, g.$$

Let $\theta_b = \text{argmin } \overline{J}_i$.

If $\sum_{i=1}^{g} N_i^l \geq T$ then go to Step 6.

Step 4 Increase the computational budget by Δ and compute $N_1^{l+1}, \ldots, N_g^{l+1}$ using Theorem 1

Step 5 Perform $\max\{N_i - n_0, 0\}$ additional samples for each design i, $i = 1, \ldots, g$. Let $l = l + 1$ and go to Step 3

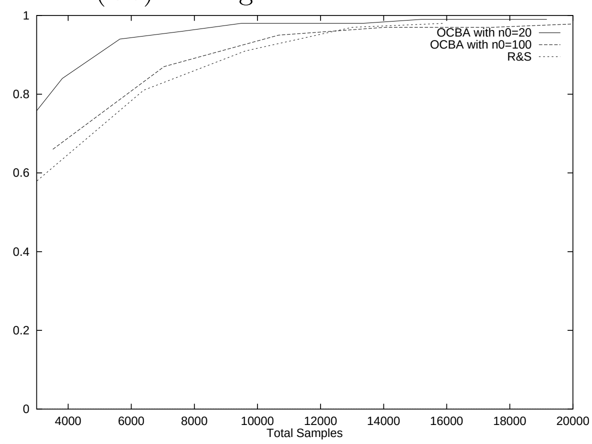
Step 6 Let $TS = TS + \sum_{i=1}^{g} N_i^l$. If $TS \geq TT$ then stop, otherwise, select a set G' randomly from $\Theta - \{\theta_b\}$, where $|G'| = g - 1$. Let $G = G' \cup \{\theta_b\}$ and go to Step 1. Otherwise, stop and return θ_b as the best design.

6. Numerical Experiments

We test our algorithms in the following generic example. Suppose we have 200 alternative designs. where $J(\theta_i) = \frac{i}{100}, i = 1, \ldots, 200$, and $L(\theta_i, Y) \sim N(J(\theta_i, 1))$. The optimization problem is $\min_{i=1,\ldots,200} J(\theta_i)$. The objective is to select a design that belongs to the best 10% good designs. It is obvious that designs $\theta_1, \theta_2, \ldots, \theta_{20}$ are the actual best 10% designs.

We apply algorithm 1 where we use the R&S in each iteration with significance level $p^* = 0.95$, the initial sample size $n_0 = 20$, and the indifference zone $d^* = 0.2$. We also apply algorithm 2 where we use OCBA with a maximum computing budget of $TT = 20,000$ samples, and the selected set G consists of 10 designs in each iteration with two different values of n_0, $n_0 = 20$ and $n_0 = 100$. The results are depicted in Figure (1). The results show that it is better to use small initial sample size in the OCBA approach. It is also clear that P(CS) converges very quickly to 1 in both algorithms. However, the OCBA is much faster than R$S. We have also applied both algorithms for solving a larger size problem where the search space Θ contains 1,000 different designs. Algorithm 1 is applied in which we use two values of n_0, $n_0 = 20$, and $n_0 = 40$. However, Algorithm 2 is applied using $n_0 = 20$ since it works better in the first experiment. The results are depicted in Figure 2. It is clear that in both algorithms the probability of correct selection converge to one. It is also clear that the performance of Algorithm 2 is much better than the performance of Algorithm 1.

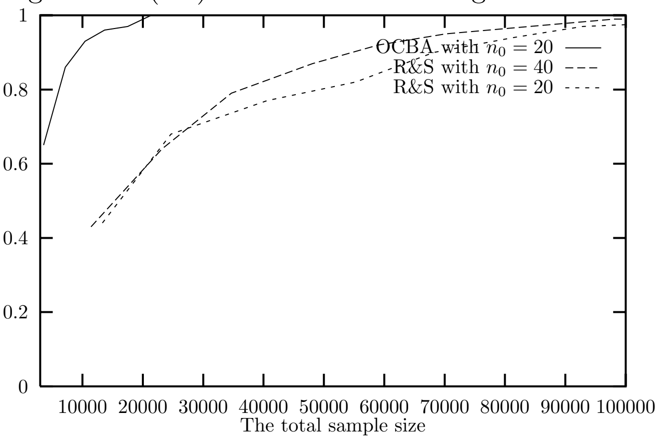
Figure 1. P(CS) for Algorithms 1 and 2.



7. Conclution

In this paper, we have proposed two new sequential algorithms for solving discrete simulation optimization problems. The objective in the optimization problems is to select a design that belongs to the top m% best designs. These algorithms are based on merging the idea of ordinal optimization together with the ranking and selection procedure in the first algorithm. In the second algorithm we use the optimal computing budget allocation procedure. The probability of correct selection P(CS) for the first algorithm is derived.

Figure 2. P(CS) when $\Theta = 1000$ designs.



However, the value of P(CS) for the second algorithm cannot be derived analytically, but we allocate the samples in a way that it maximizes P(CS), so it performs better than Algorithm 2. The numerical results indicate that these algorithms converge very quickly (the probability of correct selection converge to one). It is clear that Algorithm 2 works better than Algorithm 1. We see that algorithm 2 may be more effective when the design space is bigger.

REFERENCES

1. Alrefaei, M. H. and A. J. Alawneh, Selecting the best stochastic system for large scale problems in DEDS. Mathematics and Computers in Simulation. Vol.64, 2004 pp.237-245.
2. Chen, C. H., E. Yucesan, and S. E. Chick "Simulation Budget Allocation for Further Enhancing the Efficiency of Ordinal Optimization".*Discrete Event Dynamic Systems: Theory and Application* Vol.10, 2000, pp.251–270.
3. Ho, Y. C., R. S. Sreenivas, and P.Vakili, Ordinal optimization of DEDS. *Journal of discreate Evente dynamic System* Vol.2, 1992, pp.61-88.
4. Ho, Y. C. An explanation of Ordinal Optimization: Soft Computing for Hard Problems.*Information Sciences* Vol.113, 1999 pp.169-192.
5. Law, A. M. and W. D. Kelton. *Simulation Modeling and Analysis* Third Edition. McGraw-Hill, New York. 2000.
6. Nelson B. L. , J. Swann, D. Goldsman, and W. Song, Simple Procedures for Selecting the Best Simulated System when the Number of Alternatives is Large, *Operations Research*, Vol.49 2001 pp.950–963.
7. Rinott, Y. On two–stage selection procedures and related probability inequalities. *Communications in Statistics: Theory and Methods*, Vol.A7, 1978, pp.799–811.
8. Wilcox, R. R., A table for Rinott's selection procedure. *Journal of Quality Technology*, Vol.16, No.2, 1984, pp.97–100.